Published in final edited form as:

Ann Hum Genet. 2018 November ; 82(6): 437–447. doi:10.1111/ahg.12273.

Analysis of Causal Effect of *APOA5* Variants on Premature Coronary Artery Disease

FAN WANG^{#1,2}, ISABEL Z. WANG^{#1,3}, STEPHEN ELLIS⁴, STEPHEN ARCHACKI^{1,2}, JOHN BARNARD⁵, CARLOS HUBBARD⁴, ERIC J. TOPOL⁶, QIUYUN CHEN^{1,2,*}, and QING. K WANG^{1,2,7,*}

¹Department of Molecular Cardiology, Lerner Research Institute, Department of Cardiovascular Medicine, Cleveland Clinic, Cleveland, Ohio, 44195, USA;

²Department of Molecular Medicine, Department of Genetics and Genome Science, Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Cleveland, OH, 44195, USA;

³Shaker Heights High School, Shaker Heights, OH 44120, USA; Present address: Stanford University, Stanford, CA 94305, USA;

⁴Department of Cardiovascular Medicine, Sydell & Arnold Miller Family Heart & Vascular Institute, Cleveland Clinic, Cleveland, OH 44195, USA;

⁵Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, OH 44195, USA;

⁶Scripps Translational Science Institute, Scripps Research Institute, Scripps Clinic, La Jolla, CA 92037, USA;

⁷The Key Laboratory of Molecular Biophysics of the Ministry of Education, College of Life Science and Technology, Center for Human Genome Research, Cardio-X Institute, Huazhong University of Science and Technology, Wuhan, 430074, Hubei Province, P. R. China.

These authors contributed equally to this work.

Abstract

Apolipoprotein A5 (APOA5) regulates the metabolisms of triglyceride and HDL. *APOA5* variants have been linked to coronary artery disease (CAD), but their causal roles are not well studied yet. This study aims to identify the causal effects of *APOA5* variants on premature CAD. Sequencing analysis of *APOA5* in 128 premature, familial CAD patients from GeneQuest identified 11 genomic variants, including p.S19W (rs3135506). SKAT analysis showed that all sequenced variants, in aggregate, significantly increased the risk of premature CAD ($P_{\text{skat}}=0.037$).

*Corresponding Authors: Department of Molecular Cardiology, Cleveland Clinic Lerner Research Institute, 9500 Euclid Ave, Cleveland, OH 44195, USA. wangq2@ccf.org or qkwang@hust.edu.cn (Q.K. Wang); chenq3@ccf.org (Q. Chen).

Conflict of interest statement

All authors have declared no conflicts of interest.

Author Contribution

Q.C. and Q.K.W. conceived and designed the experiments. F.W., I.Z.W., S.A., Q.C., and Q.K.W. performed all data analyses, reports, and interpretations. S.E., C.H. E.J.T., Q.C., and Q.K.W. participated in sample collection. F.W., I.Z.W., Q.C., and Q.K.W. drafted the manuscript. F.W., I.Z.W., J.B., S.E., E.J.T., Q.C., and Q.K.W. contributed critical comments and revised the manuscript. All authors have reviewed and approved the final version of this manuscript.

Individually, the p.S19W variant was significantly associated with risk of premature CAD (OR=2.30, $p=0.008$) in an independent set of 342 premature CAD patients and 537 controls after adjusting for covariates of sex, age, hypertension, body mass index, triglycerides, total-, LDL-, and HDL-cholesterol levels. Meanwhile, p.S19W significantly correlated with HDL-C levels ($p=0.048$) and TG levels ($p=0.025$). Mediation analysis yielded a mediation effect of p.S19W on risk of premature CAD through HDL-C (OR=0.98, $p=0.040$) and TG (OR=0.98, $p=0.042$), suggesting a causal relationship between p.S19W and premature CAD partially through its effects on HDL-C and TG levels. These results suggest that *APOA5* variation regulates TG and HDL levels, thus displaying a causal role in the development of CAD.

Keywords

Apolipoprotein A5 (APOA5); High Density Lipoprotein-cholesterol (HDL-C); triglyceride; Mediation Analysis; Coronary Artery Disease (CAD)

Introduction

Coronary artery disease (CAD) and its major complication myocardial infarction (MI) are the leading causes of death globally (Benjamin et al., 2017). The heritability of CAD has been estimated to be 40%–60% (Chan and Boerwinkle, 1994), suggesting that genetic factors are critical to the development of CAD and MI. Since 2007, genome-wide association studies (GWAS) by us and others have discovered >65 genetic variants showing significant association with the risk of CAD or MI (Wang et al., 2011b, McPherson and Tybjaerg-Hansen, 2016). However, the most identified risk variants by GWAS exhibit a modest risk and lie in non-coding regions or gene deserts. Therefore, it is not clear how these risk variants affect the development of CAD and MI (Boyle et al., 2017). The identification of causal pathways regulated by CAD risk variants is critical to further our understanding of the mechanisms for the pathogenesis of CAD. To date, only a few GWAS risk variants have been validated to be causal. One interesting example is the identification of a causal role of a common non-coding variant, rs12740374 on chromosome 1p13, which was mainly associated with dyslipidemia by affecting hepatic SORT1 expression via the creation of a C/EBP transcription factor binding site. The rs12740374 was found also to be associated with the risk of CAD/MI due to dyslipidemia (Musunuru et al., 2010).

The *APOA5* gene encodes Apolipoprotein A5, which is a major determinant of plasma triglyceride (TG) levels (Nilsson et al., 2008). It is also an important component of lipoproteins such as high density lipoprotein (HDL) and very lower density lipoprotein (VLDL). *APOA5* directly controls TG levels by increasing the regulating lipoprotein lipase (LPL), thus stimulating the breakdown of TG-rich lipoproteins. It indirectly controls TG levels by reducing VLDL biosynthesis (Garelnabi et al., 2013). The overexpression of human *APOA5* gene in mice resulted in decreased plasma TG levels by 65%, while *APOA5* knockout mice caused a 4-fold increase in TG levels (Pennacchio et al., 2001). Extensive genetic studies have demonstrated that *APOA5* variants are strongly correlated with plasma TG (Lanktree et al., 2009, Pi et al., 2012) and HDL levels (Chien et al., 2009, Shahid et al., 2017) across multiple race populations.

