Original Article

Genetic Diagnosis via Whole Exome Sequencing in Taiwanese Patients with Hypertriglyceridemia

Kuan-Rau Chiou^{1,2}, Chung-Yung Chen³ and Min-Ji Charng^{2,4}¹Division of Cardiology, Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan²School of Medicine, National Yang-Ming University, Taipei, Taiwan³Department of Bioscience Technology, Chung Yuan Christian University, Chung-Li, Taiwan⁴Division of Cardiology, Taipei Veterans General Hospital, Taipei, Taiwan

Aim: Whole exome sequencing (WES) is a recently developed method for discovering rare mutations associated with hereditary disorders. However, the feasibility and utilization of this method in identifying familial hypertriglyceridemia is not well known. The purpose of the study was to identify the genetic locus that causes hypertriglyceridemia and assess its prevalence in Taiwanese subjects with hypertriglyceridemia.

Methods: We performed WES among two individuals with hypertriglyceridemia and one control subject in an index family (22 members). Based on the WES findings, we extended the study to genotype 65 unrelated adult index patients with a fasting serum triglyceride level of >500 mg/dL and 125 normal controls using polymerase chain reaction.

Results: Using WES alignment, variant calling and annotation, 15 presumptive causal variants were initially identified, including 13 cases by the autosomal dominant model and two cases by the autosomal recessive model. Only *APOA5* c.553 G>T (rs2075291), resulting in the amino acid mutation Gly185Cys, co-segregated well with hypertriglyceridemia in terms of autosomal recessive inheritance (homozygote TT: mean triglyceride level: 1,071 mg/dL vs non TT (GT and GG): mean triglyceride level: 118 mg/dL; $p < 0.001$) in the index family. In the unrelated cohorts, the frequency of the TT genotype of rs2075291 was 12.3% in the hypertriglyceridemic group; however, no TT genotype was found in the control group.

Conclusions: Our results demonstrate that WES is feasible for identifying the genetic locus that causes hypertriglyceridemia. The TT genotype of *APOA5* c.553G>T acts as an important indicator of hypertriglyceridemia in patients in Taiwan.

See editorial vol. 22: 881-885

J Atheroscler Thromb, 2015; 22: 887-900.

Key words: Apolipoprotein A5, Exome sequencing, Triglycerides, Taiwanese; Chinese

Introduction

A very high triglyceride (TG) level is defined as a fasting serum TG level of >500 mg/dL and is prevalent in approximately 1% of the adult population

Address for correspondence: Min-Ji Charng, Division of Cardiology, Department of Medicine, Taipei Veterans General Hospital, No. 201, Sec. 2, Shih-Pai Road, Taipei, 112, Taiwan
E-Mail: mjcharng@vghtpe.gov.tw

Received: January 19, 2015

Accepted for publication: February 17, 2015

worldwide^{1, 2}). This finding is a well-established cardiovascular risk factor that is independent of high-density lipoprotein cholesterol (HDL-C). Recent progress in human genetics has furthered our understanding of the genetic susceptibility to hypertriglyceridemia (HTG), which can be attributed to either monogenic dyslipidemia or complex diseases with strong polygenic components^{3, 4}). Currently, more than 30 candidate genes have been reported to be associated with HTG³⁻¹³). The prevalence of each causative gene is different among multiethnic populations, and the

differences in functional variants are specific to ethnic groups. The expanding list of disease loci makes confirming the genetic diagnosis in cases of HTG challenging. Traditionally, *ad hoc* screening for causative mutations in each candidate gene has been performed using a Sanger-based sequencing strategy. This approach is used to screen a number of genes, including the *LPL*, *APOC2*, *LMF1*, *APOA5* and *GPIHBP1* genes, followed by *AGPAT2*, *APOE*, *PPARG*, *AKT2* and *USF1*. At present, screening multiple genes is time-consuming and expensive and is a limitation to the identification of HTG-associated loci for personalized health management. Whole exome sequencing (WES) has emerged as a successful tool for detecting single-nucleotide substitutions and insertions as well as small deletions up to 10 nucleotides¹⁴. This technique has enabled the novel identification and molecular characterization of genetic defects in a wide spectrum of rare and genetically heterogeneous Mendelian disorders¹⁴⁻¹⁷. However, the feasibility and utilization of this method for identifying familial HTG is not well known¹⁸⁻²⁰. Accordingly, the aims of this study were: 1) to test the hypothesis that WES can be used to identify the causative gene in patients with clinical manifestations of the familial form of HTG and 2) to use these findings to explore the prevalence of adults with very high TG levels among Taiwanese²¹.

Methods

Study Design and Subjects

First, WES was performed to identify the causative gene in two affected individuals and a normal individual from the index family. Second, allelic discrimination of the causative gene was identified using polymerase chain reaction (PCR) in an extended cohort consisting of 65 unrelated adult patients with very high TG levels, namely an overnight fast-serum TG level of >500 mg/dL, after a 12-week low-fat diet and the exclusion of the following factors: diabetes mellitus, hypothyroidism, liver and kidney disease, drugs known to interfere with lipid metabolism, alcohol drinking, a body mass index (BMI) of >30 kg/m² or a age younger than 18 years. The 125 age/sex-matched control group included healthy, non-smoking subjects with a serum TG level below 150 mg/dL. The purpose of the study was carefully explained to all participants, and written informed consent was obtained prior to participation. The Institutional Review Board of our hospital provided ethics approval of the study protocol, which was carried out according to the principles of the Helsinki Declaration.

