

A novel mutation in USF1 gene is associated with familial combined hyperlipidemia

Article (Accepted Version)

Taghizadeh, Eskanda, Mirzaei, Farzaneh, Jalilian, Nazanin, Mobarhan, Majid Ghayour, Ferns, Gordon A and Pasdar, Alireza (2019) A novel mutation in USF1 gene is associated with familial combined hyperlipidemia. IUBMB Life. pp. 1-8. ISSN 1521-6543

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/88654/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

A novel mutation in *USF1* gene is associated with familial combined hyperlipidemia

Eskandar Taghizadeh^{1,2}, Farzaneh Mirzaei¹, Majid Ghayour Mobarhan^{3*}, Alireza Pasdar^{1,3,8*},

Gordon A. Ferns

- 1- Departments of Medical Genetics, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran
- 2- Cellular and Molecular Research Center, Yasuj University of Medical Sciences, Yasuj, Iran
- 3- Metabolic Syndrome Research Centre, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran
- 4- Medical Genetics Research Centre, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran
- 5- Division of Applied Medicine, Medical School, University of Aberdeen, Forester hill, Aberdeen, UK

*Drs. Pasdar and Ghayour Mobarhan are corresponding authors to this article

Abstract

Background: Familial combined hyperlipidemia or FCHL is one of the most common genetic causes of hyperlipidemia and is associated with elevation of cholesterol, triglycerides or both, and increased serum apolipoprotein B (apoB). Linkage analysis and next generation sequencing have been successfully used for identifying rare genetic variants that have moderate-to large effects.

Methods: We characterized a large pedigree from a proband identified following recruitment into the MASHAD study, in northeast Iran, with FCHL accompanied by early-onset coronary artery disease. We used linkage analysis for several candidate regions in previous studies such as 1q21-23, 11q23 and 8p, and then whole-exome sequencing to identify the disease-associated gene in this family

Results: we identified a novel variant in the *USF1* gene, leading to a substitution of a tryptophan for arginine at position 196. Arg196Trp co-segregated in all the affected family members in this

pedigree with clinical syndrome, and was not found in any unaffected family members of this pedigree, nor in unrelated controls.

Conclusions: We speculate that this mutation [Arg196Trp] in the USF1 gene might be associated with FCHL and early-onset coronary heart disease in this family. However, the substantial mechanism requires further investigation. These findings indicate that USF1 plays an important role in the biological pathways associated with lipid metabolism

Key words: Familial Combined Hyperlipidemia, FCHL, Cardiovascular Disease, Cholesterol, Triglycerides

Introduction

Familial combined Hyperlipidemia (FCHL) is the most common genetic and metabolic form of hyperlipidemia with prevalence of 0.5% to 2% worldwide and about 10% of these patients suffer from cardiovascular disease. This number has increased by approximately 11.3% in young survivors of myocardial infarction (MI) and by 40% among all the people who survived an MI (1-4). Based on previous studies about 3.5 million people in Europe and 7.2 million people in the US are affected by this disorder(5, 6). Biochemically, FCHL is associated with increased levels of triglyceride, total cholesterol, or both, and increased levels of VLDL, LDL, decreased HDL, and increased apo B level(1, 7, 8). Based on a combination of past and recent definitions, FCHL is now been recognized as a common metabolic disorder, and is characterized in a family by an increased serum triglyceride and LDL in two members of a family accompanied by increased risk of premature coronary artery disease(4, 9). FCHL is genetically complex; in addition to the role of environmental factors, the FCHL is genetically heterogeneous(10, 11).

Three important genetic loci for FCHL have been identified: 1q21-23, apolipoprotein A-I/C-III/A-IV cluster on chromosome 11 and lipoprotein lipase (LPL) gene on 8p that in several different studies had significant LOD scores with FCHL. Pajukanta et al have reported a linkage between FCHL and chromosome 1q markers in a Finnish families(12, 13).