APOA5 variants were found to be significantly associated with CAD, too (Zhou et al., 2013). The most studied *APOA5* variants include: rs662799 (known as c.-1131T>G) in the promoter region of *APOA5*, and rs3135506 (known as p.S19W or c.56C>G) that alters the signal peptide and regulates *APOA5* plasma secretion (Talmud et al., 2005). A large meta-analysis confirmed a significant association between *APOA5* variant rs662799 and CAD/MI with an odds ratio (OR) of 1.18 per minor allele (Triglyceride Coronary Disease Genetics et al., 2010). Recently, large exome sequencing studies have identified rare variants in the LDL receptor gene (*LDLR*) and *APOA5*. Patients carrying rare, non-synonymous (missense) mutations in *APOA5* were at a 2.2-fold increased risk for early-onset MI (Do et al., 2015). *LDLR* mutation carriers had higher levels of plasma LDL cholesterol levels, whereas *APOA5* mutation carriers had higher levels of plasma TG (Do et al., 2015). These results combined suggest that *APOA5* is a CAD-risk gene.

Despite many studies that have highlighted the increased risk of CAD due to *APOA5* variants, some controversial associations exist in a number of studies across multiple race populations. The variants rs662799 and rs3135506 were not associated with the risk of CAD in Italian (Martinelli et al., 2007) and Brazilian populations (Prochaska et al., 2010). A recent meta-analysis identified a significant association of rs662799 with CAD, but it failed to detect significant association for rs3135506 (p.S19W) among 18,254 individuals across 6 studies due to a significant heterogeneity across the six studies (Zhou et al., 2013). Meta-analysis is a simple yet powerful tool to evaluate the controversial results in genetic studies, but it may not effectively control for hidden confounding factors that may introduce spurious signals and cause large heterogeneity among individual studies. An investigation of causal relationships between *APOA5* variants and the risk of CAD would help to understand the above inconsistent results. Recent Mendelian randomization studies supported a causal role of *APOA5* in risk of CAD, and TG was considered as a primary intermediary to bridge that gap between *APOA5* and CAD (Triglyceride Coronary Disease Genetics et al., 2010, Holmes et al., 2015). However, given the several *APOA5* variants associated with HDL- and LDL- C levels, it implies that the causal role of *APOA5* in CAD may be explained by other intermediators rather than TG levels. Moreover, causal roles of many CAD-related biomarkers such as HDL-C and C-reactive protein have not been confirmed yet (Khera and Kathiresan, 2017).

To date, no study investigated the causal pathway mediated by *APOA5* variants in premature CAD through HDL levels. In this study, we carried out a Sanger sequencing analysis of the full *APOA5* gene with a set of 128 premature, familiar CAD patients to identify genomic variants associated CAD. For the sequenced variants in *APOA5*, we explored all putative biological effects of *APOA5*-sequenced variants through the use of bioinformatic analyses with massive public datasets. Then we examined the association between single *APOA5* variant and premature CAD in a separate case-control population of 342 premature CAD patients and 547 non-CAD controls. Finally, we performed a mediation analysis to determine if any causal mediation pathway existed between *APOA5* variants and premature CAD through four lipid traits and BMI. We found that a causal role of an *APOA5* variant in CAD risk development through the HDL levels.

Materials and Methods

Study subjects

Two panels of study subjects were selected from a large U.S. cohort, referred to as GeneQuest, which were recruited at the Center for Cardiovascular Genetics and Department of Cardiovascular Medicine at the Cleveland Clinic. The GeneQuest cohort consisted of 2,036 subjects from 428 unrelated, multiplex families with early-onset (premature) CAD and MI. The premature CAD utilized in the study is defined as any previous or current evidence of significant atherosclerotic CAD occurring in males of ≥ 45 years of age and females of ≥ 50 years of age. Each nuclear family had at least two affected siblings, and the majority of the families also had one normal sibling. The details of this diagnostic criteria and ascertainment of GeneQuest participants have been described by us previously (Guo et al., 2017, Gao et al., 2014, Shen et al., 2014). In brief, the presence or absence of CAD was evaluated according to a coronary angiography with a $>70\%$ stenosis, history of revascularization procedures, such as percutaneous coronary angioplasty (PCA) or coronary artery bypass (CABG), and previous diagnoses of MI as described (Guo et al., 2017, Gao et al., 2014, Shen et al., 2014). Patients with hypercholesterolemia, insulin-dependent diabetes, childhood hypertension, congenital heart disease, cardiomyopathy, valvular disease, and renal or hepatic disease were excluded from this study.

The first genetic study consisted of 128 probands, which were used for sequencing analyses of gene *APOA5*. All 128 patients were Caucasian/white, based on self-reported race. The second genetic study contained an independent set of 342 premature CAD patients and 537 non-CAD controls, which was used to examine the genetic association between *APOA5*-sequenced variants with premature CAD. Each CAD patient was selected from an individual GeneQuest family, and the 537 controls were each selected from $>10,000$ subjects who underwent coronary angiography in Cardiac Catheterization Laboratories at Cleveland Clinic (Cleveland GeneBank). Only Caucasian individuals without angiography-detectable, atherosclerotic lesions were included to the CAD patient group. This study was reviewed and approved by the Cleveland Clinic Institutional Review Board on Human Subject Research, and it conformed to the guidelines set forth by the Declaration of Helsinki. Written informed consent was obtained from all study subjects.

