Published in final edited form as:

Atherosclerosis. 2010 August ; 211(2): 397–398. doi:10.1016/j.atherosclerosis.2010.04.011.

Identification of a novel *LDLR* mutation (c.261_262invGA, p.Trp87X): Importance of specifying DNA and protein mutations

David Ng¹, Emma Spaulding¹, NIH Intramural Sequencing Center², James C. Mullikin^{1,2}, and Leslie G. Biesecker¹

¹ National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

² National Institutes of Health Intramural Sequencing Center (NISC), National Institutes of Health, Bethesda, Maryland 20892, USA

Abstract

ClinSeq is a large-scale medical sequencing (LSMS) project at the National Institutes of Health (NIH), the goal of which is to pilot the feasibility of using high throughput genome sequencing for clinical research and eventually to improve the delivery of healthcare. In phase one, 1000 participants are being clinically evaluated for cardiovascular phenotypes and DNA is being collected for sequencing of 400 candidate genes to identify genetic variants that may predispose to the early development of atherosclerosis. We report on an individual with familial hypercholesterolemia (OMIM #143890) who has a novel mutation, c.261_262invGA that predicts a premature stop (p.Trp87X) in the *LDLR* gene. Although the p.Trp87X predicted protein mutation has been reported, c.261_262invGA is distinct from mutations reported in prior families and emphasizes the importance of describing mutations at the DNA level. It is important to describe mutations according to the underlying DNA change as multiple nucleotide changes may underlie a single predicted protein change.

ClinSeq^{TM*} is a pilot to study the technical, medical, and genetic counseling issues associated with large-scale medical sequencing and its application to personalized medicine. The *ClinSeq* study was reviewed and approved by the NIH National Human Genome Research Institute and Suburban Hospital (Bethesda, Maryland) Institutional Review Boards. In this initial phase, 1000 participants are being recruited and details of the study design and cohort have been published [1].

Clinical findings were analyzed and correlated with sequencing of candidate genes with a known or suspected role in lipid metabolism, diabetes, hypertension, or obesity [2–4]. Sequence data were generated at the NIH Intramural Sequencing Center (NISC). In the initial phase, NISC used semiautomated Sanger Sequencing on 3730XL instruments (Applied Biosystems) to analyze over 200 candidate genes. In the current phase, ClinSeq participants will have whole exome sequencing on the Illumina/SOLEXA sequencing platform. We have sequenced 251 genes in more than 500 individuals, collecting over 1.36 Gb of sequencing data [1]. The majority of the variants discovered are rare and not previously identified, including 22 nonsense variants not in the dbSNP database

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

*ClinSeqTM is a registered trademark. The trademark symbol is not repeated throughout the paper for simplicity.

(<http://www.ncbi.nlm.nih.gov/projects/SNP/>). Among the variants, we have identified variants of known clinical significance in *LDLR* and *APOB*. We report one case highlighting the importance of describing variants using DNA nomenclature, rather than the predicted protein changes.

The proband was a 65 year-old Hispanic female (fig. 1a, II–VI) with a history of elevated cholesterol (LDL 530 mg/dL prior to therapy), angina, and a possible MI in her mid-30s. She was obese (BMI 31.7 kg/m²) with hypertension. Despite lipid lowering therapy, her fasting LDL (151 mg/dL) was abnormal. She had two sons and a daughter (III-I, III-II, III-III) in their mid-30s with reported elevated cholesterol. Her father (I-I), with a history of elevated cholesterol died at 59 of a MI. She had four older brothers with elevated cholesterol; one had a two-vessel coronary artery bypass graft (CABG) at 68 (II-III); a second brother had a history of two or three vessel CABG and an MI at 62 (II-V). Her sister had elevated cholesterol and a history of carotid endarterectomy (II-VII). We identified a two base pair inversion (c.261_262invGA) sequence variant in *LDLR*, which predicts p.Trp87X. The *LDLR* p.Trp87X mutation has been reported twice and presumed to be a null mutation. In one case, this protein change was predicted from the nonsense change c.260G>A [5]. In the second case [6], the DNA change was not specified in the paper, but we have confirmed this change was also c.260G>A (personal communication, AD Laurie). The previous reported cases (n=3) are of European ancestry (personal communication, I Day and AD Laurie).

