

OPEN

Exome Sequencing in a Family Identifies RECQL5 Mutation Resulting in Early Myocardial Infarction

Xiang Xie, PhD, Ying-Ying Zheng, PhD, Dilare Adi, MM, Yi-Ning Yang, MD, PhD, Yi-Tong Ma, MD, PhD, FACC, Xiao-Mei Li, PhD, Zhen-Yan Fu, PhD, Xiang Ma, PhD, Fen Liu, MM, Zi-Xiang Yu, PhD, You Chen, PhD, and Ying Huang, PhD

Abstract: Coronary artery disease (CAD) including myocardial infarction (MI) is the leading cause of death worldwide and is commonly caused by the interaction between genetic factors and environmental risks. Despite intensive efforts using linkage and candidate gene approaches, the genetic etiology for the majority of families with a multigenerational early CAD/MI predisposition is unknown.

In this study, we used whole-exome sequencing of 10 individuals from 1 early MI family, in which 4 siblings were diagnosed with MI before the age of 55, to identify potential predisposing genes.

We identified a mutation in the *RECQL5* gene, 1 of the 5 members of the RECQ family which are involved in the maintenance of genomic stability. This novel mutation, which is a TG insert at position 73,626,918 on the 13 chromosome and occurs before the last nucleotide of the introns 11 acceptor splice site affecting splicing of *RECQL5*. RT-PCR suggested the control subject had a full-length mRNA including exon 12, but the patients with *RECQL5* mutation had a shorter mRNA form involving splicing of exons 11 to 13 directly, with skipping of exon 12. Quantitative RT-PCR analysis of *RECQL5* exon 12 demonstrated that individuals whose genotype is mutant homozygote had only trace amounts of mRNA containing this exon and the family members who carry the heterozygous genotype had a level at 48% to 55% of the control's level.

These findings provide insight into both the pathogenesis of MI and the role of *RECQL5* gene in human disease.

(*Medicine* 95(5):e2737)

Abbreviations: AMI = acute myocardial infarction, BMI = body mass index, BWA = Burrows–Wheeler alignment, CAD = coronary artery disease, GATK = Genome Analysis Toolkit, GWAS = genome-wide association studies, MI = myocardial infarction.

Editor: Patrick Wall.

Received: May 21, 2015; revised: January 12, 2016; accepted: January 15, 2016.

From the Department of Coronary Artery Disease, Heart Center, First Affiliated Hospital of Xinjiang Medical University, Urumqi, PR China.

Correspondence: Yi-Tong Ma, No. 137, Liyushan Road, Urumqi, China (e-mail: myt_xj@sina.com).

Yi-Ning Yang, No. 137, Liyushan Road, Urumqi, China (e-mail: yangyn5126@163.com).

XX and Y-YZ equally contributed to this study.

Funding: this work has been supported financially by grants from the National Natural Science Foundation of China (81470014).

Disclosure: None.

The authors have no conflicts of interest to disclose.

Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.

This is an open access article distributed under the Creative Commons Attribution-NonCommercial License, where it is permissible to download, share and reproduce the work in any medium, provided it is properly cited.

The work cannot be used commercially.

ISSN: 0025-7974

DOI: 10.1097/MD.0000000000002737

INTRODUCTION

Coronary artery disease (CAD), including myocardial infarction (MI), has been considered as a leading cause of mortality and disability worldwide.^{1,2} Currently, MI has been considered a multifactorial disease resulting from the interaction between several susceptibility genes, such as *C5L2*,³ *SAAI*,⁴ and *CYP17A1* gene,⁵ and multiple environmental factors.^{6,7} Although recent genome-wide association studies (GWAS) have identified several loci associated with CAD and MI,⁸ a large fraction of the genetic risk for CAD/MI have not been identified. Therefore, it is necessary to identify additional susceptibility loci focusing on family constellation. Recently, 4 new loci which were strongly associated with CAD in Chinese Han ethnicity were identified using the GWAS analysis.⁸ Although these findings have provided new insights into pathways contributing to the susceptibility for CAD and MI, we did not find any mutation of these loci in a family whose 4 members were attacked by acute MI before 55 years. Therefore, there might be some new gene mutations related to MI that have not been identified in the past decades. In light of this observation, we examined cases without previously identified pathogenic mutations in an attempt to identify new mutations that can cause early MI.