Sequencing analysis and SNP genotyping

Direct Sanger sequencing was performed for a panel of 128 premature CAD patients. For each CAD patient, all 4 exons and exon-intron boundaries of gene *APOA5* (Refseq: NM_052968) were amplified using primers APOA5bF1 and APOA5gR1 obtained from a previous sequencing study conducted by Marcais et al (Marcais et al., 2005). PCR was carried out in a 50 μ l reactions with commercial PCR buffer (Denville Scientific, Holliston, USA) containing 1.5 mM $MgCl_2$, 0.2 mM deoxynucleotide triphosphates, 0.2 μ M of each PCR primer, 0.4 unit of *Taq* polymerase, 4 μ l DMSO, and 10 ng of genomic DNA. The PCR cycles were: 95°C for 3 min followed by 35 cycles of 95°C for 40 seconds, 68°C for 1 min, and 72°C for 3 min. After a 72°C extension for 10 min, the reaction was held at 4°C. PCR products of 2.5 kb were inspected by agarose gels and further purified using the QIAquick PCR Purification Kit (QIAGEN, Germany). The purified PCR products were sequenced

using ABI PRISM 3100 Genetic Analyzer by Cleveland Clinic Genomics Core services, following the manufacture protocol. Following sequencing primers were designed for Sanger sequencing analysis: Seq1, GAC TGA GGT AGC AAG AGC AC; Seq2, GAGGCAGGTCATCATGGCATG; Seq3, TGG GGA CAA AGG AGA TGA T; Seq4, GGG GGA AGA CAC CAA; Seq5, TCC ACC CAT ACG CCG AGA GC. All sequenced variants in gene *APOA5* were confirmed by reverse Sanger sequencing to ensure accuracy of sequencing analysis.

SNP genotyping was performed using a Taqman genotyping assay as described in our previous studies (Wang et al., 2011b, Shen et al., 2014, Shen et al., 2007, Wang et al., 2011a). For *APOA5* variant p.S19W, the TaqMan MGB fluorescent probes were purchased from the company (Thermo Fisher Scientific, Waltham, USA). A total volume of 5 µl PCR including 2.5 µl of TaqMan Universal PCR Master Mix, 0.125 µl of Taqman genotyping assay and 25 ng of genomic DNA was performed, and then PCR product was scanned and analyzed by ABI PRISM 7900HT system and SDS software version 2.1.

Statistical analysis

In this study, all statistical analyses were conducted by R version 3.3.2 (<https://www.r-project.org>) unless otherwise specified. Two types of genetic association analysis were employed to study the associations between *APOA5* sequenced variants and premature CAD. The Hardy-Weinberg disequilibrium test was carried out for the non-CAD controls for *APOA5* variant p.S19W (Xu et al., 2014). Statistical power was estimated using the formula from the Power and Sample Size website (<http://powerandsamplesize.com/Calculators/Test-Odds-Ratio/Equality>), which was developed to compute the sample size and power for case-control association studies. The SNP-set sequence kernel association test (SKAT) (Ionita-Laza et al., 2013) was used to examine whether all sequenced variants, in aggregate, of gene *APOA5* contribute to the risk of premature of CAD. Afterward, the overall p-value was computed with 2,000 times of permutations. A logistic regression model was used to examine the association between *APOA5* variant p.S19W and risk of premature CAD with adjusting for covariates such as sex, age, Hypertension, Tch, HDL-C, LDL-C, TG and BMI. The above analyses were repeated under three genetic models: allelic, C vs G; additive, CC vs CG vs GG; recessive, CG+CC vs GG. The dominant model was omitted due to a low frequency of genotype CC. In each genetic model, OR, 95% confidence intervals (CI), and corresponding p values were calculated accordingly.

The mediation analysis was used to detect a possible causal relationship between *APOA5* variant p.S19W and the risk of premature CAD through analyses of a set of intermediators (Tingley et al., 2014). Prior to analysis, the data on four lipid traits and BMI were log2-transformed and examined for normality by pair-wise correlation analysis and quantile-quantile plotting. Correlations between variant p.S19W and four lipid traits/BMI were analyzed by linear regression modeling with adjustment for sex, age and hypertension. In general, a valid mediation analysis requires that four assumptions are correctly specified (Baron and Kenny, 1986): (i) the genetic variant is associated with CAD; (ii) the genetic variant is associated with a mediator; (iii) the mediator is associated with CAD after adjusting for the genetic variant; (iv) the genetic variant is independent of CAD after

adjusting for the mediator. The mediation analyses were carried out using the R package mediation program (Tingley et al., 2014), which employs 1000 times of nonparametric bootstrapping to estimate the effects of variant p.S19W: the average causal mediation effect through a mediator and the average direct effect not through the mediator. The proportion of the mediation effect over the total effect was also reported to quantify the extent to which the observed mediator effect was relative to the total p.S19W-CAD risk. Covariates of sex, age and hypertension were adjusted in the mediation analysis. Statistical significance was considered as $p < 0.05$.

Bioinformatic analysis and public databases

Putative biological functions of all *APOA5* sequenced variants were explored with a variant annotation tool WGSa (Liu et al., 2016), which has been integrated into massive public databases (<https://sites.google.com/site/jpopgen/wgsa>), such as SnpEff (Cingolani et al., 2012), GRASP (Leslie et al., 2014), SIFT (Kumar et al., 2009), Polyphen2 (Adzhubei et al., 2010), RegulomeDB (Boyle et al., 2012), CADD (Kircher et al., 2014), DANN (Quang et al., 2014), PhyloP46way (Pollard et al., 2010), PhastCons46way (Siepel et al., 2005). Genotypes from the 1000 Genome project Phase 3 (Sudmant et al., 2015) were downloaded through the website <http://www.internationalgenome.org/data>. From the super-population of Europeans in the 1000 Genome project, 503 individuals were selected to match the 128 Caucasian CAD patients. They were then integrated into SKAT analyses. Finally, linkage disequilibrium (LD) patterns for *APOA5* variants were explored with an online LD analysis tool NIH LDlink (<https://analysistools.nci.nih.gov/LDlink>).