Biochemical Analyses and DNA Extraction

Blood samples were collected after an overnight fast for the lipid analysis. The total serum cholesterol, TG, HDL-C and low-density lipoprotein cholesterol (LDL-C) levels were determined using commercial kits (Wako, Osaka, Japan) on a Hitachi 7600 clinical analyzer (Hitachi High-Tech, Tokyo, Japan). If the TG level was >1,000 mg/dL, the quantitative determination of the HDL-C and LDL-C levels was also corrected for measurements obtained using commercial kits (SPIFE Vis Cholesterol System, Beaumont, Texas) on a Helena SPIFE 3000 device (Helena Laboratories, Beaumont, Texas). Genomic DNA of leukocytes was extracted using the DNA Midi Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

Whole Exome Sequencing, Alignment, Variant Calling and Annotation

Two affected individuals (I-1, I-3) and a normal control subject (I-4) from distinct branches of the family (see **Supplementary Fig. 1**) were selected for exome sequencing. DNA was fragmented and enriched for coding exons using the Illumina TruSeq Exome Enrichment Kit for samples (Illumina Inc., San Diego, CA) to generate exome libraries. The libraries were sequenced on an Illumina GAIIX (Illumina Inc., San Diego, CA) with 101-bp paired-end reads, according to the manufacturer's instructions. The image analysis and base calling were processed with the Sequence Control Software with Real-Time Analysis (Illumina) and CASAVA software (Illumina) programs.

The paired sequencing reads were first aligned to the human reference genome reference assembly GRCh37 using a Burrows-Wheeler Aligner²². The post-alignment process, including SAM conversion, sorting and PCR duplicate removal, was performed using Picard (<http://picard.sourceforge.net/>) and SAMtools (<http://samtools.sourceforge.net/>). The Genome Analysis Toolkit (GATK) was then used to process local realignment around small insertions-deletions (INDELs) and for Base Quality Score Recalibration²³. For each sample, a processed alignment file was saved in bam format.

In order to identify Single Nucleotide Variants (SNVs) and INDELs, we subsequently performed GATK pipeline, followed by Variant Quality Score Recalibration. The SNVs and INDELs were then annotated via the Variant Effect Predictor to record, for example, the position in the genes and amino acid changes relative to the human Refseq genome, the consequence of the mutation (i.e., non-synonymous,

Fig. 1. Pedigree of the index family. Square symbols indicate men; circles, women; X mark, death. Filled symbols indicate homozygosity (rs2075291 TT). Half-filled symbols indicate heterozygosity (rs2075291 TG). The pedigree number, age, body mass index and lipoprotein levels are indicated below each symbol. The lipid profiles are indicated in the following order: triglycerides, high-density lipoprotein cholesterol, low-density cholesterol and total cholesterol. No data were available for individuals with an asterisk.

nonsense, etc.). All added annotations were saved in a database including dbSNP, PolyPhen2, SIFT and COSMIC²⁴⁻²⁷. Then, the variants were acquired based on the following criteria: 1) a different genotype between the groups; 2) both a SIFT score of ≤ 0.05 and Polyphen-2 score of ≥ 0.2 ; 3) a minor allele frequency of ≤ 0.05 from the 1,000 Genome ASN population; and 4) containing Variant Clinical Significance annotations.

Both inheritance patterns, namely autosomal dominant vs. autosomal recessive, were analyzed. Candidate genes were acquired using DAVID (<http://david.abcc.ncifcrf.gov/>) and Gestalt (<http://bioinfo.vanderbilt.edu/webgestalt/>), which allowed the genes to be clustered based on function and different gene ontologies. Moreover, the Pathway Interaction Database (PID) (<http://pid.nci.nih.gov/>) and IPA (Ingenuity® Systems, www.ingenuity.com) defined the genes based on the pathway interactions.

Genotyping

In order to confirm the causative mutations identified using the WES analysis, the causative gene was amplified via PCR and sequenced according to Sanger sequencing (3130xL Genetic Analyzer, Applied Bio-

systems, Foster City, CA) using DNA collected from index family members. For the analysis of the genotype of the causative gene, allelic discrimination of the SNPs was performed in individuals from the index family and the extended cohort.

Statistical Analysis

The frequencies of the alleles were calculated with gene counting. The statistical analysis was performed using the SPSS software program version 10.0 for Windows (SPSS Inc., Chicago, USA). Genotype frequencies with Hardy-Weinberg equilibrium expectations were tested using Fisher's exact test. Differences in quantitative variables were calculated according to the Student's unpaired *t*-test, and the difference in the serum triglyceride levels among each genotype was tested using an analysis of variance (ANOVA). All *p* values were two-tailed, and *p* values of < 0.05 were considered to be statistically significant.

Results

Characteristics and Exome Sequencing of the Index Family

The characteristics of the index family are shown

Table 1. Overall reads, coverage statistics and identified variants

Sample	I-1	I-3	I-4
Reads statistics			
Total read count	201589662	202065903	237775313
Mapping read count	195608711	196710642	231805932
Unmapping read count	5980951	182394839	214594720
Coverage statistics			
% Target bases ≥ 1 -fold Coverage	0.96	0.96	0.96
% Target bases ≥ 10 -fold Coverage	0.93	0.93	0.93
% Target bases ≥ 20 -fold Coverage	0.91	0.91	0.91
% Target bases ≥ 100 -fold Coverage	0.57	0.64	0.66
Mean coverage for target bases	138.4	155.65	165.36
Identified variants			
Total no.	69661	69131	69692
and known variants	56420	55841	56512
and in dbSNP138	55613	55000	55692
and overlapped genes	23325	23361	23436