In this study, we used whole-exome sequencing methods to identify potential predisposing genes in an early MI family and found a novel mutation in the *RECQL5* gene is a pathogenicity-related locus resulting in early MI.

METHODS

Ethics Statement

This study was approved by the ethics committee of the first affiliated hospital of Xinjiang Medical University. Written informed consent was obtained from all subjects or their respective guardians.

Human Subjects

We focused on a family in which 4 siblings were diagnosed with MI (Figure 1A). The affected siblings (II-2, II-3, III-1, and III-3) are alive and have received the coronary stent implantation treatment. All 4 siblings had multivessel disease and >4 stents were implanted for recurrent attacks of acute myocardial infarction (AMI). Also, sibling II-2 and sibling II-3 were attacked by AMI twice and thrice, respectively. This family is of Chinese Han ancestry, ascertained via subject II-3, who presented with AMI at 49 years. He had smoked and taken alcohol before his first hospitalization because of AMI and his body mass index (BMI) was 25.4. He also had some other risk factors such as diabetes and hypercholesterolemia. His father died at 45 years of age for an unknown cause, and his mother suffered CAD and was alive. Detailed clinical data were

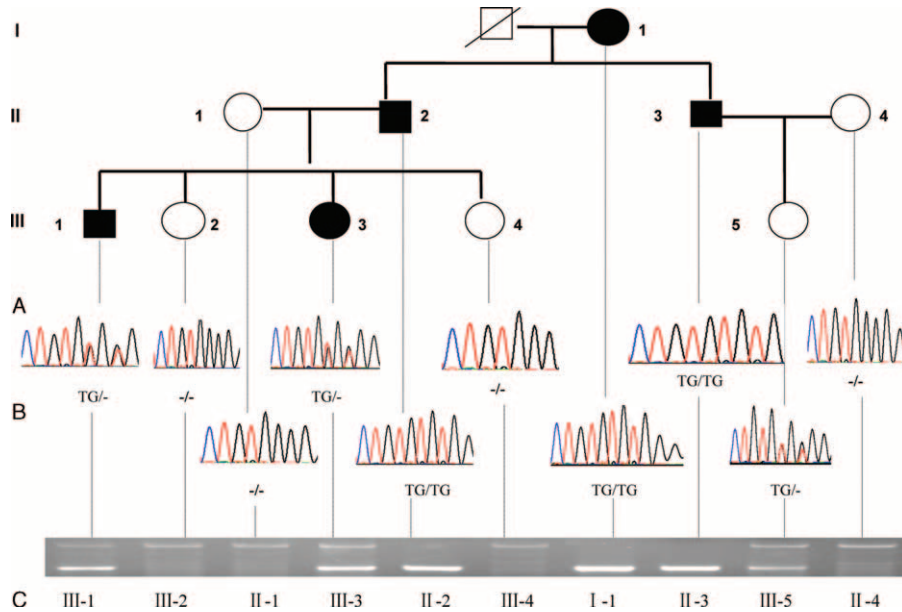


FIGURE 1. Identification of RECQL5 mutations in MI. (A) Pedigree of a family with CAD/MI in 4 children and their mother/grandmother. (B) Sanger sequencing chromatograms from the 3'-end of exon 13 and 5'-end of intron 12 on RECQL5 in the family members. (C) RT-PCR analysis of RECQL5 exons 9 and 14 from peripheral blood RNA samples from the family members. CAD = coronary artery disease, MI = myocardial infarction, RT-PCR = reverse transcription polymerase chain reaction.

obtained for all available kindred members, including 5 affected with MI or CAD, 4 free of CAD, and 1 younger asymptomatic member (CAD phenotype unknown; age 27 years). The clinical characteristics of all these 10 members are shown in Table 1.

