Ann Ist Super Sanità 2020 | Vol. 56, No. 1: 122-127
DOI: 10.4415/ANN_20_01_17

BRIEF NOTE

Identification of two novel LDLR variants by Next Generation Sequencing

Simona Moffa[1,2], Giorgia Mazzuccato[3], Maria De Bonis[3], Elisa De Paolis[3], Maria Elisabetta Onori[3], Alfredo Pontecorvi[1,2], Andrea Urbani[3,4], Andrea Giaccari[1,2], Ettore Capoluongo[5] and Angelo Minucci[3]

[1]*Endocrinologia e Diabetologia, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy*
[2]*Istituto Patologia Speciale Medica, Università Cattolica del Sacro Cuore, Rome, Italy*
[3]*Unità di Diagnostica Molecolare e Genomica, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy*
[4]*Istituto di Biochimica e Biochimica Clinica, Università Cattolica del Sacro Cuore, Rome, Italy*
[5]*Università Federico II-CEINGE, Biotecnologie Avanzate, Naples, Italy*

Abstract

Introduction. Familial hypercholesterolemia (FH) is an autosomal dominant inherited disease characterized by elevated plasma low-density lipoprotein cholesterol (LDL-C). Targeted Next Generation Sequencing (NGS) is a new opportunity to expand the existing pathogenic variants (PVs) spectrum associated to FH. Our aim was to report a diagnostic NGS-based approach to detect variants associated to FH.

Methods. We report two patients: a 48-year-old Asian woman, without known history of hypercholesterolemia and a 46-year-old Caucasian man, with childhood hypercholesterolemia.

Results. An effective NGS-based pipeline, *FH-Devyser kit/Amplicon Suite*, beginning from sequencing to data analysis, did not identify known PVs in the *LDLR*, *APOB*, *APOE, LDLRAP1*, *STAP1* and *PCSK9* genes, but revealed two *novel LDLR* variants (*c.1564A>T*, p.Ile522Phe and *c.1688C>T*, p.Pro563Leu).

Discussion and conclusions. This study showed that an effective NGS-based pipeline led to a definitive diagnosis in two FH families, allowing to plan their therapeutic treatment. Although the functional consequence of the two *LDLR* variants needs to be assessed *in vitro*, the *in silico* analysis and high preservation of the two amino acid positions observed in the LDLR protein, across different animal species, suggest that both variants are deleterious.

Key words

- next generation sequencing
- LDL-cholesterol
- FH-Devyser Kit
- SmartSeq
- *novel LDLR* variants

INTRODUCTION

Three canonical genes underlie autosomal dominant familial hypercholesterolemia (ADFH): low-density lipoprotein receptor (LDLR) (FH-type 1), apolipoprotein B (APOB) (FH-type 2) and pro-protein convertase subtilisin/kexin 9 (PCSK9) (FH-type 3). Other genes, like signal transducing adaptor family member 1 (STAP1) (FH-type 4) and apolipoprotein E (APOE) (FH-type 5) have been associated to some rare FH forms [1]. In addition, FH-like phenotypes, with a recessive transmission, are extremely rare. This group of disorders includes the classic autosomal recessive hypercholesterolemia (ARH) caused by mutations in the Low Density Lipoprotein Receptor Adaptor Protein 1 (*LDLRAP1*) gene [2].

Successful molecular diagnosis depends on the ability of the designated method to identify molecular lesion associated with FH [3]. Next Generation Sequencing (NGS) allows molecular diagnostic laboratories to increase sample throughput, to reduce the turn-around time and to analyse more disease-related genes simultaneously. As a result, recent studies demonstrated that NGS-based assays are very useful and at the same time cost effective for the genetic diagnosis of FH in primary care [4]. However, the main challenge to translate NGS in clinical practice is to develop a simple and robust diagnostics and bioinformatics pipeline, fulfilling the quality control requirements for clinical diagnosis.

We present an effective NGS-based pipeline, from sequencing to data analysis, which identified two unreported *LDLR* variants in two FH families.