in **Fig. 1** and included age, sex and lipid profiles available for 22 members. Members of the index family did not have any secondary causes of HTG, were healthy and did not suffer from cardiovascular events or acute pancreatitis. The TG levels in the family pedigree fit the clinical diagnosis of the familial form of HTG (**Fig. 1**). All four HTG members received TG-lowering medications (fenofibrate at 200 mg and acipimox at 500 mg per day), with a post-treatment TG level of around 220–300 mg/dL. Three DNA samples, two from HTG patients (I-1, I-3) and one from a normal family member (I-4), were examined using WES. The coverage and coverage statistics are summarized in **Table 1**. All variations were compared with those available in SNP databases, resulting in the identification of 55613, 55000 and 55692 SNPs for I-1, I-3 and I-4, respectively. First, we used the autosomal dominant model to analyze the variants. A total of 66 SNPs were initially detected, including 57 known SNPs and nine *de novo* variations (see **Supplementary Table 1**). Under the IPA Gene View, which demonstrated the protein function, biologic process, canonical pathway and top findings from the Ingenuity Knowledge Base, genes having a possible direct/indirect interaction with either lipids, endocrine factors, energy and/or metabolism, were selected as candidate genes. In total, 13 candidate genes were acquired (see **Supplementary Table 2**). In addition, these genes were analyzed using IPA; the prediction network is shown in **Supplementary Fig. 2**. None of the genes co-segregated with HTG in the index family. Second, we used the autosomal recessive model to analyze the variants, and two missense variants were detected

between the groups (see **Supplementary Table 3**). Only one variant, *APOA5* c.553G>T (rs2075291), which resulted in the amino acid substitution Gly185Cys, co-segregated with HTG in the affected family members (homozygote TT, mean triglyceride level: 1,071 mg/dL vs non TT (GT and GG): mean triglyceride level: 118 mg/dL; $p < 0.001$ (**Fig. 1**). Additional variants were not identified in any of the other candidate genes. Finally, Sanger sequencing of genomic DNA from all individuals in the index family confirmed that missense mutations in *APOA5* exon four co-segregated with HTG via an autosomal recessive mode of inheritance.

rs2075291 Genotyping in the Extended Cohort

A further survey of the prevalence of severe HTG adult patients carrying the *APOA5* c.553G>T mutation was performed. **Table 2** shows the general characteristics and genotypes of the index patients with severe HTG (65) and the age-matched control (125) subjects. The frequencies of the GG, GT and TT genotypes were 55.4%, 32.3% and 12.3%, respectively, in the HTG group. In the control group, the frequencies were 87.2% and 12.8% for the GG and GT genotypes, respectively. The TT genotype was not found in any individuals in the control group.

Discussion

Using WES, we confirmed that the *APOA5* TT genotype (rs2075291) is a causative mutation that co-segregates with HTG in an index family. The prevalence of this genotype was also high in the extended

Table 2. Characteristics and genotypes of the HTG patients and control subjects in the extended cohort

	HTG (<i>n</i> =65)	Control (<i>n</i> =125)	<i>p</i>
Age	54 ± 13	58 ± 15	0.109
Gender (male/female)	43/22	72/53	0.277
BMI (kg/m ²)	26.42 ± 2.53	24.03 ± 2.81	<0.001
History of CAD	5	0	0.006
History of pancreatitis	2	0	0.136
TG (mg/dL)	635 ± 196	105 ± 35	<0.001
TC (mg/dL)	246 ± 43	184 ± 35	0.004
HDL-C (mg/dL)	35 ± 9	52 ± 15	<0.001
LDL-C (mg/dL)	102 ± 31	116 ± 20	<0.001
Genotypes			<0.001
TT	8	0	
GT + GG	57	125	

All values are expressed as the mean ± SEM. *P* values were calculated according to the *t*-test

HTG: hypertriglyceridemia, BMI: body mass index, CAD: coronary artery disease, TG: triglyceride, TC: total cholesterol, HDL-C: high-density lipoprotein cholesterol, LDL-C, low-density cholesterol.

cohort. In contrast, none of the unaffected family members or normal individuals in the control group had the TT genotype in the of rs2075291 variant of *APOA5*.

Our findings demonstrate the effectiveness of WES in providing a molecular diagnosis of familial HTG. We performed a WES analysis to selectively sequence only the coding regions and neighboring intron-exon boundaries of the individual human genome in order to identify the causative mutation. Basing our analysis on an autosomal recessive mode of inheritance, the *APOA5* c.553G > T (rs2075291) SNP variant was found to be co-segregated with HTG (**Fig. 1**). According to the autosomal dominant mode of inheritance, 13 SNPs (**Supplementary Table 2**) were presumptive candidate genes. The prediction network (**Supplementary Fig. 2**) showed that APOB might be related to TNF- α , P2RY6 and NFkB, located in the extracellular region, plasma membrane and nucleus, respectively. After confirming the genotypes of all index family members using Sanger sequencing, we found that none of the 13 SNPs co-segregated with HTG.

A genetic diagnosis is the gold standard for identifying familial HTG. Techniques for obtaining a genetic diagnosis are clinically relevant, especially for screening the family members of affected individuals. However, each technique has a unique set of technical, financial and throughput limitations. Although a genome-wide association study (GWAS) can be used to identify DNA sequences as determinants of the plasma TG concentration^{4, 11}, whether these sequences lie within genomic regions containing classically estab-

lished genes and/or previously uncharacterized regions, this technique cannot be employed to assign causation to a particular gene or variant. Complementary experimental approaches, including rare variant sequencing, animal models studies, linkage and family-based studies and functional cellular and biochemical experiments, are required to more fully elucidate the genetic architecture of the plasma TG concentration.

Currently, WES can be utilized to detect rare variants in patients with a phenotype that is suspected to be a Mendelian (single-gene) genetic disorder, after known single-gene candidates have been eliminated or when testing multiple genes is prohibitively expensive¹². The limitation of WES is that this technique only allows for the identification of those variants found in coding regions and that affect the protein function. WES does not allow for the identification of structural and non-coding variants associated with the disease. These types of variants can be detected using other methods, such as whole genome sequencing, as a back-up measure if the results of exome sequencing are negative. With continuing improvements and decreasing costs, WES or targeted WES may have advantages over current Sanger-based clinical re-sequencing strategies for patients with HTG.