Construction of DNA Libraries and Sequencing

Exome sequencing was performed at Shanghai Majorbio Bio-pharm Technology Co., Ltd (Shanghai, China). Briefly, DNA was extracted from peripheral venous blood samples using whole blood purification kit (Qiagen Bio, TLD, Beijing). DNA sequencing library was constructed using 5 μg of genomic DNA. Whole-exome sequence capture was performed according to “SureSelectXT Target Enrichment System for Illumina Paired-End Sequencing Library Illumina HiSeq and MiSeq Multiplexed Sequencing Platforms Protocol Version 1.3.1, February 2012.” We used Agilent 2100 Bioanalyzer for quality

check and quantified the DNA libraries of subjects. Sequencing was performed on the Illumina Genome Analyzer II using a standard 56 cycle paired-end read sequencing protocol and Illumina’s sequencing reagents according to the manufacturer’s recommendations. Each library was sequenced individually on a single flow cell lane.

Analysis of Sequencing Data

Illumina’s Data Analysis Pipeline software (v.1.6) was used to perform the base calling and sequence reads quality assessment. And we used the Burrows–Wheeler alignment (BWA) tool⁹ to perform the alignment of sequence reads to the reference human genome (hg19, GRCh37). All sequencing data were processed according to the previous report.¹⁰ The variant filtration was performed after all variants were obtained, annotated, and assessed from the exome sequencing process.

TABLE 1. Characteristics of the Family Members

Sample	Age (years)	Gender	CHOL (mmol/L)	GLU (mmol/L)	HDLC (mmol/L)	LDL-C (mmol/L)	TG (mmol/L)	URIC (mmol/L)	Smoking	Drinking	CRP (mg/L)	BMI (kg/m ²)
I-1	86	Female	7.25	8.52	0.94	4.55	2.24	420.95	No	No	22.1	23.3
II-1	65	Female	4.41	3.61	1.02	2.04	1.58	221.47	No	No	10.3	20.3
II-2	67	Male	8.18	9.82	0.64	3.14	5.58	515.43	Yes	No	25.3	25.2
II-3	65	Male	7.72	8.1	0.78	2.39	3.83	282.22	Yes	Yes	28.2	25.4
II-4	64	Female	5.24	3.98	0.95	3.43	1.44	292.32	No	No	8.8	20.3
III-1	46	Male	6.2	7.53	0.63	2.08	7.75	520.61	Yes	No	19.3	22.4
III-2	44	Female	4.5	5.73	0.90	1.98	3.99	362.63	No	No	9.3	24.3
III-3	41	Female	7.18	9.17	0.58	2.47	4.69	454.78	No	No	13.4	25.2
III-4	31	Female	4.33	5.16	0.99	2.00	1.63	232.11	No	No	10.3	20.1
III-5	27	Female	5.18	7.17	0.77	3.49	5.08	345.11	No	No	13.3	20.9

BMI = body mass index, CHOL = cholesterol, CRP = C-reactive protein, GLU = glucose, HDLC = High-density lipoprotein, LDL = low-density lipoprotein, URIC = Uric acid.

