Title

ABCG5/G8 gene variation is associated with genetic hypercholesterolemias without mutation in candidate genes and with serum non-cholesterol sterols

Running Title

ABCG5/G8 mutations in genetic hypercholesterolemia

Authors

Itziar Lamiquiz-Moneo[a], Lucía Baila-Rueda[a] , Ana M. Bea[a], Rocío Mateo-Gallego[a], Sofía Pérez-Calahorra[a], Victoria Marco-Benedí[a], Antonio Martín-Navarro[a], Emilio Ros[b], José Carlos Rodríguez-Rey[c], Miguel Pocovi[d], Ana Cenarro[a], Fernando Civeira[a].

Affiliation

[a] Unidad Clínica y de Investigación en Lípidos y Arteriosclerosis, Hospital Universitario Miguel Servet, Instituto de Investigación Sanitaria Aragón (IIS Aragón), CIBERCV, Universidad de Zaragoza, Zaragoza, Spain.

[b] Unitat de Lipids, Servei d'Endocrinologia i Nutrició, Institut d'Investigacions Biomediques August Pi i Sunyer, Hospital Clinic and CIBER Fisiopatologia de la Obesidad y Nutrición (CIBERobn), Instituto de Salud Carlos III (ISCIII), Barcelona, Spain

[c] Departamento de Biología Molecular, Facultad de Medicina, Universidad de Cantabria, and Instituto de Investigación Valdecilla (IDIVAL), Santander, Cantabria, Spain.

[d] Universidad de Zaragoza, Departamento de Bioquímica, Biología Molecular y Celular, Instituto de Investigación Sanitaria Aragón (IIS Aragón), CIBERCV. 50009, Zaragoza, Spain

***Corresponding author**

Name: Itziar Lamiquiz-Moneo

Address: Unidad Clínica y de Investigación en Lípidos y Arteriosclerosis, Hospital Universitario Miguel Servet, Avenida Isabel La Católica 1-3, 50009, Zaragoza, Spain

Telephone number: (34) 976765500 (EXT 2895)

E-mail address: itziarlamiquiz@gmail.com

Keywords

Genetic hypercholesterolemia, *ABCG5/G8*, non-cholesterol sterols, gene variation, biomarkers of cholesterol absorption

Abstract

Objective

The aim of this work was to study the contribution of *ABCG5/G8* genes in patients with a diagnosis of familial hypercholesterolemia without a mutation in any candidate gene (mutation-negative FH), which could predispose to intestinal hyperabsorption of cholesterol, and therefore explain the mechanism of their hypercholesterolemia.

Approach and results

We sequenced the *ABCG5/G8* genes in 217 mutation-negative FH and 97 controls. Surrogate biomarkers of cholesterol absorption and phytosterols (5α-cholestanol, β-sitosterol, campesterol, stigmasterol and sitostanol) were quantified by high performance liquid chromatography tandem mass spectrometry in mutation-negative FH and controls.

We found 8 mutation-negative FH patients (3.73%) with a pathogenic mutation in *ABCG5* or *ABCG8* genes. We observed significantly higher concentration of surrogate biomarkers of cholesterol absorption in mutation-negative FH than in controls. In addition, we found significantly higher concentrations of surrogate biomarkers of cholesterol absorption in mutation-negative FH with mutation in *ABCG5/G8* genes than in mutation-negative FH without a mutation in these genes. A gene score reflecting the number of common single nucleotide variants associated with hypercholesterolemia was significantly higher in cases than in controls ($p=.032$). Subjects with a gene score above the mean had significantly higher 5α-cholestanol and stigmasterol than subjects with a lower gene score.

Conclusions

Mutation-negative FH subjects accumulate an excess of rare and common gene variations in *ABCG5/G8* genes. This variation is associated with increased intestinal absorption of surrogate makers, suggesting that these *loci* contribute to hypercholesterolemia by raising intestinal cholesterol absorption.

Abbreviations

GH: genetic hypercholesterolemias

FH: Familial hypercholesterolemia

TC: Total cholesterol

TG: triglyceride

LDLc: Low density lipoprotein cholesterol

SNVs**:** single nucleotide variations

CHD: cardiovascular disease

Mutation-negative FH: Familial hypercholesterolemia without mutation in candidate genes

HDLc: High density lipoprotein cholesterol

HPLC-MS/MS: high performance liquid chromatography tandem mass spectrometry

BMI: Body mass index

Apo: Apolipoprotein

miRNAs: microRNAs

Introduction

Genetic hypercholesterolemias (GH) are a heterogeneous group of lipid disorders caused by monogenic and polygenic defects and characterized by very high plasma concentrations of total cholesterol (TC) due to increased low-density lipoprotein cholesterol (LDLc) and high risk of premature coronary heart disease (CHD). Familial hypercholesterolemia (FH) is the most common monogenic GH [1] with autosomal codominant transmission and an estimated prevalence of about 1:250-500 in general population [2]. FH is caused by mutations in *LDLR,* the gene coding for the LDL receptor; *APOB,* coding for apolipoprotein (apo) B, *PCSK9* [3, 4]*,* which codes for the enzyme proprotein convertase subtilisin/kexin type 9, or *APOE* genes [5]. In addition, a rare recessive form of FH is also caused by mutations in the *LDLRAP1*gene [6]. However, a causative mutation in candidate genes is not found in approximately 20-40% of clinically defined FH cases [7], suggesting that there are either other genetic causative *loci*, as yet unidentified, or they are severe polygenic hypercholesterolemias. Actually, this GH mimicking the FH phenotype but without a mutation in candidate genes, or mutation-negative FH, accumulate some common single nucleotide variations (SNVs) with small LDLc raising effect [8], that do not fully explain the high LDLc phenotype of affected subjects.