The human *APOA5* gene is located within the *APOA1/C3/A4/A5* gene cluster, a short region on chromosome 11q23-q24, and is a member of the apolipoprotein gene family. APOA5 has previously been identified to be a key regulator of the serum TG concentration^{28, 29}. APOA5 indirectly regulates TG metabolism by stimulating the lipoprotein lipase (LPL) activity, which increases the rate of LPL-mediated TG

hydrolysis³⁰⁾. In previous genotyping and haplotyping studies, the *APOA5* gene was reported to be associated with an elevated TG level; however, the clinical significance and prevalence varied among different ethnicities³¹⁻³⁹⁾. In Taiwan, two haplotypes of *APOA5* (rs662799 and rs2075291) have been found to be closely related to HTG³⁹⁾. Moreover, the TT genotype of rs2075291 in the Taiwanese population was reported to be associated with HTG by Kao J. *et al.* In that report, the frequency of the TT genotype of rs2075291 was reported to be 8.7% in the HTG group⁴⁰⁾. Based on these previous studies, SNP rs2075291 (553G>T) is closely associated with HTG. Hung *et al.* used site-directed mutagenesis to introduce C553G>T *in vitro*, which reduced the activation of LPL-mediated hydrolysis of very low-density lipoprotein to levels that ranged from 31% to 81% of wild-type APOA5⁴¹⁾. Even with these data, it is still uncertain whether the SNP rs2075291 (C553G>T) is a definitive causative mutation. According to our WES findings, *APOA5* C553G>T (rs2075291) is a causative gene mutation. Additionally, in agreement with the findings of a previous report⁴⁰⁾, the prevalence of the TT genotype (rs2075291) is high in affected Taiwanese populations. According to the ExAC browser (<http://exac.broadinstitute.org/>), which contains exome sequence data for 61,486 individuals as a global “reference set,” there are 26 individuals with homozygous mutations with rs2075291 (allele frequency: 0.00637). However, in the reference set, individuals with severe pediatric diseases were removed. In our study, the enrolled patients had very high TG levels and were adults (>18 years old). The homozygous mutation may predispose adults, not children, to a very high TG level. Additionally, the majority of homozygous mutations were found in East Asians, not in the Caucasian population, in the reference set. The higher frequency of this mutation in our extended cohort highlights the importance of primary HTG in East Asians, including Japanese, Chinese, Koreans, etc. More studies are needed to better understand these issues in terms of clinical practice.

The following limitations of the study should be acknowledged: First, this study did not measure the plasma APOA5 level, LPL mass or post-heparin LPL activity in the index family members or extended cohort, which provides a direct link to this mutation in TG metabolism. Additional investigations are needed to understand the role of the APOA5 protein level and LPL lipase mass/activity modulated by this mutation. Second, besides APOA5 c.553 G>T, mutations of other genes, such as LPL, apo C-II and so on, which were not investigated in this study, may also have contributed to the development of a high trigly-

eride level in the 65 primary HTG patients.

In conclusion, this study demonstrated that WES is feasible for identifying the genetic locus that causes HTG. Furthermore, the TT genotype of *APOA5* c.553G>T acts as an important indicator for recognizing hypertriglyceridemic patients in Taiwan. Therefore, continuing efforts to investigate regional founder effects and gain insights into the evolution of the TT genotype (rs2075291) in East Asians would be worthwhile.

Acknowledgments

The authors thank Hsing-Yi Liu for her excellent technological assistance. This work was supported by the National Science Council of Taiwan (NSC 100-2314-B-075-078 and NSC 101-2314-B-075 -060).

Conflicts of Interest

No conflicts of interest, financial or otherwise, are declared by the authors.

References

- 1) Triglycerides and cardiovascular disease: a scientific statement from the American Heart Association. Miller M, Stone NJ, Ballantyne C, Bittner V, Criqui MH, Ginsberg HN, Goldberg AC, Howard WJ, Jacobson MS, Kris-Etherton PM, Lennie TA, Levi M, Mazzone T, Pennathur S; American Heart Association Clinical Lipidology, Thrombosis, and Prevention Committee of the Council on Nutrition, Physical Activity, and Metabolism; Council on Arteriosclerosis, Thrombosis and Vascular Biology; Council on Cardiovascular Nursing; Council on the Kidney in Cardiovascular Disease. *Circulation*, 2011; 123: 2292-2333
- 2) Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. *Circulation*, 2002; 106: 3143-3421
- 3) Evans D, Aberle J, Beil FU: The relative importance of common and rare genetic variants in the development of hypertriglyceridemia. *Expert Rev Cardiovasc Ther*, 2011; 9: 637-644
- 4) Johansen CT, Kathiresan S, Hegele RA: Genetic determinants of plasma triglycerides. *J Lipid Res*, 2011; 52: 189-206
- 5) Martín-Campos JM, Julve J, Roig R, Martínez S, Errico TL, Martínez-Couselo S, Escolà-Gil JC, Méndez-González J, Blanco-Vaca F: Molecular analysis of chylomicronemia in a clinical laboratory setting: diagnosis of 13 cases of lipoprotein lipase deficiency. *Clin Chim Acta*, 2014; 429: 61-68
- 6) Sprecher DL, Taam L, Gregg RE, Fojo SS, Wilson DM,