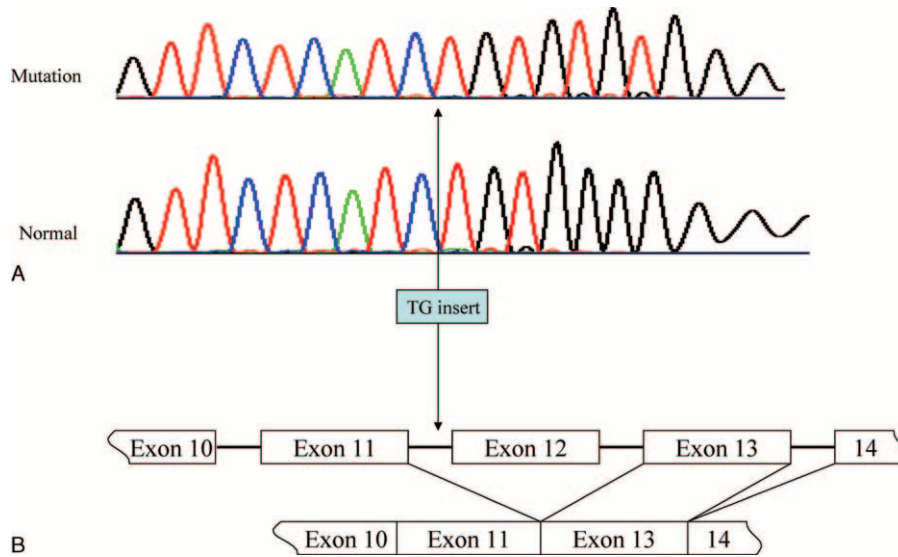


FIGURE 2. A model of how RECQL5 mutation in MI favors production of RECQL5 and the DNA sequence segment of the normal and the mutant one. (A) The sequence of the normal and the mutation. (B) Splicing pattern of RECQL5 in humans and the resulting mRNA. DNA = deoxyribonucleic acid, MI = myocardial infarction.

Raw SNP calls were filtered using the following criteria: “QUAL < 30.0 || QD < 5.0 || HRun > 5 || SB > -0.10” and raw InDel calls were filtered using the following criteria: “MQ0 ≥ 4 && ([MQ0/ {1.0 * DP}] > 0.1)” “SB ≥ -1.0” “QUAL < 10”.

follows: *RECQL5* exon 12 forward: 5’CAGGCTGACTGTG AAGG3’ 5’TTGGAGGCTCTGTGGAT3’. The primers for *GAPDH* (forward: 5’-TGCACCACCAACTGCTTAGC-3’; reverse: 5’-GGCATGGACTGTGGTCATGAG-3’) were designed referring to the previous report.⁹

Sanger Sequencing-Based Genotyping of Variants

The primer pairs were designed using the Primer Premier 5.0 software. The sense primer was 5’TCTCTGGTGTGGAT GGGAC3’ and the antisense primer was 5’GTGTTG ACTGGCGTTGCT3’. We amplify a fragment of 453 bp around this variant. The same set of primers was used for both PCR amplification from genomic DNA samples and for Sanger sequencing. Curated results were obtained from the sequencing-based genotyping using automated and manual approaches.

RT-PCR Analysis

We isolated the total RNA from peripheral blood using a Blood RNA kit (Aidlab, Beijing, China). The RNA was treated with DNase (Aidlab, Beijing, China) after purification. We carried out the RT-PCR using a One-Step RT-PCR kit (Aidlab, Beijing, China). RT-PCR primers were designed in exon 9 (forward, 5’AGGGGAACGGCTTTGAC3’) and in exon 14 (reverse, 5’CCTTGGAGGCTCTGTGGA3’) of *RECQL5*. 30 cycles with an annealing temperature of 55°C were carried out for RT-PCR, and agarose gel electrophoresis was utilized to analyze the products.

Quantitative RT-PCR

We utilized quantitative RT-PCR to detect the expression of RECQL5 exon 12. RNA was prepared as above. The PCR amplification was performed on the ABI Prism 7900HT sequence-detection system (Applied Biosystems). *GAPDH* was used as the control reference. RECQL5 expression quantification was carried out using the ΔΔCt method as described previously.¹¹ In the procedure, we utilized the primers as

RESULTS

In this study, we carried out a whole-exome sequencing on 10 members of a family in which 4 siblings were diagnosed with MI before 55 years of age. We obtained at least 40-fold coverage for >90% of the target bases. We filtered out all variants utilizing the following information: ##fileformat = VCFv4.0; ##fileDate = 20120118; ##source = dbSNP; ##dbSNP_BUILD_ID = 135; ##reference = GRCh37.p5; ## phasing = partial; ##variationPropertyDocumentationUrl = ftp://ftp.ncbi.nlm.nih.gov/ snp/specs/dbSNP_BitField_latest.pdf. After filtering, we identified a mutation in the *RECQL5* gene, which was shared by 5 affected members and 1 unknown phenotype sibling but not by the 4 unaffected members. This extreme familial clustering and segregation of phenotypes within the kindred is unlikely to be explained by chance or multifactorial determination and provides strong evidence that early MI is transmitted as a highly penetrating autosomal dominant trait (Figure 1 A and B), which was in line with the previous study.¹⁰