Next-generation sequencing (NGS) has had an important role in identifying the genes and rare variants in a huge number of diseases with moderate-to-large effects(14). NGS can be applied for identification of rare mutations in families with specific extreme phenotypes with high-speed, high-power, and low-cost (15, 16). The linkage analysis technique has been used for many years as the first method of gene discovery and genetic mapping of Mendelian and complex traits which have a familial aggregation (17). Linkage analysis can be used in conjunction with NGS filtering approaches as an important and powerful technique for discovering of some genes involved in diseases etiology(18, 19). In this study, linkage analysis and whole-exome sequencing were used to investigate the members of a large family, with a familial pattern of FCHL associated with coronary artery disease.

Material and methods

1. Study population

The protocol of this study was approved by the local ethics committee of Mashhad University of Medical Sciences and all study participants in present study signed consent form. We identified a family with a recurring familial pattern of FCHL from a cohort as name MASHAD study in northeast Iran, Mashhad who they had an unusual constellation of juvenile-onset FCHL accompanied by an early-onset coronary artery disease and elevated fasting serum triglyceride and cholesterol levels that this was accordance with FCHL diagnosis based on 2016 ESC/EAS

Guidelines for the Management of Dyslipidemias. This large pedigree was notable for having several affected members with FCHL with what appeared to be an autosomal dominant inheritance pattern. Also there were several people in this pedigree who did not have these traits (FIG 1). In this family affected members had a familial clustering and could trace their descent from a common ancestor, that suggested that the affected family members had an autosomal dominant pattern. After evaluation these family members, we collected clinical and laboratory data for all members of this family who were older than 30 years of age. Also we collected 5 mL blood contain EDTA anticoagulant and then, genomic DNA was extracted. For biochemical analysis we collected 5 mL blood without anticoagulant and after 30 minutes we separated serum with centrifuge at 5000 RPM for 10 minutes at room temperature. 5cc blood was obtained, and genomic DNA was extracted. Also we collected 5cc blood from 100 healthy people over 60 years of age for later studies.

STR markers genotyping and linkage analysis

DNA samples from 25 living family members were used for linkage studies, including 12 family member's diagnosed for FCHL disease and 13 unaffected family members. We performed a parametric analysis of linkage with STR markers for 1q21-23, 11q21 and 8P using samples obtained from affected members. Primer sequences are present in UCSC genome browser. A panel of two different STR markers were genotyped for 1q21-23 locus on chromosome 1 including D1S104 and D1S1677. We used the D11S4127 marker for apolipoprotein A-I/C-III/A-IV on chromosome 11 and D8S282 marker for LPL gene that is a candidate gene for FCHL based on past studies. Subjects were considered to affected, or unknown to ignore the reduced penetrance effect. Polymerase chain reactions (PCR) were used under standard conditions optimized for each 4 markers. Details of conditions and concentrations reactions can be provided

by the authors on request. We calculated the LOD score for each marker using the formula below:

$$\log[10] = \frac{L(\theta)}{L(0.5)}$$

Whole exome sequencing

DNA from the index patient from this family was sequenced for mutations by using the WES technique. Whole Exome Sequencing was used to enrich all exons of protein-coding regions as well as some important other genomic regions. NGS was performed using an Illumina Sequencer to sequence approximately 100 million reads. In total, this platform sequenced >95% of the targeted regions with an acceptable sensitivity of >99%. In this approach, duplications, micro-insertion/deletions, and point mutations can be detected. After this we used bioinformatics analysis for sequencing results by using international databases and standard bioinformatics software then, filtration of raw data was performed, to remove common variants that are exist in reference genomes. We also applied filters against some published data in databases and those variants with novelty that had a deleterious effect on proteins, tissue expression were selected for more investigation.