Results

Sanger sequencing analysis of the *APOA5* gene to identify CAD-associated genomic variants

The study population for Sanger sequencing analysis consisted of 128 premature CAD patients from 128 families selected from the Cleveland Clinic GeneQuest database, each of which had at least 2 affected individuals. The descriptive statistics of their demographic and clinical features were demonstrated in **Table 1**. The premature CAD patients were all Caucasian with an average age at on-set of 38.9 ± 4.5 years. Of the 128 premature CAD patients, there were 83 (64.8%) male patients, of which nearly half were diagnosed with hypertension (N=60). All 128 premature CAD patients were subjected to Sanger sequencing analyses of *APOA5* (Refseq: NM_052968). A total of 11 variants in gene *APOA5* were identified by Sanger sequencing analyses, including four missense variants (rs3135506, p.S19W; rs34282181, p.D37E; rs138033117, p.E99K; rs3135507, p.V153M), two synonymous variants (rs144178633, p.R24R; rs12287066, p.I44I; rs146781249, p.S232S), one variant in the 5'-UTR region (rs651821, c.-3G>A), one variant in the 3'-UTR region (rs619054, c.*31C>T), and two intronic variants (rs41338746 and rs41416350).

All sequenced variants were further annotated by using the bioinformatic program WGSa (Liu et al., 2016) with public databases. Their putative functions are shown in **Table 2**. Of the eleven sequenced variants, three missense variants (p.S19W, p.D37E, and p.V153M) and two variants in the UTR regions (rs651821 and rs619054) have been linked to CAD and

CAD-related traits (i.e., TG, HDL-C) by genetic association studies reported in the GRASP database (Leslie et al., 2014). Variant rs651821 in the 5'-UTR region displayed the strongest evidence of playing a regulatory role. According to the RegulomeDB database (Boyle et al., 2012), high RegulomeDB scores were also observed from two missense variants (p.S19W, p.D37E) and one synonymous variant (p.S232S). Moreover, one missense variant p.S19W and one synonymous variant (p.E99K) were each predicted to be damaging changes by programs SIFT (Kumar et al., 2009) and Polyphen2 (Adzhubei et al., 2010).

All eleven *APOA5* variants identified in the 128 premature CAD patients were analyzed for their associations with the risk of premature CAD by using the statistical program SKAT (Ionita-Laza et al., 2013). SKAT analyzes the joint-effect of multiple variants in a group on disease outcomes. By utilizing 503 European individuals from the 1000 Genome project as general controls, SKAT analyses showed that the eleven *APOA5* variants were statistically associated with premature CAD ($p_{obs}=0.039$), and the results remained significant ($p_{perm}=0.037$) in a permutation test with 2,000 sampling trials.

Associations of APOA5 variant p.S19W with lipid traits and premature CAD

The missense variant p.S19W (rs3135506) in *APOA5* is predicted to have greater biological effects (Table 2) when compared to the other sequenced variants. Moreover, because variant p.S19W has a relatively high minor allele frequency (MAF) of 4.7%, we can use a case control association study to determine whether this single variant alone confers a risk of premature CAD. We carried out a case-control study that focuses on premature, familial CAD, and contains 342 GeneQuest patients and 537 non-CAD controls. The summary of the study subjects' demographic and clinical features is shown in **Table 1**. All study subjects were Caucasian. The average age at on-set for the premature-CAD group was 39.8 ± 5.1 years, which was significantly younger than the average age for the control group subjects at the ascertainment ($p<0.05$). Significant differences between the CAD and control groups were also found in gender and four lipid traits ($p<0.05$). These results are consistent with the fact that these variables are risk factors for CAD and MI.

APOA5 variant p.S19W was genotyped with Taqman genotyping assays for both the CAD patients and non-CAD controls. The genotyping rate was 100% for variant p.S19W. The Hardy-Weinberg disequilibrium test did not detect any genotypic frequency deviation in 537 of the non-CAD controls ($p=0.391$). The frequency of allele G for variant p.S19W was 96.1% in 342 of the premature-CAD patients. This was significantly higher than that (94.5%) of the 537 non-CAD controls (OR=2.1, 95%CI=1.17–3.76, $p=0.013$) based on a logistic regression model with adjustments made for sex, age, hypertension, Tch, HDL-C, LDL-C, TG, and BMI (**Table 3**). Significant associations were also observed under two genotypic models: the additive model (OR=2.14, 95%CI=1.18–3.523.86, $p=0.012$) and recessive model (OR=2.30, 95%CI=1.24–4.25, $p=0.008$). The recessive model showed a robust association signal for p.S19W (**Table 3**). Assuming a MAF of 6.8% (1000Genome European data) and type I error of 0.05, our power analysis demonstrated that our case-control population provided >80% of statistical power to detect an association for a genetic variant with an OR of 1.63.

Mediation analysis of APOA5 variant p.S19W for risk of CAD

The major focus of this study was to identify possible causal relationships between *APOA5* variants and premature CAD through analysis of a set of intermediators such as lipid levels and BMI. The distributions of original data on lipids and BMI were shown in Figure S1, and three robust correlations were observed on Tch-LDL ($r=0.78$, $p<0.001$), Tch-TG ($r=0.53$, $p<0.001$) and HDL-TG ($r=-0.33$, $p<0.001$). Prior to analysis, the data on five CAD-related lipid and BMI traits were normalized by Log2-transformation, and the distributions of the normalized data were shown in Figure S2. Quantile-quantile plotting indicated that the lipid and BMI data were in normal distribution after Log2-transformation (Figure S3). For the mediation analysis, we first determined whether four lipid traits were associated with the risk of premature CAD. Strong associations were identified for the four lipid traits (Tch, $p=2.3\times 10^{-18}$; HDL-C, $p=6.2\times 10^{-18}$; LDL-C, $p=9.6\times 10^{-16}$; TG, $p=6.7\times 10^{-19}$) using a logistic regression model after adjustment for covariates of sex, age, and hypertension. There was a marginal, but not significant association between BMI and premature CAD ($p=0.056$).