In the present study, we identified a mutation in the *RECQL5* gene, which is a TG insert at position 73,626,918 on the 13 chromosome (Figure 1B). This mutation occurs before the last nucleotide of the introns 11 acceptor splice site of *RECQL5* gene. Therefore, this mutation was considered to be predicted to affect splicing of *RECQL5*. The results from RT-PCR demonstrated this prediction. The control subject had a full-length mRNA including exon 12, but the patients with *RECQL5* mutation had a shorter mRNA form involving splicing of exons 11 to 13 directly, with skipping of exon 12 (Figure 1C and Figure 2). Furthermore, the quantitative RT-PCR analysis indicated that only trace amounts of mRNA containing exon 12 (5%, 7%, and 9% of control levels, respectively) were observed

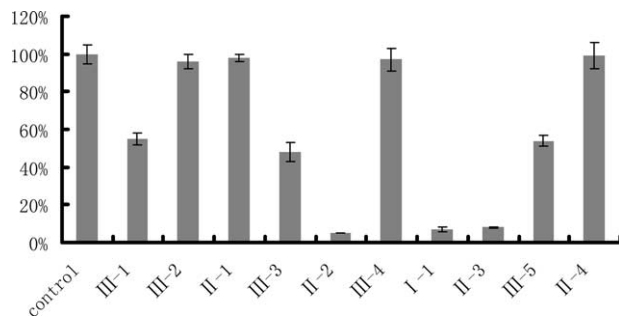


FIGURE 3. RECQL5 exon 12 mutation results in trace amounts of RECQL5 mRNA containing exon 12. Quantitative RT-PCR was performed on peripheral blood RNA samples from all the members in the family. Data are shown as the mean \pm SD. RNA = ribonucleic acid, RT-PCR = reverse transcription polymerase chain reaction, SD = standard deviation.

in individuals I-1, II-2, and II-3 whose genotype is mutant homozygote and the family members who carry the heterozygous genotype had a level at 48% to 55% of the control's level (Figure 3).

DISCUSSION

DNA helicases are a class of enzymes which can obtain energy using the hydrolysis of ATP to open the double-stranded DNA structure and generate single-stranded DNA. These enzymes play an important role in gene transcription, DNA replication, homologous recombination, and DNA repair. Among these, RECQ helicase family members are very conservative in sequence, structure, and function, which are important in maintaining genomic stability.¹² Currently, although several studies suggested that the *RECQL5* gene was associated with cancers,^{13–15} the relation between *RECQL5* gene and CHD/MI remains unclear.

The human *RECQL5* gene is of 3 transcripts (*RECQL5* α , β , and γ) through different splice modes.¹⁶ The *RECQL5* α , β , and γ contain a conserved helicase domain and a RECQC domain, respectively.¹⁷ Furthermore, the *RECQL5* β also contains a long C-terminal, which contains an NLS. However, all other RECQ helicase do not have similar C-terminal sequence, which indicates that *RECQL5* β may be of particular biological function.

Izumikawa et al¹⁸ found that the *RECQL5* gene was associated with low-density lipoprotein receptor (LDLR) gene and β -actin gene, both of which were reported to be associated with CAD. And their results showed that when *RECQL5* gene was knocked down, the LDLR gene and β -actin gene were significantly upregulated. The primary function of the membrane protein LDLR is to remove the circulating LDL by binding to and internalizing it, which are essential for the regulation of plasma cholesterol. Previous study indicated that LDLR mutations influence the LDLR residual activities to a different degree¹⁹ and associated with coronary heart disease.²⁰ Therefore, the novel mutation, a TG insert at position 73,626,918 on the 13 chromosome and occurs before the last nucleotide of the introns 11 acceptor splice site, affects splicing of *RECQL5*, which results in the dysfunction of *RECQL5* gene due to deficiency of exon 12. Furthermore, the loss-of-function *RECQL5* may affect the expression of LDLR gene according to the report from Izumikawa et al.¹⁸ This explanation may answer the question why the *RECQL5* gene mutation can increased the

MI susceptibility in this family. However, it is emphasized that CAD, including MI, is a multifactorial disease resulting from the interaction between genetic background and environmental factors. In the present study, we also found all the subjects with MI have increased serum concentration of TC, LDL-C, and CRP comparing to the healthy individuals. Therefore, there may be other mechanisms to explain the pathogenesis of early MI in this family, which remain unclear but were worthy to exploring in the future.