Different GWAS have shown that at least one hundred *loci* are associated with LDLc concentration in the population [9, 10]. One of these *loci* is *ABCG5/G8*; this gene complex encodes the proteins ABCG5 and ABCG8, that form a heterodimer conveyor located on the membrane of enterocytes and hepatocytes. It has been shown that ABCG5/G8 inhibits the intestinal absorption of cholesterol and phytosterols and promotes their hepatobiliary secretion. Severe functional mutations in *ABCG5/G8* cause sitosterolemia, a rare autosomal disorder characterized by an increase of phytosterols in blood, up to 30 times greater than normal [11]. In addition, many studies have shown that *ABCG5/G8* variation is associated with cholesterol and non-cholesterol sterol plasma

levels [12-16]. However, the contribution of *ABCG5/G8 loci* variation in mutation-negative FH has not been previously analyzed.

Considering that mutation-negative FH subjects usually have higher non-cholesterol sterol concentrations than other types of GH, probably due to sterol intestinal hyperabsorption, we hypothesized that genetic variations in *ABCG5/G8* are involved in some forms of mutation-negative FH. Hence, the aim of this study is to analyze common and rare mutations in *ABCG5/G8* in a large cohort of mutation-negative FH and assess their association with non-cholesterol sterols and LDLc, to establish the genetic contribution of these *loci* in this type of GH.

Material and Methods

Subjects

Unrelated subjects (n = 214) 18–79 years of age with the clinical diagnosis of non-FH GH: LDLc above the 95th percentile of the Spanish population [17], triglycerides (TG) below 200 mg/dL, primary cause and familial presentation (at least one first-degree relative with the same phenotype) from the Lipid Unit at Hospital Universitario Miguel Servet, Zaragoza and Unitat de Lipids, Servei d'Endocrinologia i Nutrició, Institut d'Investigacions Biomediques August Pi i Sunyer, Hospital Clinic, Barcelona were selected for this study. In all subjects, the absence of a pathogenic mutation in *LDLR, APOB* and *PCSK9* genes was confirmed by the Lipochip ® platform [18]. Exclusion criteria were the presence of ε2/ε2 genotype or the p.(Leu167del) mutation in *APOE* gene. Secondary causes of hypercholesterolemia including: obesity (body mass index >30 kg/m2), poorly controlled type 2 diabetes (HbA1c >8%), renal disease with glomerular filtration rate <30 ml/min and/or macroalbuminuria, liver diseases (ALT> 3 times upper normal limit), hypothyroidism (TSH >6 mIU/L), pregnancy, autoimmune diseases and protease inhibitors were exclusion criteria. Cardiovascular risk factors assessment, personal and family history of cardiovascular disease, consumption of drugs affecting intestinal or lipid metabolism and anthropometric measurements were performed in all participants. The normolipemic group (n= 97) consisted of healthy, unrelated men and

women volunteers aged 18–79 years, who underwent a medical examination at the Hospital Miguel Servet of Zaragoza. Exclusion criteria for control subjects were personal or parental history of premature cardiovascular disease or dyslipidemia, current acute illness, or use of drugs that might influence glucose or lipid metabolism. All subjects signed informed consent to a protocol previously approved by our local ethics committee (Comité Ético de Investigación Clínica de Aragón, Zaragoza, Spain).

Clinical and laboratory determinations

Cases and controls were assessed for personal and familial history of cardiovascular disease, medication use and cardiovascular risk factors. EDTA plasma and serum samples were collected after 10-12 hours of fasting in all participants after 6 weeks without lipid-lowering drugs in absence of cardiovascular disease. TG levels were determined by standard enzymatic methods. High density lipoprotein cholesterol (HDLc) was measured directly by an enzymatic reaction using cholesterol oxidase (UniCel DxC 800; Beckman Coulter, Inc., Brea, CA, USA). Lipoprotein (a), Apo A1, Apo B and C-reactive protein were determined by IMMAGE kinetic nephelometry (Beckman-Coulter, Inc.). In patients with cardiovascular disease, lipid values were obtained from medical records.

Genetic analysis

Genomic DNA from whole blood samples was isolated by using standard methods. Promoters, coding regions and intron-exon boundaries of *ABCG5* (NM_022436) and *ABCG8* (NM_022437) were amplified by PCR and purified by ExoSap-IT (USB). Amplified fragments were sequenced by the Sanger method using the BigDye 3.1 sequencing kit (Applied Biosystems) in an automated ABI 3500xL sequencer (Applied Biosystems). DNA sequences were analyzed using VariantReporter™ software (Applied Biosystems). *APOE* genotypes were determined by DNA sequencing of exon 4 as previously described [19].

To evaluate the pathogenicity of new identified genetic variants, we used PolyPhen-2 [20] and Mutation Taster [21]. The effect of variants in potential splicing sites was predicted with NetGene2 [22] and NNSplice [23]. We refer to non-synonymous variants as sequence variations causing amino acid substitutions (missense variants) or introducing a

premature stop codon (nonsense variants). Intronic variants were considered when they were located in intron-exon boundaries.

In order to compare the frequency of identified variants with the general population we have compiled the allele frequencies of the identified variants from the 1000 Genomes Project [24].

Mutation was defined as those genetic variants whose frequency is lower than 1% in the general population. A mutation was defined as pathogenic when it was not present in controls and bioinformatic analysis prediction defined it as "damaging" (Polyphen-2) or "disease causal" (MutationTaster).

Gene score

A gene score using the sum of the risk alleles of 27 genetic variants with allelic frequencies statistically different between cases and controls was calculated for each subject. (Aquí tendrias que poner la referencia del gene score de 27 variantes)

Serum sterol quantification

Serum concentrations of cholesterol, 5α-cholestanol, β-sitosterol, campesterol, stigmasterol and sitostanol were quantify by high performance liquid chromatography tandem mass spectrometry (HPLC-MS/MS) in 206 non-FH GH subjects without lipid-lowering drugs and 97 controls according to the method previously described [25]. In 8 subjects we did not quantified non-cholesterol sterols because 7 subjects had CHD and one subject did not have a serum sample. Data was expressed as mg/dL, as well as normalized to total cholesterol. Briefly, 100 µL of serum were transferred to screw-capped vials, and 6 µl of deuterium-labeled internal standard, [2H6] cholesterol-26,26,26,27,27,27 (7.9 mM) for non-cholesterol sterols, and [2H7] cholesterol-25,26,26,26,27,27,27 for cholesterol, were added. Alkaline hydrolysis was performed for 20 min at 60ºC in an ultrasound bath, followed by a double extraction with 3 µl of hexane. The extract was loaded onto the solid-phase extraction cartridge (1 mg, Discovery DSC-18, Supelco, Spain), which was preconditioned with 400 µl of methanol and gravity eluted. The non-cholesterol sterols and cholesterol were desorbed with 1.4 µL

of 2-propanol by gravity, and 40 µL of the final mixture was injected into the HPLC-MS/MS system.