Second, multiple linear regression models were used to examine the correlations between *APOA5* variant p.S19W, four lipid traits, and BMI after adjusting for sex, age and hypertension. Under the additive model, variant p.S19W was significantly associated with TG levels ($\beta=-0.216$, $p=0.027$). Under the recessive model, variant p.S19W was significantly associated with TG levels ($\beta=-0.223$, $p=0.025$) and HDL-C levels ($\beta=0.098$, $p=0.048$). No significant association was detected for Tch, LDL or BMI. Details of these results were shown in **Table 4**.

Third, the mediation analyses were performed to determine whether there is a causal relationship between *APOA5* variant p.S19W and the risk of premature CAD through HDL-C or TG. With 1,000 times of bootstrapping (Tingley et al., 2014), our mediation analysis identified a significant mediation effect of variant p.S19W on reducing CAD risk through HDL-C (additive model: OR=0.98, 95%CI=0.96–1.00, $p=0.048$; recessive model: OR=0.98, 95%CI=0.96–1.00, $p=0.040$) (**Table 5**). The HDL-C mediated effect accounted for 52.7% (additive model) and 51.8% (recessive model) of total p.S19W-CAD risk. Similar results were observed for TG (additive model: OR=0.98, 95%CI=0.96–1.00, $p=0.040$; recessive model: OR=0.98, 95%CI=0.95–1.00, $p=0.042$) (**Table 5**). The TG-mediated effect accounted for 50.6% and 48.8% of total p.S19W-CAD risk under additive and recessive models, respectively. These data suggest that there is a causal relationship between variant p.S19W and risk of CAD, which is partially mediated by HDL-C and TG.

Discussion

The causal role of *APOA5* variants in CAD/MI was not well studied. Our sequence analysis identified eleven variants in *APOA5* (**Table 2**). The variants, in aggregate, contributed to the risk of premature CAD ($p=0.037$ by permutation). Furthermore, we found that the *APOA5* missense variant p.S19W (rs3135506) contributed to the risk of premature CAD in a separate case-control study of premature CAD (**Table 3**). Further analysis showed that p.S19W was correlated with both TG and HDL levels (Table 4), thus, we performed mediation analyses using HDL-C and TG as separate intermediators to detect a possible causal relationship between the *APOA5* variant and risk of CAD through intermediate

mediators. We detected a modest yet significant causal, indirect effect of p.S19W on reducing the risk of premature CAD through TG and HDL-C (**Table 5**). Our analysis showed that 48.8–50.6% of the total p.S19W-CAD risk was attributed to TG levels (**Table 5**). Our data further support the earlier Mendelian randomization finding supporting a causal role of *APOA5* in risk of CAD via TG, (Triglyceride Coronary Disease Genetics et al., 2010, Holmes et al., 2015). Most interestingly, our mediation analysis showed that 51.8–52.7% of the total p.S19W-CAD risk was attributed to HDL-C levels and there was a negative correlation between HDL-C levels and premature CAD (**Table 5**). To the best of our knowledge, this is the first report showing that *APOA5* variation confers risk of premature CAD by regulating HDL levels. It is equally interesting to note that the identified HDL-C mediated effect on CAD by *APOA5* variant p.S19W did not conflict with the current concept of HDL's roles in CAD. Moreover, although we found that *APOA5* variant p.S19W is causal to CAD partially due to its effects on HDL-C and TG, the HDL-C and TG levels in each participant are also affected by other genetic variants. These other genetic variants may explain the partial effects of HDL-C and TG on the causal relationship between *APOA5* variant p.S19W and CAD.

Our data showed that *APOA5* variant p.S19W was significantly associated with HDL-C levels (Table 4). Our finding was supported by other studies. Priore Oliva et al. (2008) reported a 17-year-old patient with hypertriglyceridemia and low plasma HDL-C levels carried an *APOA5* variant p.Q97X. A genetic study of nearly 4,000 samples from American families demonstrated that the interaction between variant p.S19W and 25-hydroxy-vitamin D (25OHD) influenced in HDL-C levels (Shirts et al., 2012). The molecular mechanism by which *APOA5* variant p.S19W affects HDL-C levels is unknown. The p.S19W variant caused a substitution of a tryptophan residue (Trp, W) for a serine residue (Ser, S) at codon 19. The variant is located within the signal peptide. The functional analysis of variant p.S19W showed that W19 enlarged the angle of the signal peptide insertion into the water/lipid interface of the cell membrane by 25°, which reduced the secretion of the *APOA5*-secretory alkaline phosphatase (SEAP) fusion protein (a reporter) into the plasma by 50% (Talmud et al., 2005). *APOA5* was shown to be in a complex with VLDL and thereby serves as a key regulator of plasma TG levels (Garelnabi et al., 2013). *APOA5* was also shown to be in a complex with HDL (Garelnabi et al., 2013), therefore, may also regulate HDL-C levels. Moreover, our bioinformatic analysis also showed that p.S19W is a damaging variant by multiple programs (**Table 2**). Moreover, p.S19W may affect the binding of multiple proteins (i.e., IKZF1, ZNF263, and ETS1), as stated in the RegulomeDB (<http://www.regulomedb.org/snp/chr11/116662406>). The *IKZF1* gene encodes a zinc-finger DNA-binding protein, and has a function in chromatin remodeling. It was identified as a locus for HDL-C GWAS ($p=1.0\times10^{-8}$) (Global Lipids Genetics et al., 2013). The *ZNF263* gene encodes zinc finger protein 263, which plays an important role in basic cellular processes as a transcriptional repressor. It was also recently found to be involved in the fine-tuning of TGFB1 expression in atherosclerotic lesions (Dhaouadi et al., 2014). The transcription factor Ets1, encoded by *ETS1*, was demonstrated to be critical for VE-cadherin expression in endothelial cells in an early study (Lelievre et al., 2000). One recent study showed that the Ets1-mediated, VE-cadherin pathway was activated in the presence of Tyrosine kinase receptor κ , leading to a described risk of CAD (Jiang et al., 2015). These interesting findings

support that p.S19W is a causal variant for CAD. Based on the 1000Genome data, the frequencies of many *APOA5* variants including variant p.S19W vary in different ethnic populations. It may be interesting to explore whether the findings in this study can be expanded to non-European ancestry populations in the future.