Variant Validation Studies with sanger sequencing

After WES annotation we found a mis-sense mutation in exon 8 USF1 gene which is reported to be an important gene for susceptibility to FCHL in previous studies. We confirmed this variant in the index patient by PCR technology and sanger sequencing technique (PCR primers and their characteristics are shown in table 1). Also we carried out segregation studies in other members of this family who were not applied for WES and polymerase chain reaction was used for sequencing target regions especially the variants of interest (FIG2). The PCR was performed with mastermix solutions contain PCR buffer, MgCL2, deoxyribonucleotide triphosphates, Taq

DNA polymerase and 10 pmol of each primer, 100 ng genomic DNA in a volume of 25 μ l. we set up thermocycler for an initial denaturation at 95 $^{\circ}$ C for 5 min in step 1 followed by 35 cycles in step 2 for an denaturation at 95 $^{\circ}$ C for 1 min, annealing at 59 $^{\circ}$ C for 45 sec, an extension at 72 $^{\circ}$ C for 1 min and a final extension in step 3 at 72 $^{\circ}$ C for 10 min. Then the products of PCR were run and separate on 1% agarose gel and visualized with green viewer. The amplified products were sequenced by sanger sequencing technique and data were analyzed with the help of Snap gene software. In addition, we tested 100 healthy people in over 60 years of age for this variant from general population of this area.

Results

Study population

The clinically characterized members of this family included 13 affected family members with FCHL. All 13 affected family members had CAD at a mean (\pm SE) age of 40 \pm 5 years. All the affected family members had a serum total cholesterol >240 mg/dL, LDL> 160 mg/dL, TG>120 mg/dL and apoB > 120 mg/dL and based on 2016 ESC/EAS Guidelines for the Management of Dyslipidaemias they met the standard definition of a metabolic syndrome called FCHL (Table 2).

Linkage study

We carried out a parametric analysis of linkage with STR markers for 1q21-23, 11q21 and 8P using genotypes obtained from affected and non-affected members of this pedigree. The condition was considered as an autosomal dominant trait. Analysis of linkage showed significant evidence for of 1q21-23 linked to FCHL patients. The maximum LOD score was 3.05 and other loci showed a LOD score less than 1.0. linkage analysis for 1q21-23 in this pedigree represented a founder mutation. No individual in this family was homozygous for this locus. The PCR conditions for STR markers are shown in Table 3

Whole-Exome Sequencing and validation using sanger sequencing

With filtration raw data from WES in the index patient, removing common SNPs and other variants that were found in databases as non-pathogenic, we identified a rare protein-altering variants in the index patient in *USF1* gene on chromosome 1 in 1q21-23 locus which was linked to FCHL in all affected members of this family. This mutation led to a substitution of tryptophan for arginine at position 196 of *USF1* (p. Arg196Trp) (FIG3). segregation genotyping showed a cosegregation of this variant in all affected members. Also these variant was not found in healthy members of this pedigree.

The Arginine 196 of the *USF1* gene is highly conserved between orthologues and paralogues in some species including *Ptrogodytes*, *Mmulatta*, *Fcatus*, *Mmusculus*, *Drerio* and humans. (Fig4). We genotyped 100 ethnically healthy control from matched Iranians for Arg196Trp variant and this mutation was absent in all samples. This variant was not found in databases such as ExAC, gnomAD and iranome. Several programs such as Mutation taster, SIFT, PROVEAN, FATHMM, DANN and FATHMM-MKL predicted the variant to be damaging (Table 4).

Discussion

FCHL is one of the most common type of hyperlipidemia worldwide with increased risk of premature CHD and due to increased VLDL and LDL levels affected persons have an elevation of both cholesterol and triglycerides in the blood.

The present study showed an association between a non-conservative mutation in the *USF1* gene and FCHL patients associated with early-onset CHD in a family with an informative pedigree. The *USF1* locus on 1q21-23 has been linked to FCHL and CHD in some previous studies(12, 20, 21).

USF1 is a transcription factor belong to the bHLH-Zip class, and the domain affected here is important for DNA binding and dimerization. The functional protein forms homo-dimers, or heterodimers with USF2 and through binding at distal E-box elements, activates the transcription of some target genes(22, 23). Thus some post translational modifications such as phosphorylation in USF1 can change transactivation activity(24). This gene can also play a role in chromatin barrier insulator function and protect euchromatin regions from heterochromatin-induced gene silencing(25). Reduced expression of USF1 gene may lead to increased production and reduced metabolism of lipoproteins and plasma lipids because USF1 is involved in regulation of several genes of glucose and lipid metabolism(24). Whole-genome ChIP analysis identified 2518 binding sites for USF1 in HepG2 cells related to chromatin context which were strongly re correlated with expression level of some target genes and this suggested key roles for USF1 gene in transcription activation(26).