The mutation in the present report showed two adjacent nucleotide changes (in the third and first positions of codons 87 and 88, respectively (fig. 1b). It is important to recognize that this sequencing result could represent either a two base pair change in one allele or two single base pair changes, one in each allele. Although compound heterozygous familial hypercholesterolemia was unlikely, given the clinical presentation in the proband, precise determination of the mutation is important for recurrence risk estimations and to facilitate allele-specific mutation detection. The PCR product was also subcloned into a TA-TOPO vector (Invitrogen) and 23 clones were sequenced on an ABI 3100 instrument and compared to GenBank reference (GenBank accession nos. AC011485.6, and NM_000527.3). The two-nucleotide change was found in 18 clones and wild-type sequence in five clones, supporting the hypothesis that the change was in *cis*, and therefore correctly described as c.261_262invGA (fig. 1b). We concluded that the patient has heterozygous familial hypercholesterolemia. *LDLR* p.Trp87X resulting from c.261_262invGA has not been previously reported in the literature. This finding illustrates the importance of describing mutations in terms of the genomic change, which is less ambiguous than protein descriptors. There are three theoretical mutations of *LDLR* codon 87 that would lead to a prediction of p.Trp87X (c.260G>A, c.261G>A, and c.260_261delinsAA). Also, there are innumerable mutations involving multiple codons that could lead to p.Trp87X. Nucleotide descriptors are important because many mutation assays (allele-specific oligonucleotide hybridization, TaqMan assays, etc.) detect specific nucleotide sequences; therefore a custom assay detecting c.260G>A would not detect the other two mutations. The p.Trp87X designation was not based on specific data derived from the mutation in those two cases, but was a prediction, based on assumptions. Protein sequencing is technically feasible, but rarely performed. Because it is the nucleotide change that is determined experimentally and the protein is a prediction, it is the former that should be reported.

In summary, the *ClinSeq* project is designed to identify novel genetic variants that influence clinical phenotypes. Direct sequencing has the advantage of identifying rare or novel DNA sequence variants. It is important to describe sequence variants according to the underlying DNA changes, as opposed to the predicted protein changes as multiple distinct genomic DNA changes may underlie a single predicted protein change. In complex DNA mutations,

it is important to determine if multi-nucleotide changes lie on the same allele (in cis) or on alternate alleles (in trans) as these data have medical and genetic counseling implications for the patients.

Acknowledgments

This study was supported by funds from the Intramural Research Program of the National Human Genome Research Institute, NIH. The opinions expressed here are those of the authors and do not necessarily reflect the opinions of the institutions to which they are affiliated.

References

1. Biesecker LG, Mullikin JC, Facio FM, et al. The ClinSeq Project: Piloting large-scale genome sequencing for research in genomic medicine. *Genome Res* 2009;19:1665–74. [PubMed: 19602640]
2. Cohen JC, Kiss RS, Pertsemlidis A, et al. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* 2004;305:869–72. [PubMed: 15297675]
3. Cohen JC, Boerwinkle E, Mosely TH, et al. Sequence variations in PCSK9, low LDL, and protecting against coronary artery disease. *NEJM* 2006;354:1264–72. [PubMed: 16554528]
4. Frayling TM, Timpson NJ, Weedon MN, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 2007;316:889–94. [PubMed: 17434869]
5. Day IN, Whittall RA, O'Dell SD, et al. Spectrum of LDL receptor gene mutations in heterozygous familial hypercholesterolemia. *Hum Mutat* 1997;10:116–27. [PubMed: 9259195]
6. Laurie AD, Scott RS, George PM. Genetic screening of patients with familial hypercholesterolemia (FH): a New Zealand perspective. *Atheroscler Suppl* 2004;5:13–5. [PubMed: 15556094]

Fig. 1.

Fig 1a. Pedigree of family with familial hypercholesterolemia. The proband is indicated by the arrow. ■ = elevated cholesterol.

Fig 1b. Sequence tracing showing the c.261_262invGA mutation on the mutant allele compared to the wildtype allele from individual II-VI.