Besides the p.S19W variant, our study also identified ten other variants in 128 premature-CAD patients. Most variants were predicted to have regulatory roles by bioinformatic analyses with the exception of two intronic variants (**Table 2**). The rs651821 variant in the 5'-UTR region is not conservative, but it has the largest RegulomeDB score of 1f among eleven sequenced variants, suggesting its important regulatory role. Other studies have shown that rs651821 was significantly correlated with plasma lipid levels (Gombojav et al., 2016) as well as CAD (Paré et al., 2007). In 2005, deep sequencing of over 13,000 MI patients identified 46 variants in *APOA5*, and these variants aggregately increased risk of MI by 2.2 fold (Do et al., 2015). Consistent with this, our SKAT analysis showed that all eleven sequenced variants identified in our GeneQuest population conferred an accumulative risk of CAD ($p=0.037$). A recent meta-analysis of multi-ethnic populations from Southeast Asia identified a missense variant, p.G185C (rs2075291), in *APOA5* that was significantly associated with CAD and acute MI as well as HDL-C levels (Han et al., 2017). Another recent study supports the association between variant p.G185C and MI in North Indian populations (Kashyap et al., 2018). However, we did not detect variant p.G185C in our GeneQuest population, therefore, different *APOA5* variants may be associated with risk of CAD, TG and HDL-C levels in different ethnic populations. It may be interesting to use mediation analyses to explore whether variant p.G185C or other *APOA5* variants are causal to CAD through intermediators such as TG and HDL-C in other non-European ancestry populations in the future.

In summary, the data in this study show that rare variants in *APOA5*, in aggregate, confer a significant risk of premature CAD. Most interestingly, for the first time our study identifies a causal role of *APOA5* variation in the development of CAD by regulating the HDL metabolism. The same strategy may be used to analyze other CAD/MI variants identified to date, in particular, those found by GWAS, to determine whether or not they are causal to CAD and MI. Identification of causal variants and genes is expected to greatly facilitate the elucidation of molecular pathogenic mechanisms of CAD and MI.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the patients and their relatives for their enthusiastic participation in the study. This study was supported by NIH/NHLBI grants R01 HL121358 and R01 HL126729, the China National Natural Science Foundation grants (81630002, 91439129 and 31430047), the Chinese National Basic Research Programs (973 Programs 2013CB531101 and 2012CB517801), Hubei Province Natural Science Key Program (2016CFB224 and 2014CFA074), Hubei Province's Outstanding Medical Academic Leader Program, and Specialized Research Fund for the Doctoral Program of Higher Education from the Ministry of Education. I.Z.W. was supported by the Cleveland Clinic Summer Internship Program.