Statistical Analysis

Analyses were performed using SPSS version 20.0 (Chicago, Ilinois, USA). The nominal level for significance was $p< 0.05$. Normal distribution of variables was analyzed by the Kolmogorov–Smirnov test. Quantitative variables with normal distribution were expressed as mean ± standard deviation and were analyzed with a Student T test. Variables with a skewed distribution were expressed as median and interquartile range and were analyzed with the Mann-Whitney U test. Qualitative variables were expressed as a percentage and were analyzed by Chi-square test.

Results

The main clinical and biochemical characteristics of the two groups: 214 mutation-negative FH patients and 97 normolipemic controls are presented in Table 1. Mutation-negative FH participants had statistically significant higher values of total cholesterol, triglycerides, HDLc, LDLc, Lipoprotein (a), Apo B, hypertension and cardiovascular disease than normolipemic controls. No differences in age, body mass index (BMI) and *APOE* genotype were found between patients and controls.

Non-cholesterol sterol to TC ratios are presented in Table 2. 5α-cholestanol, stigmasterol and sitostanol ratios were significantly higher in cases than in controls (*p=.023, p<.001 and p=.003*, respectively). The sum of all cholesterol absorption surrogate markers and phytosterols was also significantly higher in mutation-negative FH than in controls (*p= .028*).

The frequency of SNVs in *ABCG5/G8* genes found in mutation-negative FH was compared to the frequency in controls and in the 1000 Genomes Project. Table 3 shows the 29 SNVs identified whose allelic frequencies were statistically different from those identified in controls or described in the 1000 Genomes Project. They include 27 SNPs and two mutations: p.(Gly269Arg) in *ABCG5* and p.(Gly512Arg) in *ABCG8*.

Table 4 shows the clinical characterization of patients carrying pathogenic mutations identified by sequencing in *ABCG5* and *ABCG8.* A total of eight patients out of 214 (3.73%) were carriers of 6 pathogenic mutations: p.(Asn578Ser), p.(Gly288Arg), p.(Arg198Gln), p.(Gly269Arg), and p.(Asn285Ser) in *ABCG5* and p.(Gly512Arg) in *ABCG8.* One of them was described for the first time in this work: p.(Asn578Ser) in *ABCG5.* None of these pathogenic mutations were present in controls.

Subjects with pathogenic mutation in *ABCG5/G8* genes had significantly higher non-cholesterol sterol to TC ratios (5α-cholestanol, β-sitosterol, campesterol, and stigmasterol) and cholesterol absorption than subjects without mutation in *ABCG5/G8* genes (*p=.042, p=.045, p=.034, p=.047* and *p=.030*, respectively). TC concentrations did not show significant differences between both studied groups (Table 5).

Among common genetic variants with statistically different frequencies between cases and 1000 Genomes Project, only c.*380T>G, located in the 3´ UTR of *ABCG5* gene was associated with statistically significant differences in non-cholesterol sterol to TC ratios (5α-cholestanol, β-sitosterol, stigmasterol and campesterol) and in cholesterol absorption. T allele carriers had higher intestinal cholesterol absorption than subjects carrying the G allele (*p=.011, p=.00, p=.002, p=.022* and *p=.002,* respectively), in spite of TC concentrations were not significantly different between genotypes (Table 6).

A gene score reflecting the number of common SNVs associated with hypercholesterolemia was significantly higher in cases than in controls (*p=.032*). Subjects with scores above the mean (>33 points) had significantly higher non-cholesterol sterol to TC ratios (5α-cholestanol and stigmasterol) than subjects with the lowest score (*p=.034* and *p=.029*, respectively). The cholesterol absorption showed a trend to be higher in subjects with the highest score than subjects with the lowest score (*p=.077*) (Table 7).

Discussion

This study shows that the *ABCG5/G8* genes variation plays a substantial role in the pathogenesis of genetic hypercholesterolemia unrelated to *LDLR*, *APOB* and *PCSK9* genes. Subjects with these GH have, as previously described, higher surrogate biomarkers of cholesterol intestinal absorption than controls [26], and three new important results can be highlighted from this work.

First, *ABCG5/G8* genetic variation contributes to this group of GH through rare pathogenic mutations with large effects and with common variants with small effects. Patients with rare pathogenic mutations present, approximately, have double serum concentrations of surrogate markers of cholesterol intestinal absorption than subjects not carrying mutations. However, the percentage of mutation-negative FH subjects carrying these rare pathogenic mutations is low, approximately 4%. We did not find any pathogenic mutation in the control group, so we consider that these major defects are responsible for most of the high LDLc phenotype. We found 8 patients carrying six pathogenic mutations. Five mutations were in *ABCG5* and one mutation in *ABCG8.* The mutations p.(Arg198Gln), p.(Gly269Arg), p.(Asn296Ser) and p.(Gly288Arg) in ABCG5 are located in the cytosolic N-terminal region, prior to the first transmembrane domain. In this N-terminal region of ABCG5, one mutation causing sitosterolaemia has been previously described [27]. The mutation p.(Asn578Ser) in ABCG5 is located in the extracellular domain, and this position seems to interact with the amino acid Arg419. The mutations p.(Arg419Pro) and p.(Arg419His) at the same residue, have been shown to be causal for sitosterolaemia. The mutation p.(Gly512Arg) in *ABCG8* is located in the transmembrane α-helix domain; a close mutation in this domain, p.(Leu501Pro), causes sitosterolaemia [28].