We have found that FCHL is linked to 1q21-23 and this is confirmed in some previous linkage analysis such as studies done by Pajukanta et al (12, 27), Huertas-Vazquez (28) and Pei et al(20). Pei et al. investigated familial combined hyperlipidemia (FCHL) families from non-isolated regions in China and Germany to investigate some evidence for linkage to a chromosome 1q locus and they observed a linkage between FCHL patients and 1q21–q23 locus (20). Pajukanta et al. also carried out a linkage analysis using several markers for 10 chromosomal regions and they observed strongly support that a FCHL gene can be on chromosome 1q (12). It seemed that FCHL linkage to 1q21-23 is due to that 1q21-23 locus has several genes associated with lipid metabolism and usually based on recombination roles, it co-segregates in affected members of a pedigree if the genes in this locus are a causes of the diseases.

We carried out WES for one patients of this pedigree and data analysis with more focus on 1q21-23 region showed a missense mutation in exon 8 of *USF1* gene (p. Arg196Trp).

Initially, Pajukanta et al. showed that FCHL is linked to the *USF1* gene and they observed that FCHL is associated with a common haplotype contain non-coding SNPs within the *USF1* gene(27). It was also shown that individuals with the risky allele, have a lack of insulin-induced increase of *USF1* expression in fat tissue and skeletal muscle. The *USF1* gene has an important role in adipose tissue metabolism and through mediating the glucose-regulated expression of hormone-sensitive lipase (HSL) influences de novo lipogenesis by hormone-sensitive lipase, which is an important enzyme in the regulation of lipid source in adipose tissue(29, 30). *USF1* also has an important role in transcription of fatty acid synthase which is involved in the synthesis of fatty acids and can plays an important role in the transcription of several apolipoproteins such as APOCIII, APOAII, and APOE, hepatic lipase, angiotensinogen, glucokinase, and ABCA1 (30, 31). To date, 121 SNPs in the *USF1* gene have been identified in the dbSNP database and only the rs4126997 T>C causes a non-synchronous mutation but data on functional studies and allele frequency are not available for this SNP. rs2073658 A>G in intron 7 and rs3737787 C>T in the 3'-UTR are two functional SNP of *USF1* gene that are in linkage disequilibrium which the minor allele is accompanied with normal expression of *USF1* in fat tissues and human muscles but there is a loss of insulin-induced upregulation of *USF1* mRNA and related target genes of *USF1* gene and as a result insulin-mediated anti-lipolytic activity reduces(26, 32, 33).

As *USF1* gene has an important role in the regulation of numerous genes involved in the metabolism of lipids and glucose, non-responsive *USF1* expression can adversely affect the metabolism of lipids and lipoproteins resulting in increased plasma lipids, thus *USF1* is an ideal

candidate gene for involvement with clinical features of FCHL and hyperlipidemia and some problems associated with CHD such as hypertension and we can say given that *USFI* gene affects the complex lipid phenotype of familial combined hyperlipidemia. Arg196Trp mutation *USFI* gene might be associated with FCHL and early-onset coronary heart disease in this family and can be a cause for related high triglycerides, ApoB, total cholesterol but for further confirmation it is necessary to undertake a functional study for this variant and can be later confirmed in controlled studies in animal models.

Conclusion

We have confirmed that 1q21-23 locus is linked to FCHL in an Iranian pedigree, and a new mutation in *USFI* gene which is within this locus may be related with some clinical characteristics which is specified by elevation cholesterol, triglyceride, CAD. Therefore, our results indicate that *USFI* gene plays an important role in some biologic pathways associated with lipid metabolism.

Conflict of interest

No potential conflict of interest relevant to this article was reported.