LIMITATIONS

There are several limitations in the present study. First, we included relative small number of patients in this family. Second, we included a younger asymptomatic member (CAD phenotype unknown; age 27 years) who did not perform in this family. Therefore, this fact to identify healthy subject using a self-stated record rather than coronary angiography is a limitation. Finally, we did not verify the causal relationship between *RECQL5* and MI by knocking out the *RECQL5* gene in mice. Therefore, our results need further verification in both animal model and large-sample population.

REFERENCES

- Roger VL, Go AS, Lloyd-Jones DM, et al. Heart disease and stroke statistics-2011 update: a report from the American Heart Association. *Circulation*. 2011;123:e18–e209.
- He J, Gu D, Wu X, et al. Major causes of death among men and women in China. *N Engl J Med*. 2005;353:1124–1134.
- Zheng YY, Xie X, Ma YT, et al. A novel polymorphism (901G > a) of C5L2 gene is associated with coronary artery disease in Chinese Han and Uyghur population. *Lipids Health Dis*. 2013;12:139.
- Xie X, Ma YT, Yang YN, et al. Genetic polymorphisms of serum amyloid A1 and coronary artery disease risk. *Tissue Antigens*. 2015;85:168–176.
- Dai CF, Xie X, Yang YN, et al. Relationship between CYP17A1 genetic polymorphism and coronary artery disease in a Chinese Han population. *Lipids Health Dis*. 2015;14:16.
- Zee RY, Cheng S, Hegener HH, et al. Genetic variants of arachidonate 5-lipoxygenase-activating protein, and risk of incident myocardial infarction and ischemic stroke: a nested case-control approach. *Stroke*. 2006;37:2007–2011.
- Marenberg ME, Risch N, Berkman LF, et al. Genetic susceptibility to death from coronary heart disease in a study of twins. *N Engl J Med*. 1994;330:1041–1046.
- Lu X, Wang L, Chen S, et al. Genome-wide association study in Han Chinese identifies four new susceptibility loci for coronary artery disease. *Nat Genet*. 2012;44:890–894.
- Sankaran VG, Ghazvinian R, Do R, et al. Exome sequencing identifies GATA1 mutations resulting in Diamond–Blackfan anemia. *J Clin Invest*. 2012;122:2439–2443.
- Mani A, Radhakrishnan J, Wang H, et al. LRP6 mutation in a family with early coronary disease and metabolic risk factors. *Science*. 2007;315:1278–1282.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009;25:1754–1760.
- Larsen NB, Hickson ID. RECQ helicases: conserved guardians of genomic integrity. *Adv Exp Med Biol*. 2013;767:161–184.
- Dong YZ, Huang YX, Lu T. Single nucleotide polymorphism in the *RECQL5* gene increased osteosarcoma susceptibility in a Chinese Han population. *Genet Mol Res*. 2015;14:1899–1902.

14. He YJ, Qiao ZY, Gao B, et al. Association between RECQL5 genetic polymorphisms and susceptibility to breast cancer. *Tumour Biol.* 2014;35:12201–12204.
15. Zhi LQ, Ma W, Zhang H, et al. Association of RECQL5 gene polymorphisms and osteosarcoma in a Chinese Han population. *Tumour Biol.* 2014;35:3255–3259.
16. Seki M, Tada S, Enomoto T. Function of RECQ family helicase in genome stability. *Subcell Biochem.* 2006;40:49–73.
17. Sultan SJ, Sultan ST. Bloom syndrome in two siblings. *Pediatr Dermatol.* 2010;27:174–177.
18. Izumikawa K, Yanagida M, Hayano T, et al. Association of human DNA helicase RECQ5beta with RNA polymerase II and its possible role in transcription. *Biochem J.* 2008;413: 505–516.
19. Etxebarria A, Palacios L, Stef M, et al. Functional characterization of splicing and ligand-binding domain variants in the LDL receptor. *Hum Mutat.* 2012;33:232–243.
20. Ye H, Zhao Q, Huang Y, et al. Meta-analysis of low density lipoprotein receptor (LDLR) rs2228671 polymorphism and coronary heart disease. *Biomed Res Int.* 2014;2014:564940.