Second, the study shows that mutation-negative FH subjects accumulate risk alleles of *ABCG5/G8* with small effects on non-sterols cholesterol levels, suggesting that, in the vast majority of these GH, the contribution of *ABCG5/G8* genetic variation has a limited effect within a polygenic background, as occurs in other genetic dyslipidemias [29]. Among common SNVs frequently found in mutation-negative FH, only one SNVs in the 3´UTR showed

significant variations in non-cholesterol sterols by itself. Variants in 3´UTR have been previously involved in the pathogenesis of some mutation-negative FH. An important mechanism of post-transcriptional regulation linked to the 3´UTR implicates the binding of microRNAs (miRNAs), small non-protein coding RNAs that regulate gene expression at the post-transcriptional level. [30]. The frequency of the variant c.*380T>G in *ABCG5*, located in 3´UTR, is significantly lower in mutation-negative FH than in controls, and T allele carriers had higher intestinal cholesterol absorption than subjects carrying the G allele. The c.*380T>G variant has been previously associated with sitosterolaemia [31]. Bioinformatic analysis of this variant showed that T allele carriers have mir-494, but it is not present in G allele carries. For this reason, we propose that mir-494 could regulate the absorption of non-cholesterol sterols and could explain in part the pathogenesis of the hypercholesterolemia in these subjects.

Finally, our study supports the multifactorial origin of most mutation-negative FH, in which the intestinal hyperabsorption of sterols plays a minor role that does not fully explain the etiology of these hypercholesterolemias, and with different contributions among subjects depending on the number of risk alleles they carry. Those subjects that accumulate many risk alleles of *ABCG5/G8* show a clearer hyperabsorption mechanism that could explain some LDLc variation. We have recently analyzed a group of mutation-negative FH families with hyperabsorption in the proband. The cosegregation analysis showed a substantial contribution of hyperabsorption on the LDLc concentration, although with high variation among the families [26].

An interesting issue is the potential relationship between *ABCG5/G8* gene variation and the CHD risk. The C allele of the variant c.1199C>A, p.(Thr400Lys) in *ABCG8* has been previously associated with CHD [32]. In our study, the frequency of this risk allele variant was significantly higher in mutation-negative FH subjects than in 1000 Genomes Project, suggesting that the association with CHD previously reported could be mediated by its effect on LDLc. However, the association of plasma phytosterols with CHD has not been clearly established [33]. Actually, none of 8 subjects with severe pathogenic mutations in *ABCG5/G8* have CHD in

our study. Therefore, further studies are required to establish the potential association of *ABCG5/G8* variation and atherosclerosis.

Our study has limitations: We have not performed functional studies, but due to the large number of variants found, their study would require a different approach. We have analyzed surrogate markers for cholesterol intestinal absorption; although very well validated, they have some limitations, and may be influenced by *APOE* genotype, sex, age and BMI. However, the differences in non-cholesterol sterols between mutation-negative FH subjects and controls and between patients with pathogenic mutations *ABCG5/G8* and patients without pathogenic mutations were not modified by any of these confounder factors.

In conclusion, the sequencing analysis of a large group of subjects with genetic hypercholesterolemia and without a pathogenic mutation in *LDLR*, *APOB*, *PCSK9* and *APOE* genes, shows that these subjects accumulate an excess of rare and common gene variations in *ABCG5/G8* genes. This variation is associated with increased intestinal absorption surrogate markers, suggesting that these *loci* contribute to their hypercholesterolemia by raising intestinal cholesterol absorption.

Acknowledgments

a) Acknowledgments

Genetic analyses were performed in the Sequencing and Functional Genomics Facility of Servicios Científico Técnicos of CIBA (IACS-Universidad de Zaragoza), Zaragoza, Spain.

b) Sources of Funding

This study was supported by grants from the Spanish Ministry of Economy and Competitiveness PI15/01983, PI13/02507 and CIBERCV. These projects “PI15/01983, PI13/02507 and CIBERCV” are co-financed by the Carlos III Health Institute and the European Regional Development Fund (ERDF) of the European Union "A way to make Europe". (Acuerdate de agradecer a Fundacion Cuenca Villorio)

c) Disclosure

I certify that neither I nor my co-authors have a conflict of interest as described above that is relevant to the subject matter or materials included in this work.

References

1. Goldstein JK, Hobbs HH, Brown MS. Familial hypercholesterolemia. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds). *The Metabolic & Molecular Bases of Inherited Disease*. 8th ed. New York: McGraw-Hill; 2001. p. 2863–2913.
2. Marks D, Thorogood M, Neil HA, Humphries SE. A review on the diagnosis, natural history, and treatment of familial hypercholesterolaemia. *Atherosclerosis*. 2003;168:1-14.
3. Civeira F. International Panel on Management of Familial Hypercholesterolemia. Guidelines for the diagnosis and management of heterozygous familial hypercholesterolemia. *Atherosclerosis*. 2004; 173:55-68.
4. Futema M, Plagnol V, Li K, Whittall RA, Neil HA, Seed M, Simon Broome Consortium, Bertolini S, Calandra S, Descamps OS, Graham CA, Hegele RA, Karpe F, Durst R, Leitersdorf E, Lench N, Nair DR, Soran H, Van Bockxmeer FM[1]; UK10K Consortium, Humphries SE. Whole exome sequencing of familial hypercholesterolaemia patients negative for LDLR/APOB/PCSK9 mutations. *J Med Genet*. 2014;515:37-44.
5. Cenarro A, Etxebarria A, de Castro-Orós I, Stef M, Bea AM Palacios L, Mateo-Gallego R, Benito-Vicente A, Ostolaza H, Tejedor T, Martín C, Civeira F. The p.Leu167del mutation in APOE gene causes Autosomal Dominant Hypercholesterolemia by down-regulation of LDL receptor expression in hepatocytes. *J Clin Endocrinol Metab*. 2016; 101:2113-21.
6. Ciccarese MA, Pacifico G, Tonolo G, Pintus P, Nikoshkov A, Zuliani G, Fellin R, Luthman H and Maioli M. A new locus for autosomal recessive hypercholesterolemia maps to human chromosome 15q25-q26. *Am J Hum Genet*. 2000; 66:453-460.