CASES' PRESENTATION

Case A

Index case A is a 48-year-old Asian woman, without known history of hypercholesterolemia. During a rou-

Address for correspondence: Angelo Minucci, Molecular and Genomic Diagnostics Laboratory, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Largo Agostino Gemelli 8, 00168 Rome, Italy. E-mail: angelo.minucci@policlinicogemelli.it.

tine blood examination, high LDL-C levels (324 mg/dl) were revealed *(Table 1)*. Deeper insights on her family history disclosed two already deceased brothers from heart attack at the age of 48 and 54, respectively. Her parents' medical history remained unclear. No tendon xanthoma or corneal arcus were diagnosed during physical examination. The patient had no other cardio-vascular (CV) risk factors as smoking, diabetes and CHD history. Her body mass index was 21 kg/m^{2}and she was normotensive.

The patient's Dutch Lipid Clinic Network diagnostic criteria for FH score was 10, indicating a definite FH diagnosis. Therefore, we performed genetic evaluation, according with EAS/ESC guidelines [5]. Furthermore, her 13-year-old son and 14-year-old daughter, whose lipid profile *(Table 1)* was clearly indicative of FH, underwent genetic testing.

Case B

Index case B is a 46-year-old Caucasian man, suffering from hypercholesterolemia since his childhood, in absence of CV history. The patient complained about muscle pain and cramps due to statin therapy, which was, therefore, suspended for about a month. His blood examination showed high LDL-C levels (313 mg/dl), despite ongoing therapy with ezetimibe and fibrate *(Table 1)*. Family history revealed that his mother suffered from severe hypercholesterolemia, even if no familial CV events were reported. No tendon xanthoma or corneal arcus were diagnosed during physical examination. He did not show additional CV risk factors, such as smoking and diabetes, while he presented with a normal BMI of 22 kg/m^2 and normotensive status.

Considering that his Dutch Score was 9, which is related to a definite diagnosis of FH, according with EAS/ESC guidelines, the genetic evaluation was performed.

The two patients gave their written informed consent prior to blood sampling. Genomic DNA was isolated from peripheral blood by an automatic device (MagCore HF16 Plus, Diatech Lab Line, Jesi, Italy). We used 30 ng input DNA for the analysis with the NGS-based method (Devyser FH kit, Devyser, Hägersten, Sweden). All exons, the exon–intron boundaries, 5' and 3' untranslated regions of six FH-related genes (*LDLR*, *APOB*, *PCSK9*, *LDLRAP1*, *APOE* and *STAP1)* and polymorphisms, associated with polygenic FH form and statin treatment effects, were examined [6]. The design was based on the human (Hg19) reference genome to generate 200-bp amplicons by using 192 primer pairs designed and mixed in a single tube. The Qubit® 2.0 Fluorometer 8 (Life Technologies) was used to normalize the genomic DNA concentrations to 2 ng/µl. The final libraries were quantified on Qubit®, diluted to a concentration of approximately 7 pM and denatured to ensure the efficient use of the MiSeq Reagent Kit (Illumina, CA, USA).

MATERIALS AND METHODS

NGS pipeline and Sanger sequencing

Each NGS run consisted of eight FH samples loaded on MiSeq Reagent Kit v2 Nano cartridge, using 2 × 150-bp paired-end chemistry, according to manufacturer instructions. MiSeq-generated FASTQ files were downloaded and processed using Amplicon Suite software (SmartSeq s.r.l, Novara, Italy) providing integrated tools for the analysis, visualization and interpretation of NGS data. Proprietary algorithms based on coverage depth and uniformity were applied for copy number variation (CNV) detection in the *LDLR* gene. Sanger sequencing was performed with an Applied Biosystems 3500 Genetic Analyzer (Life Technologies, Carlsbad, CA, USA). Sequence analysis was carried out with the SeqScape Software v2.5. The primers were designed by Primers 3 software (http://primer3.ut.ee/). Their sequences were: forward (F) 5'-ACTGGATCCACAGCAACATCT-3' and reverse (R) 5'-TGGGATTACAGGTGCTTTGAG-3' and (F) 5'-AGCTATTCTCTGTCCTCCCA-3' and (R) 5'-CTTCAGGGAGCAGCTTGG-3', able to amplify and sequence the *LDLR* exons 10 and 11, respectively.

Prediction of variant effects and species sequence alignment

Four different programs were used to predict pathogenicity of the two *LDLR* variants identified in this study: *Polymorphism Phenotyping version 2* (http://genetics.bwh.harvard.edu/pph2/), *Sorting Intolerant From Tolerant* (http://sift.jcvi.org/www/SIFT_enst_submit.html), *Mutation Taster* (http://www.mutationtaster.org) and *Provean* (http://provean.jcvi.org/index.php). In addition, multiple sequence alignment was created using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/). The reference sequence used for LDLR protein was P01130.1 (SwissProt).