- Kashyap ML, Brewer HB Jr: Identification of an apoC-II variant (apoC-II^{Bethesda}) in a kindred with apoC-II deficiency and type I hyperlipoproteinemia. *J Lipid Res*, 1988; 29: 273-278
- 7) Péterfy M, Ben-Zeev O, Mao HZ, Weissglas-Volkov D, Aouizerat BE, Pullinger CR, Frost PH, Kane JP, Malloy MJ, Reue K, Pajukanta P, Doolittle MH: Mutations in LMF1 cause combined lipase deficiency and severe hypertriglyceridemia. *Nat Genet*, 2007; 39: 1483-1487
 - 8) Beigneux AP, Franssen R, Bensadoun A, Gin P, Melford K, Peter J, Walzem RL, Weinstein MM, Davies BS, Kuivenhoven JA, Kastelein JJ, Fong LG, Dallinga-Thie GM, Young SG: Chylomicronemia with a mutant GPIHBP1 (Q115P) that cannot bind lipoprotein lipase. *Arterioscler Thromb Vasc Biol*, 2009; 29: 956-962
 - 9) Surendran RP, Visser ME, Heemelaar S, Wang J, Peter J, Defesche JC, Kuivenhoven JA, Hosseini M, Péterfy M, Kastelein JJ, Johansen CT, Hegele RA, Stroes ES, Dallinga-Thie GM: Mutations in LPL, APOC2, APOA5, GPIHBP1 and LMF1 in patients with severe hypertriglyceridaemia. *J Intern Med*, 2012; 272: 185-196
 - 10) Mendoza-Barberá E, Julve J, Nilsson SK, Lookene A, Martín-Campos JM, Roig R, Lechuga-Sancho AM, Sloan JH, Fuentes-Prior P, Blanco-Vaca F: Structural and functional analysis of APOA5 mutations identified in patients with severe hypertriglyceridemia. *J Lipid Res*, 2013; 54: 649-661
 - 11) Johansen CT, Hegele RA: Allelic and phenotypic spectrum of plasma triglycerides. *Biochim Biophys Acta*, 2012; 1821: 833-842
 - 12) Nakamura T, Suehiro T, Yasuoka N, Yamamoto M, Ito H, Yamano T, Hashimoto K: A novel nonsense mutation in exon 1 and a transition in intron 3 of the lipoprotein lipase gene. *J Atheroscler Thromb*, 1996; 3: 17-24
 - 13) Yoshida T, Gotoda T, Okubo M, Iizuka Y, Ishibashi S, Kojima T, Murakami T, Murase T, Yamada N: A Japanese patient with lipoprotein lipase deficiency homozygous for the Gly188Glu mutation prevalent worldwide. *J Atheroscler Thromb*, 2000; 7: 45-49
 - 14) Biasecker LG, Green RC: Diagnostic clinical genome and exome sequencing. *N Engl J Med*, 2014; 370: 2418-2425
 - 15) Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, Shendure J: Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet*, 2011; 12: 745-755
 - 16) Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, Gildersleeve HI, Beck AE, Tabor HK, Cooper GM, Mefford HC, Lee C, Turner EH, Smith JD, Rieder MJ, Yoshiura K, Matsumoto N, Ohta T, Niikawa N, Nickerson DA, Bamshad MJ, Shendure J: Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet*, 2010; 42: 790-793
 - 17) Lupski JR, Reid JG, Gonzaga-Jauregui C, Rio Deiros D, Chen DC, Nazareth L, Bainbridge M, Dinh H, Jing C, Wheeler DA, McGuire AL, Zhang F, Stankiewicz P, Halperin JJ, Yang C, Gehman C, Guo D, Irikat RK, Tom W, Fantin NJ, Muzny DM, Gibbs RA: Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med*, 2010; 362: 1181-1191
 - 18) Rosenthal EA, Ranchalis J, Crosslin DR, Burt A, Brunzell JD, Motulsky AG, Nickerson DA; NHLBI GO Exome Sequencing Project, Wijsman EM, Jarvik GP: Joint linkage and association analysis with exome sequence data implicates SLC25A40 in hypertriglyceridemia. *Am J Hum Genet*, 2013; 93: 1035-1045
 - 19) Johansen CT, Dubé JB, Loyzer MN, MacDonald A, Carter DE, McIntyre AD, Cao H, Wang J, Robinson JF, Hegele RA: LipidSeq: a next-generation clinical resequencing panel for monogenic dyslipidemias. *J Lipid Res*, 2014; 55: 765-772
 - 20) Gonzaga-Jauregui C, Mir S, Penney S, Jhangiani S, Midgen C, Finegold M, Muzny DM, Wang M, Bacino CA, Gibbs RA, Lupski JR, Kellermayer R, Hanchard NA: Whole-exome sequencing reveals GPIHBP1 mutations in infantile colitis with severe hypertriglyceridemia. *J Pediatr Gastroenterol Nutr*, 2014; 59: 17-21
 - 21) Lin M, Chu CC, Chang SL, Lee HL, Loo JH, Akaza T, Juji T, Ohashi J, Tokunaga K: The origin of Minnan and Hakka, the so-called "Taiwanese", inferred by HLA study. *Tissue Antigens*, 2001; 57: 192-199
 - 22) Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 2009; 25: 1754-1760
 - 23) DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernysky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ: A framework or ariation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*, 2011; 43: 491-498
 - 24) McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F: Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics*, 2010; 26: 2069-2070
 - 25) Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR: A method and server for predicting damaging missense mutations. *Nat Methods*, 2010; 7: 248-249
 - 26) Kumar P, Henikoff S, Ng PC: Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*, 2009; 4: 1073-1081
 - 27) Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, Jia M, Shepherd R, Leung K, Menzies A, Teague JW, Campbell PJ, Stratton MR, Futreal PA: COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res*, 2011; 39: D945-D950
 - 28) Elosua R, Ordovas JM, Cupples LA, Lai CQ, Demissie S, Fox CS, Polak JF, Wolf PA, D'Agostino RB Sr, O'Donnell CJ: Variants at the APOA5 locus, association with carotid atherosclerosis, and modification by obesity: the Framingham Study. *J Lipid Res*, 2006; 47: 990-996
 - 29) Pennacchio LA, Olivier M, Hubacek JA, Cohen JC, Cox DR, Fruchart JC, Krauss RM, Rubin EM: An apolipoprotein influencing triglycerides in humans and mice revealed by comparative sequencing. *Science*, 2001; 294: 169-173
 - 30) Schaap FG, Rensen PC, Voshol PJ, Vriens C, van der Vliet HN, Chamuleau RA, Havekes LM, Groen AK, van Dijk KW: ApoAV reduces plasma triglycerides by inhibiting