7. Palacios L, Grandoso L, Cuevas N, Olano-Martín E, Martinez A, Tejedor D, Stef M. Molecular characterization of familial hypercholesterolemia in Spain. *Atherosclerosis*. 2012;221:137-42.
8. Talmud PJ, Shah S, Whittall R, Futema M, Howard P, Cooper JA, Harrison SC, Li K, Drenos F, Karpe F, Neil HA, Descamps OS, Langenberg C, Lench N, Kivimaki M, Whittaker J, Hingorani AD, Kumari M, Humphries SE. Use of low-density lipoprotein cholesterol gene score to distinguish patients with polygenic and monogenic familial hypercholesterolemia: a case-control study. *Lancet*. 2013; 381:1293-301.
9. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM Koseki M et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010; 466:707–13.
10. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S Kanoni S et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet*. 2013; 45:1-24.
11. Yu , L. , J. Li-Hawkins , R. E. Hammer , K. E. Berge , J. D. Horton , J. C. Cohen , and H. H. Hobbs. Overexpression of ABCG5 and ABCG8 promotes biliary cholesterol secretion and reduces fractional absorption of dietary cholesterol. *J. Clin. Invest.* 2002; 110**:**671–80 .
12. Lily Jakulj , 1, Maud N. Vissers , Michael W. T. Tanck , Barbara A. Hutten , Frans Stellaard , John J. P. Kastelein , Geesje M. Dallinga-Thie. *ABCG5*/*G*8 polymorphisms and markers of cholesterol metabolism: systematic review and meta-analysis. *J Lipid Res*. 2010;51:3016-26.
13. Berge , K. E., K. von Bergmann, D. Lutjohann, R. Guerra, S. M. Grundy, H. H. Hobbs, and J. C. Cohen. Heritability of plasma non-cholesterol sterols and relationship to DNA sequence polymorphism in ABCG5 and ABCG8. *J. Lipid Res* 2002; 43**:** 486–94.
14. Hubacek , J. A. , K. E. Berge , J. Stefkova , J. Pitha , Z. Skodova , V. Lanska , and R. Poledne..Polymorphisms in ABCG5 and ABCG8 transporters and plasma cholesterol levels. *Physiol. Res*. 2004;53**:**395–401.

15. Weggemans RM, Zock PL, Tai ES, Ordovas JM, Molhuizen HO, Katan MB. ATP binding cassette G5 C1950G polymorphism may affect blood cholesterol concentrations in humans. *Clin. Genet.* 2002;62**:**226–229.
16. Chen ZC, Shin SJ, Kuo KK, Lin KD, Yu ML, Hsiao PJ. Significant association of ABCG8:D19H gene polymorphism with hypercholesterolemia and insulin resistance. *J. Hum. Genet.* 2008;53**:**757–63.
17. Wu G, Li GB, Yao M, Zhang DQ, Dai B, Ju CJ, Han M. *ABCG5/8* variants are associated with susceptibility to coronary heart disease. *Mol Med Rep*. 2014;6:2512-20.
18. Stef MA, Palacios L, Olano-Martín E, Foe-A-Man C, van de Kerkhof L, Klaaijsen LN, Molano A, Schuurman EJ, Tejedor D, Defesche JC. A DNA microarray for the detection of point mutations and copy number variation causing familial hypercholesterolemia in Europe. *J Mol Diagn.* 2013;15:362-72.
19. Solanas-Barca M, De Castro-orós I, Mateo-Gallego R, Cofán M, Plana N Puzo J et al. Apolipoprotein E gene mutations in subjects with mixed hyperlipidemia and a clinical diagnosis of familial combined hyperlipidemia. *Atherosclerosis.* 2012; 222::449-55.
20. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010;7:248–49.
21. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods*. 2010;7:575-6.
22. Brunak S, Engelbrecht J, Knudsen S. Prediction of human mRNA donor and acceptor sites from the DNA sequence. *J Mol Biol.* 1991;220:49–65.
23. Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection inGenie. *J Comput Biol.* 1997;4:311–23.
24. 1000 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM. An integrated map of genetic variation from 1,092 human genomes. *Nature.* 2012;491:56–65.

25. Baila-Rueda L, Cenarro A, Cof_an M, et al. Simultaneous determination of oxysterols, phytosterols and cholesterol precursors by high performance liquid chromatography tandem mass spectrometry in human serum. Analytical Methods. 2013;5:2249–2257.
26. Baila-Rueda L, Cenarro A, Lamiquiz-Moneo I, Mateo-Gallego R, Bea AM, Perez-Calahorra S, Marco-Benedí V, Civeira F. Bile acid synthesis precursors in subjects with genetic hypercholesterolemia negative for LDLR/APOB/PCSK9/APOE mutations. Association with lipids and carotid atherosclerosis. *J Steroid Biochem Mol Biol.* 2016;16:30263-71.
27. Hubacek JA, Berge KE, Cohen JC, Hobbs HH. Mutations in ATP-cassette binding proteins G5 (ABCG5) and G8 (ABCG8) causing sitosterolemia. Hum Mutat. 2001;4:359-60.
28. Jyh-Yeuan Lee, Lisa N. Kinch, Dominika M. Borek, Jin Wang, Junmei Wang, Ina L. Urbatsch, Xiao-Song Xie, Nikolai. Grishin, Jonathan C. Cohen, Zbyszek Otwinowski, Helen H. Hobbs, and Daniel M. Rosenbaum. Crystal structure of the human sterol transporter ABCG5/ABCG8. *Nature*. 2016;533:561–64.
29. De Castro-Orós I, Cenarro A, Tejedor MT, Baila-Rueda L, Mateo-Gallego R, Lamiquiz-Moneo I, Pocoví M, Civeira F. Common genetic variants contribute to primary hypertriglyceridemia without differences between familial combined hyperlipidemia and isolated hypertriglyceridemia. *Circ Cardiovasc Genet.* 2014;7:814-21.
30. Perez-Campo F; De Castro-Orós I; Noriega A; Cofán M; Lamiquiz-Moneo I; Cenarro A; Ros E; Civeira F; Pocoví M; Rodríguez-Rey JC. Functional analysis of new 3´UTR genetic variants in genes associated with hypercholesterolemias. *J Clin Lipidol*. 2017 [In press].
31. NCBI ClinVar. https://www.ncbi.nlm.nih.gov/clinvar/variation/336031/
32. Wu G, Li GB, Yao M, Zhang DQ, Dai B, Ju CJ, Han M. *ABCG5/8* variants are associated with susceptibility to coronary heart disease. *Mol Med Rep.* 2014;6:2512-20.