Acknowledgment

Supported by grants from the Mashhad University of Medical Sciences (grant number: 961251)

References

1. Gaddi A, Cicero A, Odo F. Practical guidelines for familial combined hyperlipidemia diagnosis: an up-date. *Vascular health and risk management*. 2007;3(6):877.
2. Shoulders C, Jones E, Naoumova R. Genetics of familial combined hyperlipidemia and risk of coronary heart disease. *Human molecular genetics*. 2004;13(suppl_1):R149-R60.
3. Cortner JA, Coates PM, Gallagher PR. Prevalence and expression of familial combined hyperlipidemia in childhood. *The Journal of pediatrics*. 1990;116(4):514-9.
4. Bello-Chavolla OY, Kuri-García A, Ríos-Ríos M, Vargas-Vázquez A, Cortés-Arroyo JE, Tapia-González G, et al. Familial Combined Hyperlipidemia: Current Knowledge, Perspectives, and Controversies. *Revista de Investigación Clínica*. 2018;70(5):224-36.

5. Taghizadeh E, Mardani R, Rostami D, Taghizadeh H, Bazireh H, Hayat SMG. Molecular mechanisms, prevalence, and molecular methods for familial combined hyperlipidemia disease: A review. *Journal of cellular biochemistry*. 2019;120(6):8891-8.
6. Minicocci I, Prisco C, Montali A, Di Costanzo A, Ceci F, Pigna G, et al. Contribution of mutations in low density lipoprotein receptor (LDLR) and lipoprotein lipase (LPL) genes to familial combined hyperlipidemia (FCHL): a reappraisal by using a resequencing approach. *Atherosclerosis*. 2015;242(2):618-24.
7. Goldstein JL, Schrott HG, Hazzard WR, Bierman EL, Motulsky AG. Hyperlipidemia in coronary heart disease II. Genetic analysis of lipid levels in 176 families and delineation of a new inherited disorder, combined hyperlipidemia. *The Journal of clinical investigation*. 1973;52(7):1544-68.
8. Eckel RH. Familial combined hyperlipidemia and insulin resistance: distant relatives linked by intra-abdominal fat? : *Am Heart Assoc*; 2001.
9. Veerkamp MJ, de Graaf J, Hendriks JC, Demacker PN, Stalenhoef AF. Nomogram to diagnose familial combined hyperlipidemia on the basis of results of a 5-year follow-up study. *Circulation*. 2004;109(24):2980-5.
10. Eurlings PM, van der Kallen CJ, Geurts JM, van Greevenbroek MM, de Bruin TW. Genetic dissection of familial combined hyperlipidemia. *Molecular genetics and metabolism*. 2001;74(1-2):98-104.
11. Taghizadeh E, Esfehiani RJ, Sahebkar A, Parizadeh SM, Rostami D, Mirinezhad M, et al. Familial combined hyperlipidemia: An overview of the underlying molecular mechanisms and therapeutic strategies. *IUBMB Life*.
12. Pajukanta P, Nuotio I, Terwilliger JD, Porkka KV, Ylitalo K, Pihlajamäki J, et al. Linkage of familial combined hyperlipidaemia to chromosome 1q21–q23. *Nature genetics*. 1998;18(4):369.
13. Coon H, Myers RH, Borecki IB, Arnett DK, Hunt SC, Province MA, et al. Replication of linkage of familial combined hyperlipidemia to chromosome 1q with additional heterogeneous effect of apolipoprotein AI/C-III/A-IV locus: the NHLBI family heart study. *Arteriosclerosis, thrombosis, and vascular biology*. 2000;20(10):2275-80.
14. Teer JK, Mullikin JC. Exome sequencing: the sweet spot before whole genomes. *Human molecular genetics*. 2010;19(R2):R145-R51.
15. Agrawal N, Frederick MJ, Pickering CR, Bettgowda C, Chang K, Li RJ, et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science*. 2011;333(6046):1154-7.
16. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature genetics*. 2011;43(5):491.
17. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. *American journal of human genetics*. 1996;58(6):1347.
18. Ott J, Wang J, Leal SM. Genetic linkage analysis in the age of whole-genome sequencing. *Nature Reviews Genetics*. 2015;16(5):275.
19. Lathrop GM, Lalouel J, Julier C, Ott J. Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. *American journal of human genetics*. 1985;37(3):482.