- very low density lipoprotein-triglyceride (VLDL-TG) production and stimulating lipoprotein lipase-mediated VLDL-TG hydrolysis. *J Biol Chem*, 2004; 279: 27941-27947
- 31) Aouizerat BE, Kulkarni M, Heilbron D, Drown D, Raskin S, Pullinger CR, Malloy MJ, Kane JP: Genetic analysis of a polymorphism in the human apoA-V gene: effect on plasma lipids. *J Lipid Res*, 2003; 44: 1167-1173
 - 32) Lai CQ, Demissie S, Cupples LA, Zhu Y, Adiconis X, Parnell LD, Corella D, Ordovas JM: Influence of the APOA5 locus on plasma triglyceride, lipoprotein subclasses, and CVD risk in the Framingham Heart Study. *J Lipid Res*, 2004; 45: 2096-2105
 - 33) Ken-Dror G, Goldbourt U, Dankner R: Different effects of apolipoprotein A5 SNPs and haplotypes on triglyceride concentration in three ethnic origins. *J Hum Genet*, 2010; 55: 300-307
 - 34) Martinelli N, Trabetti E, Bassi A, Girelli D, Friso S, Pizzolo F, Sandri M, Malerba G, Pignatti PF, Corrocher R, Olivieri O: The -1131 T>C and S19W APOA5 gene polymorphisms are associated with high levels of triglycerides and apolipoprotein C-III, but not with coronary artery disease: an angiographic study. *Atherosclerosis*, 2007; 191: 409-417
 - 35) Talmud PJ, Hawe E, Martin S, Olivier M, Miller GJ, Rubin EM, Pennacchio LA, Humphries SE: Relative contribution of variation within the APOC3/A4/A5 gene cluster in determining plasma triglycerides. *Hum Mol Genet*, 2002; 11: 3039-3046
 - 36) Wright WT, Young IS, Nicholls DP, Patterson C, Lytle K, Graham CA: SNPs at the APOA5 gene account for the strong association with hypertriglyceridaemia at the APOA5/A4/C3/A1 locus on chromosome 11q23 in the Northern Irish population. *Atherosclerosis*, 2006; 185: 353-360
 - 37) Hishida A, Morita E, Naito M, Okada R, Wakai K, Matsuo K, Nakamura K, Takashima N, Suzuki S, Takezaki T, Mikami H, Ohnaka K, Watanabe Y, Uemura H, Kubo M, Tanaka H, Hamajima N: Associations of apolipoprotein A5 (APOA5), glucokinase (GCK) and glucokinase regulatory protein (GCKR) polymorphisms and lifestyle factors with the risk of dyslipidemia and dysglycemia in Japanese - a cross-sectional data from the J-MICC Study. *Endocr J*, 2012; 59: 589-599
 - 38) Yin RX, Li YY, Lai CQ: Apolipoprotein A1/C3/A5 haplotypes and serum lipid levels. *Lipids Health Dis*, 2011; 140
 - 39) Chien KL, Fang WH, Wen HC, Lin HP, Lin YL, Lin SW, Wu JH, Kao JT: APOA1/C3/A5 haplotype and risk of hypertriglyceridemia in Taiwanese. *Clin Chim Acta*, 2008; 390: 56-62
 - 40) Kao JT, Wen HC, Chien KL, Hsu HC, Lin SW: A novel genetic variant in the apolipoprotein A5 gene is associated with hypertriglyceridemia. *Hum Mol Genet*, 2003; 12: 2533-2539
 - 41) Huang YJ, Lin YL, Chiang CI, Yen CT, Lin SW, Kao JT: Functional importance of apolipoprotein A5 185G in the activation of lipoprotein lipase. *Clin Chim Acta*, 2012; 413: 246-250

Supplementary Fig. 1.
(A) Pedigree of the index family. Square symbols indicate men; circles, women X mark, death. The pedigree number, age, body mass index and lipoprotein levels are indicated below each symbol. The lipid profiles are indicated in the following order: triglycerides, high-density lipoprotein cholesterol, low-density cholesterol and total cholesterol. No data were available for individuals with an asterisk.

Supplementary Table 1.
Candidate genes for SNPs of hypertriglyceridemia based on genomic-wide screening