33. Jeemon P, Pettigrew K, Sainsbury C, Prabhakaran D, Padmanabhan S. Implications of discoveries from genome-wide association studies in current cardiovascular practice. *World J Cardiol.* 2011;3:230-47.

Highlights

- Mutation-negative FH subjects accumulate an excess of rare and common gene variations in *ABCG5/G8* genes
- *ABCG5/G8* variants modulate the absorption non-cholesterol sterols
- Patients with variants or mutations in *ABCG5/G8* gene present an increased in concentration non-cholesterol sterols.

Table 1. Clinical and biochemical characteristics of controls and mutation-negative FH.

		Controls (n = 97)	Mutation-negative FH (n = 214)	*p*
Age, years		44.5 ± 17.3	47.0 ±12.4	0.135
Men, n (%)		37 (38.1)	112 (47.6)	0.020
Body mass index, kg/m^2		25.3 ± 4.59	24.5 ± 2.69	0.149
Total cholesterol by HPLC MS/MS, mg/dL		201 ± 45.3	302 ± 59.8	<0.001
Triglycerides, mg/dL		75.0 (49.0-103)	112 (86.0-157)	<0.001
LDL cholesterol, mg/dL		122 (114-146)	224 (205-244)	<0.001
HDL cholesterol, mg/dL		53.0 (45.0-64.3)	55.0 (47.0-68.0)	0.042
Lipoprotein(a), mg/dL		15.7 (7.26-30.6)	44.3 (16.1-98.4)	<0.001
Apolipoprotein B, mg/dL		98.8 ± 20.0	154 ± 32.3	<0.001
Apolipoprotein A1, mg/dL		157 (138-177)	157 (137-180)	0.503
Glucose, mg/dL		85.5 (79.0-92.0)	87.0 (82.0-96.0)	0.169
Hypertension, n (%)		6 (6.38)	32 (15.0)	0.035
Diabetes, n (%)		0 (0)	1 (0.47)	0.509
Cardiovascular disease, n (%)		0 (0)	7 (3.27)	0.073
Tobacco, n (%)	Non smoker	53 (54.6)	100 (46.7)	0.330
	Current smoker	22 (22.7)	61 (28.5)	
	Former smoker	22 (22.7)	53 (24.8)	
APOE genotype, n (%)	E2/E3	10 (10.3)	14 (6.54)	0.107
	E3/E3	69 (71.1)	136 (63.5)	
	E4/E3	15 (15.5)	57 (26.6)	
	E4/E4	1 (1.03)	6 (2.8)	
	E4/E2	2 (2.06)	1 (0.46)	

Quantitative variables are expressed as mean ± standard deviation, except for variables not following normal distribution that were expressed as median (interquartile range). Qualitative variables were expressed as %. The *p* value was calculated by Student's t test or Mann-Whitney U and Chi-square, as appropriate.

HPLC-MS/MS: high performance liquid chromatography tandem mass spectrometry.

Table 2. Non-cholesterol sterols in controls and mutation-negative FH.

	Controls (n = 97)	Mutation-negative FH (n = 206)	*p*
5α-cholestanol to TC x 10^3	2.06 (1.57-2.47)	2.11 (1.66-2.89)	0.023
β-sitosterol to TC x 10^3	1.60 (0.985-2.05)	1.84 (1.33-2.74)	<0.001
Campesterol to TC x 10^3	0.935 (0.703-1.22)	0.939 (0.699-1.33)	0.435
Stigmasterol to TC x 10^3	0.102 (0.064-0.148)	0.171 (0.097-0.291)	<0.001
Sitostanol to TC x 10^3	0.076 (0.040-0.154)	0.046 (0.027-0.107)	0.008
Cholesterol absorption*	4.88 (3.70-6.10)	5.24 (4.14-7.34)	0.028

*Cholesterol absorption denotes the sum of 5α-cholestanol, β-sitosterol, campesterol, stigmasterol and sitostanol.

Quantitative variables with not following normal distribution were expressed as median (interquartile range). The *p* value was calculated by Mann-Whitney U.

Table 3**.** Allele frequencies of genetic variants whose allelic frequencies were statistically different in mutation-negative FH from those identified in controls or described in the 1000 Genomes Project.

Gene	Nucleotide change	Predicted aminoacid change	Bioinformatic Analysis	Frequency in mutation-negative FH	Frequency in controls	p^1	Frequency in the 1000 Genomes population	P^2	Frequency in ExAc	P^3
ABCG5	rs6756629 c.148C>T	p.(Arg50Cys)	Damage	0.047	0.067	0.243	0.076	0.0457	0.066	0.1094
	rs9789463 c.502-395T>G	-	Benign	0.210	0.242	0.378	0.347	<0.001	0.224	0.665
	rs4549146 c.339A>G	p.(Val113Val)	Benign	0.276	0.351	0.062	0.422	<0.001	0.4319	<0.001
	rs4557032 c.502-192C>T	-	Benign	0.259	0.273	0.717	0.344	0.0017	-	-
	rs114938914 c.502-256G>A	-	Benign	0.040	0.062	0.225	0.077	0.0099	0.1416	0.007
	rs552803459 c.805G>A	p.(Gly269Arg)	Damage	0.005	0.000	0.339	0.000	0.0296	<0.001	<0.001
	rs4148187 c.1324+164C>T	-	Benign	0.264	0.253	0.761	0.334	0.0093	-	-
	rs200839584 c.1570G>A	p.(Val524Ile)	Benign	0.005	0.000	0.339	0.000	0.0296	<0.001	<0.001
	rs150002815 c.1320T>A	p.(Asn440Lys)	Benign	0.012	0.000	0.129	0.000	0.0006	<0.001	<0.001
	rs2278356 c.*380 T>G	-	Benign	0.270	0.340	0.1183	0.320	0.0339	-	-
	rs2278357 c.*416G>A	-	Benign	0.092	0.124	0.226	0.157	0.0011	-	-