Table 1
Lipid profile of the two index cases and their family's members. The values beyond the normal reference ranges are highlighted in bold

Lipid fraction	Case A			Case B	Laboratory reference ranges* (mg/dl)
	Proband	Son	Doughter	Proband	-
Total cholesterol	**406**	**319**	**302**	**402**	< 200
HDL-C	66	50	51	73	> 45
Triglycerides	78	62	72	82	20-170
LDL-C	**324**	**258**	**237**	**313**	< 130
Apo-B	**145**	**169**	**167**	85	55-130

*target therapy ranges were instead defined according to ESC/EAS guidelines [5]

RESULTS

FH testing obtained by NGS did not reveal any known PVs in all genes investigated and no bioinformatics algorithms were suggestive for the presence of CNV in the *LDLR* gene. However, we found two *LDLR* variants (*c.1564A > T*, coverage: 325/643X, allele frequency: 50%, p.Ile522Phe) and *c.1688C > T*, coverage: 219/433X, allele frequency: 50%, p.Pro563Leu) in case A and B, respectively *(Figure 1)*. The nomenclature of the variants is based on the *LDLR* sequence (NCBI Reference Sequence: NM_000527.4; *GRCh37*), according to the recommendations of the Human Genome Variation Society (https://www.hgvs.org) *(Table 2)*. Sanger sequencing confirmed the presence of both variants on a second independent patient sample.

These variants were considered as *novel* since they were neither found in *ClinVar* (https://www.ncbi.nlm.nih.gov/clinvar/), *Ensemble Human Genome Mutation* (https://www.ensembl.org/index.html), *ExAC* (http://exac.broadinstitute.org) nor *1000G* (http://www.internationalgenome.org/1000-genomes-browsers/) browsers. In addition, among more than 200 alleles routinely analysed by NGS-based molecular screening, we did not identify the two variants in other FH or healthy individuals.

Other family members were not testable for the variant p.(Pro563Leu). Instead, we conducted family screening on two hypercholesterolemic subjects (the patient's sons) for the variant p.(Ile522Phe). Both subjects resulted to be carriers of the variant, providing evidence for his co-segregation genotype/phenotype.

In silico analysis of the p.(Ile522Phe) and p.(Pro563Leu) variants, performed with four different prediction tools, suggested that both *LDLR* variants are deleterious *(Table 1)*. In addition, the alignment of the amino acids in the LDLR protein across seven different animal species, ranging from *Homo sapiens* to *Ovis aries (sheep*), showed that two amino acids (in humans: Ile522, and Pro563) are highly preserved *(Figure 2)*.

DISCUSSION

FH diagnosis is usually based on clinical features, such as physical findings of tendon xanthomas or corne-