Chromosome	SNP position	SNP ID	Gene Name	Allele	Description
1	20141528	rs1541185	RNF186	C>T	ring finger protein 186 [Source:HGNC Symbol;Acc:25978]
1	227192752	rs2297417	CDC42BPA	T>C	CDC42 binding protein kinase alpha (DMPK-like) [Source:HGNC Symbol;Acc:1737]
1	236381825	rs1055851	ERO1LB	G>C	ERO1-like beta (<i>S. cerevisiae</i>) [Source:HGNC Symbol;Acc:14355]
1	20141528	rs1541185	RNF186	C>T	ring finger protein 186 [Source:HGNC Symbol;Acc:25978]
1	236381825	rs1055851	ERO1LB	G>C	ERO1-like beta (<i>S. cerevisiae</i>) [Source:HGNC Symbol;Acc:14355]
2	21225281	rs1042034	APOB	C>T	apolipoprotein B (including Ag(x) antigen) [Source:HGNC Symbol;Acc:603]
2	170411698	rs12618227	FASTKDI	C>G	FAST kinase domains 1 [Source:HGNC Symbol;Acc:26150]
2	21225281	rs1042034	APOB	C>T	apolipoprotein B (including Ag(x) antigen) [Source:HGNC Symbol;Acc:603]
2	179545859	rs36051007	TTN	C>T	titin [Source:HGNC Symbol;Acc:12403]
2	183594666	ND	DNAJC10	C>A	DnaJ (Hsp40) homolog, subfamily C, member 10 [Source:HGNC Symbol;Acc:24637]
2	238419714	rs2292880	MLPH	C>T	melanophilin [Source:HGNC Symbol;Acc:29643]
2	240982047	rs148163158	PRR21	A>G	proline rich 21 [Source:HGNC Symbol;Acc:33866]
3	41759288	rs61745795	ULK4	T>C	unc-51-like kinase 4 (<i>C. elegans</i>) [Source:HGNC Symbol;Acc:15784]
3	52551973	rs77803411	STAB1	C>T	stabilin 1 [Source:HGNC Symbol;Acc:18628]
3	155520356	rs9853408	C3orf33	C>T	chromosome 3 open reading frame 33 [Source:HGNC Symbol;Acc:26434]
4	1388759	rs112233131	CRIPAK	C>T	cysteine-rich PAK1 inhibitor [Source:HGNC Symbol;Acc:26619]
5	35033605	rs180749	AGXT2	G>A	alanine-glyoxylate aminotransferase 2 [Source:HGNC Symbol;Acc:14412]
5	52240810	rs988574	ITGA1	A>G	integrin, alpha 1 [Source:HGNC Symbol;Acc:6134]
5	111066697	rs111559	C5orf13	T>C	chromosome 5 open reading frame 13 [Source:HGNC Symbol;Acc:16834]
6	26637796	rs145346169	ZNF322	C>A	zinc finger protein 322 [Source:HGNC Symbol;Acc:23640]
7	16666725	rs36104932	ANKMY2	C>T	ankyrin repeat and MYND domain containing 2 [Source:HGNC Symbol;Acc:25370]
7	29519929	rs3750103	CHN2	A>G	chimerin (chimaerin) 2 [Source:HGNC Symbol;Acc:1944]
7	47968927	rs2686817	PKDIL1	C>A	polycystic kidney disease 1 like 1 [Source:HGNC Symbol;Acc:18053]
7	87160618	rs2032582	ABCB1	A>C	ATP-binding cassette, sub-family B (MDR/TAP), member 1 [Source:HGNC Symbol;Acc:40]
7	95041016	rs12026	PON2	G>C	paraoxonase 2 [Source:HGNC Symbol;Acc:9205]
7	134216684	rs2303312	AKR1B10	C>T	aldo-keto reductase family 1, member B10 (aldose reductase) [Source:HGNC Symbol;Acc:382]
8	29927358	ND	TMEM66	G>A	transmembrane protein 66 [Source:HGNC Symbol;Acc:28789]

(Cont Supplementary Table 1)

Chromosome	SNP position	SNP ID	Gene Name	Allele	Description
8	41906095	rs3824276	KAT6A	A > G	K(lysine) acetyltransferase 6A [Source:HGNC Symbol;Acc:13013]
9	77386669	rs139669465	TRPM6	C > T	transient receptor potential cation channel, subfamily M, member 6 [Source:HGNC Symbol;Acc:17995]
9	79324955	rs11145017	PRUNE2	C > T	prune homolog 2 (Drosophila) [Source:HGNC Symbol;Acc:25209]
9	88937851	rs34431437	ZCCHC6	T > A	zinc finger, CCHC domain containing 6 [Source:HGNC Symbol;Acc:25817]
9	98638288	rs690528	C9orf102	A > G	chromosome 9 open reading frame 102 [Source:HGNC Symbol;Acc:26922]
10	17204187	rs11254413	TRDMT1	G > A	tRNA aspartic acid methyltransferase 1 [Source:HGNC Symbol;Acc:2977]
10	51568378	rs10761581	NCOA4	T > G	nuclear receptor coactivator 4 [Source:HGNC Symbol;Acc:7671]
10	88635779	rs11528010	BMPRIA	C > A	bone morphogenetic protein receptor, type IA [Source:HGNC Symbol;Acc:1076]
10	106152111	rs11192036	CCDC147	T > A	coiled-coil domain containing 147 [Source:HGNC Symbol;Acc:26676]
11	562421	rs2242182	RASSF7	G > A	Ras association (RalGDS/AF-6) domain family (N-terminal) member 7 [Source:HGNC Symbol;Acc:1166]
11	45975130	rs3736508	PHF21A	C > T	PHD finger protein 21A [Source:HGNC Symbol;Acc:24156]
11	55339652	rs1459101	OR4C16	C > T	olfactory receptor, family 4, subfamily C, member 16 [Source:HGNC Symbol;Acc:15172]
11	65487790	ND	RNASEH2C	T > C	ribonuclease H2, subunit C [Source:HGNC Symbol;Acc:24116]
12	131306278	rs7301926	STX2	T > C	syntaxin 2 [Source:HGNC Symbol;Acc:3403]
13	20797501	ND	GJB6	G > A	gap junction protein, beta 6, 30kDa [Source:HGNC Symbol;Acc:4288]
13	28143229	rs8002697	LNK2	A > G	ligand of numb-protein X 2 [Source:HGNC Symbol;Acc:20421]
13	31531009	rs12857479	Cl3orf26	G > A	chromosome 13 open reading frame 26 [Source:HGNC Symbol;Acc:28622]
14	20389606	rs77615110	OR4K5	G > A	olfactory receptor, family 4, subfamily K, member 5 [Source:HGNC Symbol;Acc:14745]
14	103566835	rs2297066	EXOC3L4	C > G	exocyst complex component 3-like 4 [Source:HGNC Symbol;Acc:20120]
15	40661782	ND	DISP2	G > A	dispatched homolog 2 (Drosophila) [Source:HGNC Symbol;Acc:19712]
15	40477831	rs1801376	BUB1B	G > A	budding uninhibited by benzimidazoles 1 homolog beta (yeast) [Source:HGNC Symbol;Acc:1149]
15	43767774	rs560191	TP53BP1	G > C	tumor protein p53 binding protein 1 [Source:HGNC Symbol;Acc:11999]
15	49217103	ND	SHC4	T > C	SHC (Src homology 2 domain containing) family, member 4 [Source:HGNC Symbol;Acc:16743]
16	55853446	rs4513095	CES1	C > A	carboxylesterase 1 [Source:HGNC Symbol;Acc:1863]
16	77246517	rs62049594	SYCE1L	C > A	synaptonemal complex central element protein 1-like [Source:HGNC Symbol;Acc:37236]
16	78458915	rs75559202	WWOX	C > G	WW domain containing oxidoreductase [Source:HGNC Symbol;Acc:12799]
16	81314496	rs7501331	BCMO1	C > T	beta-carotene 15,15'-monooxygenase 1 [Source:HGNC Symbol;Acc:13815]