ABCG8	rs3806471 c.-19T>A	-	Benign	0.237	0.227	0.7791	0.297	0.0204	0.417	<0.001
	rs72647315 c.-15A>C	-	Benign	0.050	0.050	0.9390	0.000	0.0296	0.000	<0.001
	rs11887534 c.55G>A	p.(Asp19Asn)	Benign	0.042	0.096	0.0082	0.080	0.0107	0.077	0.007
	rs4148209 c.64-21C>A	-	Benign	0.291	0.273	0.6476	0.371	0.0038	0.388	<0.001
	rs4148210 c.64-7C>T	-	Benign	0.275	0.278	0.9243	0.372	<0.001	0.386	<0.001
	rs4148211 c.161A>G	p.(Tyr54Cys)	Benign	0.279	0.273	0.8741	0.371	<0.001	0.387	<0.001
	rs4148213 c.165+105G>C	-	Benign	0.296	0.278	0.6576	0.372	0.0058	-	-
	rs4148214 c.561+43T>C	-	Benign	0.322	0.412	0.0280	0.461	<0.001	0.561	<0.001
	rs114197606 c.964+178G>A	-	Benign	0.002	0.000	0.4494	0.032	0.0057	-	-
	rs376069170 c.1534G>A	p.(Gly512Arg)	Probably damage	0.002	0.000	0.4494	0.000	0.0025	0.000	<0.001
	rs4148217 c.1199C>A	p.(Thr400Lys)	Benign	0.033	0.000	0.0105	0.201	<0.001	0.2146	<0.001
	rs4245794 c.1212-77T>C	-	Benign	0.068	0.093	0.2811	0.107	0.0209	-	-
	rs34198326 c.1411+104A>G	-	Benign	0.052	0.124	0.0015	0.090	0.0128	-	-
	rs371711306 c.1845G>C	p.(Met615Ile)	Benign	0.050	0.000	0.3403	0.000	0.030	0.000	<0.001
	rs55741639 c.965-58A>G	-	Benign	0.023	0.000	0.0314	0.016	0.2072	-	-
	rs4953027 c.1412-152A>G	-	Benign	0.289	0.117	<0.001	0.247	0.0956	-	-

	rs10709506 c.1411+90delA	-	Benign	0.132	0.240	0.0009	0.153	0.3054	-	-
	rs28517482 c.1412-8C>T	-	Benign	0.343	0.196	0.0002	0.397	0.0548	0.599	<0.001

The *p* value was calculated by Chi-square test, by comparing mutant versus wild-type allelic frequencies between mutation-negative FH and controls (p^1), between mutation-negative FH and 1000 Genomes Project (p^2) and between mutation-negative FH and ExAc Browser Data (p^3).

Table 4. Clinical characteristics of patients carrying pathogenic mutations in *ABCG5/G8* genes.

Gene	Nucleotide change	Predicted aminoacid change	Age (years)	Sex	BMI (kg/m^2)	TC mg/dL	TG* mg/dL	Apo B, mg/dL	*APOE* Genotype	Non q
ABCG5	rs146534033 c.1733A>G	p.(Asn578Ser)	49	M	19.8	308	149	157	ε3/ε3	
			49	M	25.5	312	180	118	ε3/ε4	
	rs139264483 c.862G>T	p.(Gly288Cys)	56	H	25.4	263	166	123	ε3/ε4	
	rs141828689 c.593G>A	p.(Arg198Gln)	58	M	25.0	316	156	136	ε2/ε3	
			48	H	26.7	315	211	153	ε3/ε3	
	rs552803459 c.805G>A	p.(Gly269Arg)	46	M	26.4	362	104	130	ε3/ε3	
	c. 887A>G	p.(Asn296Ser)	48	H	22.9	309	102	159	ε3/ε3	
ABCG8	rs376069170 c.1534G>A	p.(Gly512Arg)	51	H	25.0	272	189	109	ε3/ε3	

BMI: Body mass index. TC: Total cholesterol. TG: triglyceride. ApoB: Apolipoprotein B

Quantitative variables are expressed as mean, except for variables not following normal distribution that were expressed as median*.

Table 5. Lipids and non-cholesterol sterols in mutation-negative FH subjects with mutation in *ABCG5/G8* genes and in mutation-negative FH without mutation in *ABCG5/G8* genes.

	Mutation-negative FH with *ABCG5/G8* mutation* (n = 7)	Mutation-negative FH without *ABCG5/G8* mutation (n = 199)	*p*
Age (years)	49.9 ± 3.89	46.9 ±12.5	0.111
Sex, men, n (%)	4 (50%)	108 (52.4%)	0.893
Body Mass Index, Kg/m^2	24.5 ± 2.39	24.5 ± 2.72	0.946
Cholesterol by HPLC MS/MS, mg/dL	296 ± 83.2	290 ± 61.1	0.820
Triglycerides, mg/dL	157 ± 38.6	120 ± 49.8	0.039
LDL cholesterol, mg/dL	229 (206-243)	223 (205-244)	0.760
HDL cholesterol, mg/dL	47 (43.3-50.8)	56.0 (48.0-68.0)	0.015
Apolipoprotein B, mg/dL	158 ± 25.9	154 ± 32.6	0.775
5α-cholestanol to TC x 10^3	3.13 (2.05-8.59)	2.11 (1.66-2.86)	0.042
β-sitosterol to TC x 10^3	3.03 (1.86-11.5)	1.84 (1.33-2.72)	0.045
Campesterol to TC x 10^3	1.64 (0.907-4.94)	0.938 (0.691-1.32)	0.034
Stigmasterol to TC x 10^3	0.447 (0.167-0.819)	0.167 (0.097-2.89)	0.047
Sitostanol to TC x 10^3	0.084 (0.057-0.314)	0.046 (0.027-0.107)	0.088
Cholesterol absorption†	7.97 (5.48-21.3)	5.14 (4.13-7.15)	0.030

*Referred to patients described in detail in Table 4.