20. Pei W, Baron H, Müller-Myhsok B, Knoblauch H, Ali Al-Yahyaee S, Hui R, et al. Support for linkage of familial combined hyperlipidemia to chromosome 1q21–q23 in Chinese and German families. *Clinical genetics*. 2000;57(1):29-34.
21. Huertas-Vázquez A, Del Rincon J, Canizales-Quinteros S, Riba L, Vega-Hernández G, Ramírez-Jiménez S, et al. Contribution of Chromosome 1q21-q23 to Familial Combined Hyperlipidemia in Mexican Families. *Annals of human genetics*. 2004;68(5):419-27.
22. Jaiswal AS, Narayan S. Upstream stimulating factor-1 (USF1) and USF2 bind to and activate the promoter of the adenomatous polyposis coli (APC) tumor suppressor gene. *Journal of cellular biochemistry*. 2001;81(2):262-77.
23. Rada-Iglesias A, Ameer A, Kapranov P, Enroth S, Komorowski J, Gingeras TR, et al. Whole-genome maps of USF1 and USF2 binding and histone H3 acetylation reveal new aspects of promoter structure and candidate genes for common human disorders. *Genome research*. 2008;18(3):380-92.
24. Corre S, Galibert MD. Upstream stimulating factors: highly versatile stress-responsive transcription factors. *Pigment cell research*. 2005;18(5):337-48.
25. Huang S, Li X, Yusufzai TM, Qiu Y, Felsenfeld G. USF1 recruits histone modification complexes and is critical for maintenance of a chromatin barrier. *Molecular and cellular biology*. 2007;27(22):7991-8002.
26. Verhoeven AJ. USF1 (upstream transcription factor 1). *Atlas of Genetics and Cytogenetics in Oncology and Haematology*. 2011.
27. Pajukanta P, Lilja HE, Sinsheimer JS, Cantor RM, Lusis AJ, Gentile M, et al. Familial combined hyperlipidemia is associated with upstream transcription factor 1 (USF1). *Nature genetics*. 2004;36(4):371.
28. Plaisier CL, Horvath S, Huertas-Vazquez A, Cruz-Bautista I, Herrera MF, Tusie-Luna T, et al. A systems genetics approach implicates USF1, FADS3, and other causal candidate genes for familial combined hyperlipidemia. *PLoS genetics*. 2009;5(9):e1000642.
29. Laurila P-P, Soronen J, Kooijman S, Forsström S, Boon MR, Surakka I, et al. USF1 deficiency activates brown adipose tissue and improves cardiometabolic health. *Science translational medicine*. 2016;8(323):323ra13-ra13.
30. van der Vleuten GM, Isaacs A, Hijmans A, van Duijn CM, Stalenhoef AF, de Graaf J. The involvement of upstream stimulatory factor 1 in Dutch patients with familial combined hyperlipidemia. *Journal of lipid research*. 2007;48(1):193-200.
31. Ribeiro A, Pastier D, Kardassis D, Chambaz J, Cardot P. Cooperative binding of upstream stimulatory factor and hepatic nuclear factor 4 drives the transcription of the human apolipoprotein A-II gene. *Journal of Biological Chemistry*. 1999;274(3):1216-25.
32. Kantartzis K, Fritsche A, Machicao F, Stumvoll M, Machann J, Schick F, et al. Upstream transcription factor 1 gene polymorphisms are associated with high antilipolytic insulin sensitivity and show gene–gene interactions. *Journal of Molecular Medicine*. 2007;85(1):55-61.
33. Söderlund S, Lyssenko V, Vaag A, Poulsen P, Groop L, Taskinen M-R, et al. Functional Variant Disrupts Insulin Induction of USF1. 2009.