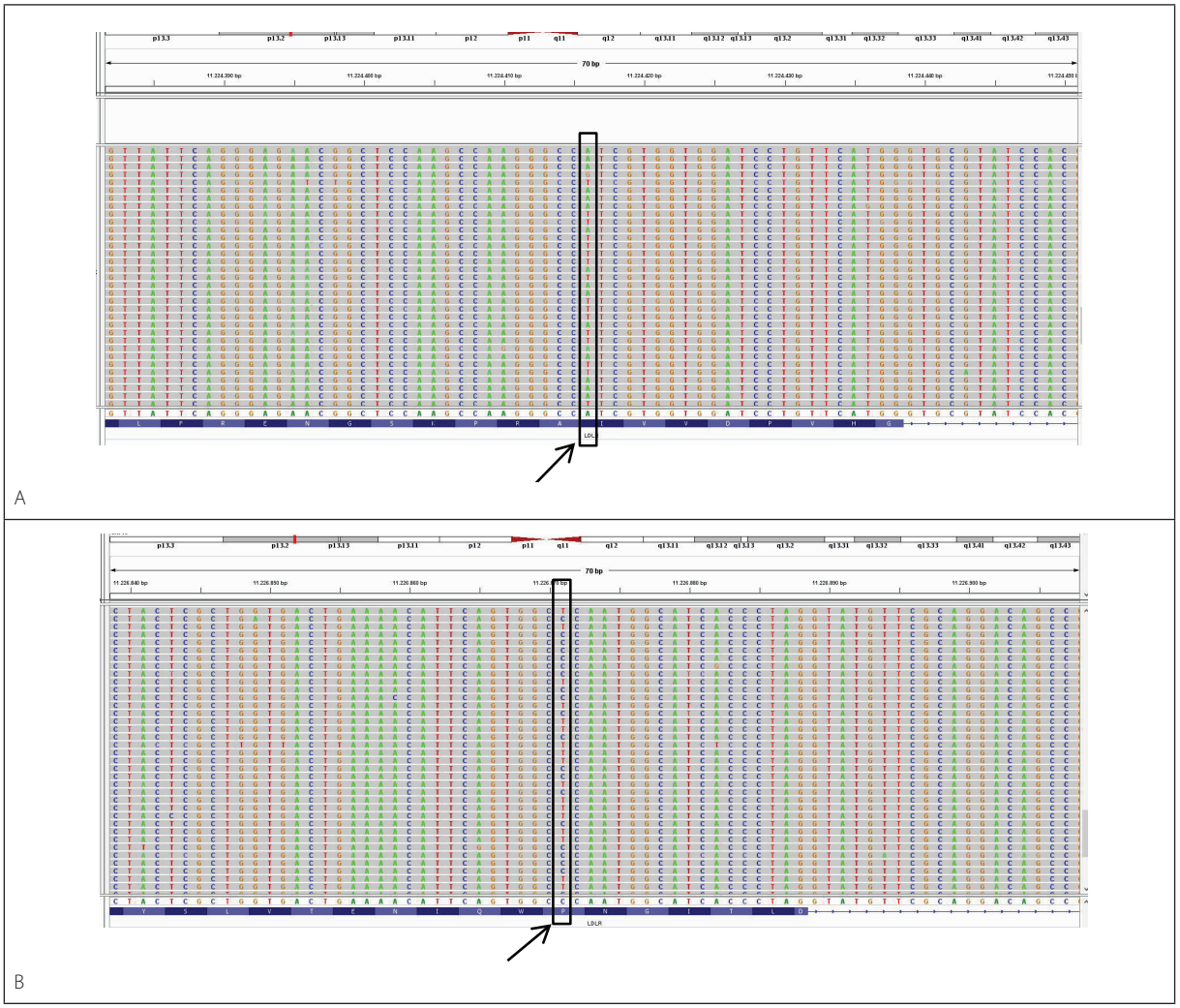

Figure 1
BAM files visualized by the Integrative Genomics Viewer show the reads associated to c.1564A>T, p.(Ile522Phe) (A) and c.1688C>T, p.(Pro563Leu) (B) with a coverage of 325/643X (variant allele frequency: 50%) and 219/433X (variant allele frequency: 50%), respectively. The arrows indicate the nucleotide position of the two LDLR variants.

Table 2
Identities of the two *LDLR* variants and *in silico* prediction of their effect at protein level. Both variants show high scores of pathogenicity for all four software used; following the ACMG guidelines, the *c.1564A>T* and the *c.1688C>T* can be classified as *likely pathogenic* variants

Transcript	Coding impact	HGVS coding	HGVS protein level	Location	Protein position	Splice distance	PolyPhen-2 (range: 0-1)	SIFT (cutoff = 0.05)	Mutation Taster	Provean (cutoff = -2.5)	ACMG classification*
NM_000527.4	M	*c.1564A>T*	I522F p.(Ile522Phe)	exon 10 of 18	206 of 228 (coding)	-23	Probably damaging (score = 0.990)	Damaging (score = 0.001)	Disease causing	Deleterious (score = -3.92)	**Likely pathogenic** (PM1+PM2+ PP1+PP3)
	M	*c.1688C>T*	P563L p.(Pro563Leu)	exon 11 of 18	102 of 119 (coding)	-18	Probably damaging (score = 1)	Damaging (score = 0)	Disease causing	Deleterious (score = -9.39)	**Likely pathogenic** (PM1+PM2+ PM5+PP3)

*Interpretation of variants pathogenicity based on the American the College of Medical Genetics and Genomics (ACMG) recommendations [15], *i.e.*, **PM1**: located in a mutational hot spot and/or in critical functional domain, **PM2**: absent from controls, **PM5**: novel missense change at amino acid residue where a different pathogenic missense change has been seen before, **PP1**: co-segregation with disease in multiple affected family members, **PP3**: multiple lines of computational evidence support a deleterious effect on the gene or gene product.

al arcus, cardio-vascular heart disease history, and high LDL-C concentrations. About 50% of heterozygous FH patients lack obvious phenotypes. Thus, many patients are underdiagnosed until they suffer from acute cardio-vascular events [7]. Previous studies indicated that in many countries less than 1% of FH individuals were correctly diagnosed with FH [8], partly due to the lack of reliable cost-effective genetic testing.