†Cholesterol absorption denotes the sum of 5α-cholestanol, β-sitosterol, campesterol, stigmasterol and sitostanol.

Quantitative variables are expressed as mean ± standard deviation, except for variables not following normal distribution that were expressed as median (interquartile range). The *p* value was calculated by Student's t test or Mann-Whitney U, as appropriate.

Table 6. Non-cholesterol sterols in mutation-negative FH subjects with rs2278356 variant and in mutation-negative FH without rs2278356 variant. .

rs2278356 c.*380T>G	Mutation-negative FH with variant c.*380T>G (n=97)	Mutation-negative FH without variant c.*380T>G (n=109)	*p*
Cholesterol by HPLC MS/MS, mg/dL	292 ± 62.8	290 ± 61.1	0.838
Triglycerides, mg/dL	118 ± 46.8	124 ± 52.8	0.317
LDL cholesterol, mg/dL	218 (205-239)	230 (206-250)	0.021
HDL cholesterol, mg/dL	56.0 (48.5-67.0)	55.0 (46.0-68.0)	0.459
Apolipoprotein B, mg/dL	151 ± 33.6	157 ± 31.3	0.176
5α-cholestanol to TC x 10^3	2.07 (1.69-2.82)	2.37 (1.68-3.41)	0.011
β-sitosterol to TC x 10^3	1.68 (1.26-2.40)	2.28 (1.42-3.40)	0.001
Campesterol to TC x 10^3	0.831 (0.642-1.31)	1.16 (0.757-1.64)	0.002
Stigmasterol to TC x 10^3	0.142 (0.082-0.278)	0.179 (0.097- 0.329)	0.022
Sitostanol to TC x 10^3	0.043 (0.025-0.088)	0.052 (0.027-0.116)	0.420
Cholesterol absorption†	4.66 (4.03-6.44)	6.12 (4.34-8.32)	0.002

†Cholesterol absorption denotes the sum of 5α-cholestanol, β-sitosterol, campesterol, stigmasterol and sitostanol.

Quantitative variables are expressed as mean ± standard deviation, except for variables not following normal distribution that were expressed as median (interquartile range). ln(5α-cholestanol to TC x 10^3), ln(β-sitosterol to TC x 10^3), ln(campesterol to TC x 10^3), ln(stigmasterol to TC x 10^3), ln(sitostanol to TC x 10^3) and ln(cholesterol absorption) were used to calculate *p* for trend adjusted by confusing factors: body mass index, sex and age.

Table 7. Non-cholesterol sterols in mutation-negative FH according to score SNV gene score.

	Mutation-negative FH with lower gene score (n = 100)	Mutation-negative FH with higher gene score (n =106)	*p*
Cholesterol by HPLC MS/MS, mg/dL	282 ± 62.5	298 ± 60.3	0.074
Triglycerides, mg/dL	121 ± 48.6	121 ± 51.3	0.967
LDL cholesterol, mg/dL	225 (204-245)	224 (205-241)	0.586
HDL cholesterol, mg/dL	58.5 (48.0-69.0)	55.0 (46.0-67.0)	0.420
Apolipoprotein B, mg/dL	153 ± 34.7	155 ± 30.5	0.659
5α-cholestanol to TC x 10^3	2.036 (1.65-2.82)	2.33 (1.78- 3.38)	0.034
β-sitosterol to TC x 10^3	1.86 (1.31- 2.67)	1.89 (1.40-3.01)	0.149
Campesterol to TC x 10^3	0.938 (0.643-1.38)	0.972 (0.761-1.56)	0.068
Stigmasterol to TC x 10^3	0.129 (0.078-0.271)	0.192 (0.101-0.321)	0.029
Sitostanol to TC x 10^3	0.050 (0.032- 0.097)	0.042 (0.027-1.04)	0.953
Cholesterol absorption*	4.99 (3.97-7.06)	5.58 (4.15-8.37)	0.077

*Cholesterol absorption denotes the sum of 5α-cholestanol, β-sitosterol, campesterol, stigmasterol and sitostanol.

Quantitative variables not following normal distribution were expressed as median (interquartile range). ln(5α-cholestanol to TC x 10^3), ln(β-sitosterol to TC x 10^3), ln(campesterol to TC x 10^3), ln(stigmasterol to TC x 10^3), ln(sitostanol to TC x 10^3) and ln(cholesterol absorption) were used to calculate *p* for trend adjusted by confusing factors: body mass index, sex and age.

Supplemental Table 1. Bioinformatic Analysis to pathogenic mutation

Gene	Nucleotide change	Predicted aminoacid change	Bioinformatic Analysis				Frequency in the 1000 Genomes population	Frequency in the ExAc
			SIFT (score)	PredictSNP2* (probability)	Polyphen-2 (score)	Mutation taster (probability)		
ABCG5	rs146534033 c.1733A>G	p.(Asn578Ser)	Tolerated (0.10)	Damaging (87%)	Probably Damaging (0.984)	Disease causing (0.984)	0.000	<0.001
	rs139264483 c.862G>T	p.(Gly288Cys)	Damaging (0.00)	Damaging (87%)	Probably Damaging (0.999)	Disease causing (0.999)	0.000	<0.001
	rs141828689 c.593G>A	p.(Arg198Gln)	Damaging (0.02)	Damaging (87%)	Probably Damaging (1.000)	Disease causing (0.999)	0.00399	0.0023
	rs552803459 c.805G>A	p.(Gly269Arg)	Damaging (0.00)	Damaging (87%)	Probably Damaging (1.000)	Disease causing (0.999)	0.000	<0.001
	c. 887A>G	p.(Asn296Ser)	Damaging (0.02)		Possibly damaging (0.635)	Disease causing (0.999)	-	-
ABCG8	rs376069170 c.1534G>A	p.(Gly512Arg)	Tolerated (0.37)	Damaging (87%)	Possibly damaging (0.682)	Disease causing (0.997)	0.000	0.000

*PredictSNP2 used CADD, DANN, FATHMM and FunSeq2 as bioinformatic predictors.