References

- ADZHUBEI IA, SCHMIDT S, PESHKIN L, RAMENSKY VE, GERASIMOVA A, BORK P, KONDRASHOV AS & SUNYAEV SR 2010 A method and server for predicting damaging missense mutations. *Nature methods*, 7, 248–249. [PubMed: 20354512]
- BARON RM & KENNY DA 1986 The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, 51, 1173. [PubMed: 3806354]
- BENJAMIN EJ, BLAHA MJ, CHIUVE SE, CUSHMAN M, DAS SR, DEO R, DE FERRANTI SD, FLOYD J, FORNAGE M, GILLESPIE C, ISASI CR, JIMENEZ MC, JORDAN LC, JUDD SE, LACKLAND D, LICHTMAN JH, LISABETH L, LIU S, LONGENECKER CT, MACKEY RH, MATSUSHITA K, MOZAFFARIAN D, MUSSOLINO ME, NASIR K, NEUMAR RW, PALANIAPPAN L, PANDEY DK, THIAGARAJAN RR, REEVES MJ, RITCHEY M, RODRIGUEZ CJ, ROTH GA, ROSAMOND WD, SASSON C, TOWFIGHI A, TSAO CW, TURNER MB, VIRANI SS, VOEKS JH, WILLEY JZ, WILKINS JT, WU JH, ALGER HM, WONG SS, MUNTNER P, AMERICAN HEART ASSOCIATION STATISTICS, C. & STROKE STATISTICS, S. 2017 Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association. *Circulation*, 135, e146–e603. [PubMed: 28122885]
- BOYLE AP, HONG EL, HARIHARAN M, CHENG Y, SCHAUB MA, KASOWSKI M, KARCZEWSKI KJ, PARK J, HITZ BC, WENG S, CHERRY JM & SNYDER M 2012 Annotation of functional variation in personal genomes using RegulomeDB. *Genome Research*, 22, 1790–1797. [PubMed: 22955989]
- BOYLE EA, LI YI & PRITCHARD JK 2017 An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell*, 169, 1177–1186. [PubMed: 28622505]
- CHAN L & BOERWINKLE E 1994 Gene-environment interactions and gene therapy in atherosclerosis. *Cardiology in Review*, 2, 130–137.
- CHIEN KL, HSU HC, CHEN YC, SU TC, LEE YT & CHEN MF 2009 Association between sequence variant of c.553 G > T in the apolipoprotein A5 gene and metabolic syndrome, insulin resistance, and carotid atherosclerosis. *Transl Res*, 154, 133–41. [PubMed: 19665689]
- CINGOLANI P, PLATTS A, WANG LL, COON M, NGUYEN T, WANG L, LAND SJ, LU XY & RUDEN DM 2012 A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w(1118); iso-2; iso-3. *Fly*, 6, 80–92. [PubMed: 22728672]
- DHAOUADI N, LI JY, FEUGIER P, GUSTIN MP, DAB H, KACEM K, BRICCA G & CERUTTI C 2014 Computational identification of potential transcriptional regulators of TGF-ss1 in human atherosclerotic arteries. *Genomics*, 103, 357–70. [PubMed: 24819318]
- DO R, STITZEL NO, WON HH, JORGENSEN AB, DUGA S, ANGELICA MERLINI P, KIEZUN A, FARRALL M, GOEL A, ZUK O, GUELLA I, ASSELT R, LANGE LA, PELOSO GM, AUER PL, PROJECT NES, GIRELLI D, MARTINELLI N, FARLOW DN, DEPRISTO MA, ROBERTS R, STEWART AF, SALEHEEN D, DANESH J, EPSTEIN SE, SIVAPALARATNAM S, HOVINGH GK, KASTELEIN JJ, SAMANI NJ, SCHUNKERT H, ERDMANN J, SHAH SH, KRAUS WE, DAVIES R, NIKPAY M, JOHANSEN CT, WANG J, HEGELE RA, HECHTER E, MARZ W, KLEBER ME, HUANG J, JOHNSON AD, LI M, BURKE GL, GROSS M, LIU Y, ASSIMES TL, HEISS G, LANGE EM, FOLSOM AR, TAYLOR HA, OLIVIERI O, HAMSTEN A, CLARKE R, REILLY DF, YIN W, RIVAS MA, DONNELLY P, ROSSOUW JE, PSATY BM, HERRINGTON DM, WILSON JG, RICH SS, BAMSHAD MJ, TRACY RP, CUPPLES LA, RADER DJ, REILLY MP, SPERTUS JA, CRESCI S, HARTIALA J, TANG WH, HAZEN SL, ALLAYEE H, REINER AP, CARLSON CS, KOOPERBERG C, JACKSON RD, BOERWINKLE E, LANDER ES, SCHWARTZ SM, SISCOVICK DS, MCPHERSON R, TYBJAERG-HANSEN A, ABECASIS GR, WATKINS H, NICKERSON DA, ARDISSINO D, SUNYAEV SR, O'DONNELL CJ, ALTSHULER D, GABRIEL S & KATHIRESAN S 2015 Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. *Nature*, 518, 102–6. [PubMed: 25487149]

- GAO H, LI L, RAO S, SHEN G, XI Q, CHEN S, ZHANG Z, WANG K, ELLIS SG, CHEN Q, TOPOL EJ & WANG QK 2014 Genome-wide linkage scan identifies two novel genetic loci for coronary artery disease: in GeneQuest families. *PLoS One*, 9, e113935. [PubMed: 25485937]
- GARELNABI M, LOR K, JIN J, CHAI F & SANTANAM N 2013 The paradox of ApoA5 modulation of triglycerides: evidence from clinical and basic research. *Clin Biochem*, 46, 12–9. [PubMed: 23000317]
- GLOBAL LIPIDS GENETICS, C, WILLER CJ, SCHMIDT EM, SENGUPTA S, PELOSO GM, GUSTAFSSON S, KANONI S, GANNA A, CHEN J, BUCHKOVICH ML, MORA S, BECKMANN JS, BRAGG-GRESHAM JL, CHANG HY, DEMIRKAN A, DEN HERTOOG HM, DO R, DONNELLY LA, EHRET GB, ESKO T, FEITOSA MF, FERREIRA T, FISCHER K, FONTANILLAS P, FRASER RM, FREITAG DF, GURDASANI D, HEIKKILA K, HYPONEN E, ISAACS A, JACKSON AU, JOHANSSON A, JOHNSON T, KAAKINEN M, KETTUNEN J, KLEBER ME, LI X, LUAN J, LYYTIKAINEN LP, MAGNUSSON PK, MANGINO M, MIHAILOV E, MONTASSER ME, MULLER-NURASYID M, NOLTE IM, O'CONNELL JR, PALMER CD, PEROLA M, PETERSEN AK, SANNA S, SAXENA R, SERVICE SK, SHAH S, SHUNGIN D, SIDORE C, SONG C, STRAWBRIDGE RJ, SURAKKA I, TANAKA T, TESLOVICH TM, THORLEIFSSON G, VAN DEN HERIK EG, VOIGHT BF, VOLCIK KA, WAITE LL, WONG A, WU Y, ZHANG W, ABSHER D, ASIKI G, BARROSO I, BEEN LF, BOLTON JL, BONNYCASTLE LL, BRAMBILLA P, BURNETT MS, CESANA G, DIMITRIOU M, DONEY AS, DORING A, ELLIOTT P, EPSTEIN SE, EYJOLFSSON GI, GIGANTE B, GOODARZI MO, GRALLERT H, GRAVITO ML, GROVES CJ, HALLMANS G, HARTIKAINEN AL, HAYWARD C, HERNANDEZ D, HICKS AA, HOLM H, HUNG YJ, ILLIG T, JONES MR, KALEEBU P, KASTELEIN JJ, KHAW KT, et al. 2013 Discovery and refinement of loci associated with lipid levels. *Nat Genet*, 45, 1274–83. [PubMed: 24097068]
- GOMBOJAV B, LEE SJ, KHO M, SONG YM, LEE K & SUNG J 2016 Multiple susceptibility loci at chromosome 11q23.3 are associated with plasma triglyceride in East Asians. *J Lipid Res*, 57, 318–24. [PubMed: 26634697]
- GUO Y, WANG F, LI L, GAO H, ARCKACKI S, WANG IZ, BARNARD J, ELLIS S, HUBBARD C, TOPOL EJ, CHEN Q & WANG QK 2017 Genome-Wide Linkage Analysis of Large Multiple Multigenerational Families Identifies Novel Genetic Loci for Coronary Artery Disease. *Sci Rep*, 7, 5472. [PubMed: 28710368]
- HAN Y, DORAJOO R, CHANG X, WANG L, KHOR CC, SIM X, CHENG CY, SHI Y, THAM YC, ZHAO W, CHEE ML, SABANAYAGAM C, CHEE ML, TAN N, WONG TY, TAI ES, LIU J, GOH DYT, YUAN JM, KOH WP, VAN DAM RM, LOW AF, CHAN MY, FRIEDLANDER Y & HENG CK 2017 Genome-wide association study identifies a missense variant at APOA5 for coronary artery disease in Multi-Ethnic Cohorts from Southeast Asia. *Sci Rep*, 7, 17921. [PubMed: 29263402]
- HOLMES MV, ASSELBERGS FW, PALMER TM, DRENOS F, LANKTREE MB, NELSON CP, DALE CE, PADMANABHAN S, FINAN C, SWERDLOW DI, TRAGANTE V, VAN IPEREN EP, SIVAPALARATNAM S, SHAH S, ELBERS CC, SHAH T, ENGMANN J, GIAMBARTOLOMEI C, WHITE J, ZABANEH D, SOFAT R, MCLACHLAN S, CONSORTIUM U, DOEVENDANS PA, BALMFORTH AJ, HALL AS, NORTH KE, ALMOGUERA B, HOOGEVEEN RC, CUSHMAN M, FORNAGE M, PATEL SR, REDLINE S, SISCOVICK DS, TSAI MY, KARCZEWSKI KJ, HOFKER MH, VERSCHUREN WM, BOTS ML, VAN DER SCHOUW YT, MELANDER O, DOMINICZAK AF, MORRIS R, BEN-SHLOMO Y, PRICE J, KUMARI M, BAUMERT J, PETERS A, THORAND B, KOENIG W, GAUNT TR, HUMPHRIES SE, CLARKE R, WATKINS H, FARRALL M, WILSON JG, RICH SS, DE BAKKER PI, LANGE LA, DAVEY SMITH G, REINER AP, TALMUD PJ, KIVIMAKI M, LAWLOR DA, DUDBRIDGE F, SAMANI NJ, KEATING BJ, HINGORANI AD & CASAS JP 2015 Mendelian randomization of blood lipids for coronary heart disease. *Eur Heart J*, 36, 539–50. [PubMed: 24474739]
- IONITA-LAZA I, LEE S, MAKAROV V, BUXBAUM JD & LIN X 2013 Sequence kernel association tests for the combined effect of rare and common variants. *Am J Hum Genet*, 92, 841–53. [PubMed: 23684009]
- JIANG H, HUANG S, LI X, LI X, ZHANG Y & CHEN ZY 2015 Tyrosine kinase receptor B protects against coronary artery disease and promotes adult vasculature integrity by regulating Ets1-