(Cont Supplementary Table 1)

Chromosome	SNP position	SNP ID	Gene Name	Allele	Description
17	39619115	rs2071563	KRT32	G > A	keratin 32 [Source:HGNC Symbol;Acc:6449]
17	56388238	rs2680704	BZRAP1	C > G	benzodiazepine receptor (peripheral) associated protein 1 [Source:HGNC Symbol;Acc:16831]
17	39274069	rs349771	KRTAP4-11	G > C	keratin associated protein 4-11 [Source:HGNC Symbol;Acc:18911]
17	39305785	rs411367	KRTAP4-5	A > T	keratin associated protein 4-5 [Source:HGNC Symbol;Acc:18899]
17	39619115	rs2071563	KRT32	G > A	keratin 32 [Source:HGNC Symbol;Acc:6449]
17	56388238	rs2680704	BZRAP1	C > G	benzodiazepine receptor (peripheral) associated protein 1 [Source:HGNC Symbol;Acc:16831]
17	67250466	ND	ABCA5	C > T	ATP-binding cassette, sub-family A (ABCI), member 5 [Source:HGNC Symbol;Acc:35]
17	72862593	rs690514	FDXR	C > T	ferredoxin reductase [Source:HGNC Symbol;Acc:3642]
18	2724930	rs633422	SMCHD1	A > T	structural maintenance of chromosomes flexible hinge domain containing 1 [Source:HGNC Symbol;Acc:29090]
19	33467515	rs3816032	Cl9orf40	T > C	chromosome 19 open reading frame 40 [Source:HGNC Symbol;Acc:28467]
19	991129	rs11538683	WDR18	G > A	WD repeat domain 18 [Source:HGNC Symbol;Acc:17956]
19	33467515	rs3816032	Cl9orf40	T > C	chromosome 19 open reading frame 40 [Source:HGNC Symbol;Acc:28467]
19	50464184	ND	SIGLEC11	T > C	sialic acid binding Ig-like lectin 11 [Source:HGNC Symbol;Acc:15622]
19	59074653	rs4756	MZF1	T > C	myeloid zinc finger 1 [Source:HGNC Symbol;Acc:13108]
22	22869742	rs361959	ZNF280A	C > A	zinc finger protein 280A [Source:HGNC Symbol;Acc:18597]
22	35782737	ND	HMOX1	C > G	heme oxygenase (decycling) 1 [Source:HGNC Symbol;Acc:5013]
X	140993912	rs12558365	MAGEC1	C > T	melanoma antigen family C, 1 [Source:HGNC Symbol;Acc:6812]

Supplementary Table 2.

Candidate genes according to the autosomal dominant model for analyzing the variants

chr	SNP pos	gene	region	type	Annotation
1	228467711	OBSCN	exon28	NON_SYNONYMOUS_CODING	obscurin, myosin light chain kinase
1	20141528	RNF186	exon1	NON_SYNONYMOUS_CODING	ring finger protein 186 [Source:HGNC Symbol;Acc:25978]
1	227192752	CDC42BPA	exon34	NON_SYNONYMOUS_CODING	CDC42 binding protein kinase alpha (DMPK-like) [Source:HGNC Symbol;Acc:1737]
1	236381825	ERO1LB	exon16	NON_SYNONYMOUS_CODING	ERO1-like beta (<i>S. cerevisiae</i>) [Source:HGNC Symbol;Acc:14355]
2	21225281	APOB	exon29	NON_SYNONYMOUS_CODING	apolipoprotein B (including Ag(x) antigen) [Source:HGNC Symbol;Acc:603]
2	170411698	FASTKD1	exon7	NON_SYNONYMOUS_CODING	FAST kinase domains 1 [Source:HGNC Symbol;Acc:26150]
7	36764123	AOAH		5PRIME_UTR	acyloxyacyl hydrolase (neutrophil)
7	100472729	TRIP6		DOWNSTREAM	thyroid hormone receptor interactor
8	10583220	SOX7		3PRIME_UTR	transcription factor SOX-7
15	40661782	DISP2	exon8	NON_SYNONYMOUS_CODING	dispatched homolog 2 (<i>Drosophila</i>) [Source:HGNC Symbol;Acc:19712]
17	39619115	KRT32	exon6	NON_SYNONYMOUS_CODING	keratin 32 [Source:HGNC Symbol;Acc:6449]
17	56388238	BZRAP1	exon19	NON_SYNONYMOUS_CODING	benzodiazepine receptor (peripheral) associated protein 1 [Source:HGNC Symbol;Acc:16831]
19	33467515	C19orf40	exon5	NON_SYNONYMOUS_CODING	chromosome 19 open reading frame 40 [Source:HGNC Symbol;Acc:28467]

Supplementary Fig. 2.
The predicted network was generated from candidate genes using IPA (Ingenuity® Systems, www.ingenuity.com).

Supplementary Table 3.
Candidate genes according to the autosomal recessive model for analyzing the variants

chr	SNP position	Gene Name	Region	Type	Description
11	116661392	APOA5	exon 4	missense_variant	apolipoprotein A-V [Source:HGNC Symbol;Acc: 17288]
19	36339044	NPHS1	exon 11	missense_variant	nephrosis 1, congenital, Finnish type (nephrin) [Source:HGNC Symbol;Acc: 7908]