Genetic testing can provide a definitive diagnosis. Possible benefits include individual patient management. For instance, heterozygous FH from CNVs seems to be a more severe phenotype than FH compared to single nucleotide variants. Genetic diagnosis could reduce delays for appropriate treatment, possibly more aggressive LDL-C lowering strategies [9, 10]. Additionally, for some monogenic dyslipidaemias, a genetic diagnosis is needed to ensure funding for newer therapies. This is the case of inhibitors of PCSK9, where, in certain jurisdictions, third party private coverage for the treatment of FH requires a genetic diagnosis [11, 12].

In different FH patient cohorts, PVs in *LDLR* gene are found in more than 85-90% of FH cases and more than 1800 PVs of this gene have been reported and annotated in the UCL database (http://www.ucl.ac.uk/ldlr/LOVDv.1.1.0/)

The heterogeneity of FH-causing variants supports the sequencing-based techniques as the primary methodology to detect the disease and recent advances improved accessibility for diagnostic use. This is the case of the targeted NGS in which only genes specifically involved in FH pathogenesis are screened.

In the present study, targeted sequencing of six genes (*LDLR, APOB, PCSK9, APOE, LDLRAP1* and *STAP1*) performed by an effective NGS-based molecular pipeline (*FH-Devyser Kit/Amplicon Suite software*) allowed to identify two *novel LDLR* variants.

The first, namely *c.1564A>T,* causes substitution of isoleucine for phenylalanine residue (p.Ile522Phe) in the LDL-receptor class B3 domain (486-528aa). The second, *c.1688C>T,* causes substitution of proline for leucine residue (p.Pro563Leu) in the LDL-receptor class B4 domain (529-572 aa). Both variants are located in the very conservative ß-propeller domain, in the YWTD 3 and 4 repeats of the LDLR receptor, respectively [13]. Both protein positions result to be extremely conserved in the species in which they were analysed. For this reason, *in silico* predictions performed with four different tools, using mainly amino acid conservation analysis, suggested that both *LDLR* variants are deleterious. Furthermore, at least in the family with the p.(Ile522Phe) variant, this FH-related mutation segregates with high serum LDL-C levels, linking the variant with the phenotype.

In addition, the p.(Pro563Leu) variant occurs at the same position as other pathogenic missense changes: the *c.1688C>A*, rs879254987, (p.Pro563His) and the *c.1687C*>T, rs879254986, p.(Pro563Ser). Regarding the *c.1564A>T* variant, we underline that it is in the proximity of three missense substitutions: *c.1567G>A,* rs28942080, p.(Val523Met), *c.1567G>T*, rs28942080, p.(Val523Leu) and *c.1561G>A*, rs879254940, p.(Ala521Thr), showing discordant pathogenicity data. We underline that these findings provide further evidence in order to consider p.(Pro563Leu) and (p.Ile522Phe) as deleterious variants.

CONCLUSIONS

In this study, targeted sequencing of six genes (*LDLR, APOB, PCSK9, APOE, LDLRAP1* and *STAP1*) performed by an effective NGS-based molecular approach (*FH-Devyser Kit/Amplicon Suite software)* allowed to identify two *novel LDLR* variants in FH families. Although the functional consequence of the p.(Ile522Phe) and the p.(Pro563Leu) variants remains to be determined *in vitro*, we believe these substitutions should be considered as *"likely pathogenic variants"*.

Finally, we underline that NGS integration with other molecular and cellular techniques, to assess the significance of VUS or *novel* variants, will improve the sensitivity of FH testing entirely, supporting personalized medicine.

Conflict of interest statement

All authors (SM, GM, MDB, EDP, MEO, AP, AU, AG, EC and AM) have read and approved submission

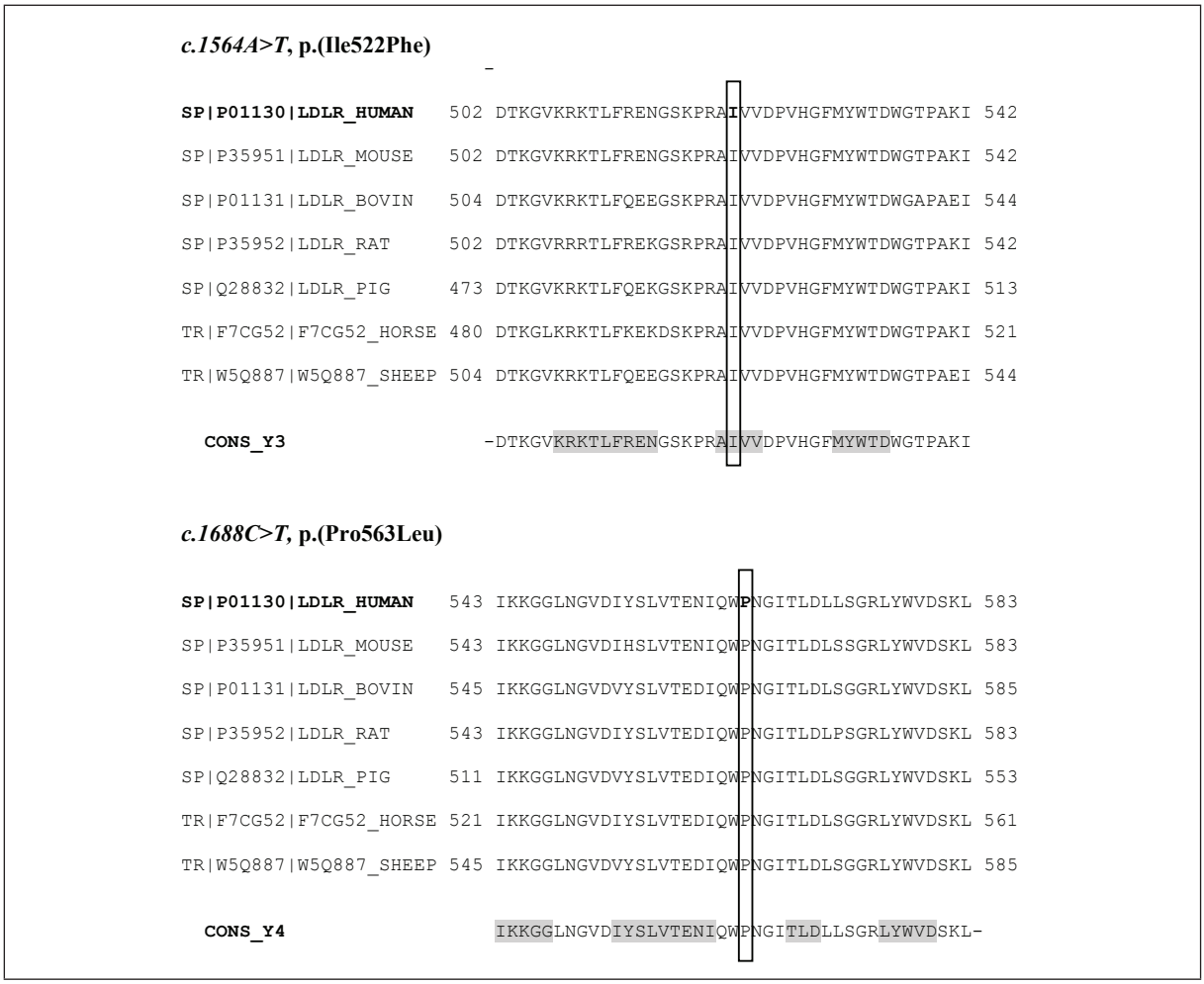

Figure 2
The alignment of the LDLR aminoacids in seven different species and a consensus of 89 YWTD domains (CONS_Y3 and Y4) are shown. The residues predicted to be in b-strand [13,14] are highlighted in grey. The amino acids (SP|P01130|LDLR_HUMAN: Ile522, and Pro563) show high conservation between the species analysed.

of the manuscript. The paper has not been published and is not being considered for publication elsewhere in whole or part in any language. All authors declare that there is no conflict of interests. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgement

We would like to thank Franziska M. Lohmeyer for critically reviewing and editing our manuscript.

Ethical approval

This study was in compliance with the Ethical Principles for Medical Research Involving Human Subjects according to the World Medical Association Declaration of Helsinki and was reported to the Committee of the Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy.

Received on 31 July 2019.
Accepted on 27 November 2019.

REFERENCES

1. Talmud PJ, Futema M, Humphries SE. The genetic architecture of the familial hyperlipidaemia syndromes: rare mutations and common variants in multiple genes. Curr Opin Lipidol. 2014;25(4):274-81. doi: 10.1097/MOL.0000000000000090
2. Paththinige CS, Sirisena ND, Dissanayake V. Genetic determinants of inherited susceptibility to hypercholesterolemia – a comprehensive literature review. Lipids Health Dis. 2017;16(1):103. doi: 10.1186/s12944-017-0488-4
3. Pirillo A, Garlaschelli K, Arca M, Averna M, Bertolini S, Calandra S, Tarugi P, Catapano AL; LIPIGEN Group. Spectrum of mutations in Italian patients with familial hypercholesterolemia: New results from the LIPIGEN study. Atheroscler Suppl. 2017;29:17-24. doi: 10.1016/j.atherosclerosissup.2017.07.002
4. Hinchcliffe M, Le H, Fimmel A, Molloy L, Freeman L, Sullivan D, Trent RJ. Diagnostic validation of a familial hypercholesterolaemia cohort provides a model for using

targeted next generation DNA sequencing in the clinical setting. Pathology. 2014;46(1):60-8. doi: 10.1097/PAT.0000000000000026
5. Catapano AL, Chapman J, Wiklund O, Taskinen MR. The new joint EAS/ESC guidelines for the management of dyslipidaemias. Atherosclerosis. 2011;217(1):1. doi: 10.1016/j.atherosclerosis.2011.06.011
6. Talmud PJ, Shah S, Whittall R, Futema M, Howard P, Cooper JA, Harrison SC, et al. Use of low-density lipoprotein cholesterol gene score to distinguish patients with polygenic and monogenic familial hypercholesterolaemia: a case-control study. Lancet. 2013;381(9874):1293-301. doi: 10.1016/S0140-6736(12)62127-8
7. Bhatnagar D, Morgan J, Siddiq S, Mackness MI, Miller JP, Durrington PN. Outcome of case finding among relatives of patients with known heterozygous familial hypercholesterolaemia. BMJ. 2000;321(7275):1497-500. doi: 10.1136/bmj.321.7275.1497
8. Benito-Vicente A, Uribe KB, Jebari S, Galicia-Garcia U, Ostolaza H, Martin C. Familial hypercholesterolemia: the most frequent cholesterol metabolism disorder caused disease. Int J Mol Sci. 2018;19(11):pii:E3426. doi: 10.3390/ijms19113426
9. Paynter NP, Ridker PM, Chasman DI. Are Genetic Tests for atherosclerosis ready for routine clinical use? Circ Res. 2016;118(4):607-19. doi: 10.1161/CIRCRESAHA.115.306360
10. Di Taranto MD, D'Agostino MN, Fortunato G. Functional characterization of mutant genes associated with autosomal dominant familial hypercholesterolemia: integration and evolution of genetic diagnosis. Nutr Metab Cardiovasc Dis. 2015;25(11):979-87. doi: 10.1016/j.numecd.2015.06.007
11. Farnier M. PCSK9: From discovery to therapeutic applications. Arch Cardiovasc Dis. 2014;107(1):58-66. doi: 10.1016/j.acvd.2013.10.007
12. Di Taranto MD, Benito-Vicente A, Giacobbe C, Uribe KB, Rubba P, Etxebarria A, Guardamagna O, Gentile M, Martín C, Fortunato G. Identification and in vitro characterization of two new PCSK9 Gain of Function variants found in patients with familial hypercholesterolemia. Sci Rep. 2017;7(1):15282. doi: 10.1038/s41598-017-15543-x
13. Springer TA. An extracellular beta-propeller module predicted in lipoprotein and scavenger receptors, tyrosine kinases, epidermal growth factor precursor, and extracellular matrix components. J Mol Biol. 1998;283(4):837-62. doi: 10.1006/jmbi.1998.2115
14. Rost B. PHD: predicting one-dimensional protein structure by profile-based neural networks. Methods Enzymol. 1996;266:525-39. doi: 10.1016/s0076-6879(96)66033-9