mediated VE-cadherin expression. *Arterioscler Thromb Vasc Biol*, 35, 580–8. [PubMed: 25633318]

- KASHYAP S, KUMAR S, AGARWAL V, MISRA DP, RAI MK & KAPOOR A 2018 The association of polymorphic variants, rs2267788, rs1333049 and rs2383207 with coronary artery disease, its severity and presentation in North Indian population. *Gene*, 648, 89–96. [PubMed: 29309886]
- KHERA AV & KATHIRESAN S 2017 Genetics of coronary artery disease: discovery, biology and clinical translation. *Nat Rev Genet*, 18, 331–344. [PubMed: 28286336]
- KIRCHER M, WITTEN DM, JAIN P, O'ROAK BJ, COOPER GM & SHENDURE J 2014 A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*, 46, 310–5. [PubMed: 24487276]
- KUMAR P, HENIKOFF S & NG PC 2009 Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols*, 4, 1073–1082. [PubMed: 19561590]
- LANKTREE MB, ANAND SS, YUSUF S, HEGELE RA & INVESTIGATORS S 2009 Replication of genetic associations with plasma lipoprotein traits in a multiethnic sample. *J Lipid Res*, 50, 1487–96. [PubMed: 19299407]
- LELIEVRE E, MATTOT V, HUBER P, VANDENBUNDER B & SONCIN F 2000 ETS1 lowers capillary endothelial cell density at confluence and induces the expression of VE-cadherin. *Oncogene*, 19, 2438–2446. [PubMed: 10828886]
- LESLIE R, O'DONNELL CJ & JOHNSON AD 2014 GRASP: analysis of genotype-phenotype results from 1390 genome-wide association studies and corresponding open access database. *Bioinformatics*, 30, i185–94. [PubMed: 24931982]
- LIU X, WHITE S, PENG B, JOHNSON AD, BRODY JA, LI AH, HUANG Z, CARROLL A, WEI P, GIBBS R, KLEIN RJ & BOERWINKLE E 2016 WGS: an annotation pipeline for human genome sequencing studies. *J Med Genet*, 53, 111–2. [PubMed: 26395054]
- MARCAIS C, VERGES B, CHARRIERE S, PRUNETA V, MERLIN M, BILLON S, PERROT L, DRAI J, SASSOLAS A, PENNACCHIO LA, FRUCHART-NAJIB J, FRUCHART JC, DURLACH V & MOULIN P 2005 ApoA5 Q139X truncation predisposes to late-onset hyperchylomicronemia due to lipoprotein lipase impairment. *J Clin Invest*, 115, 2862–9. [PubMed: 16200213]
- MARTINELLI N, TRABETTI E, BASSI A, GIRELLI D, FRISO S, PIZZOLO F, SANDRI M, MALERBA G, PIGNATTI PF, CORROCHER R & OLIVIERI O 2007 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*, 191, 409–17. [PubMed: 16682041]
- MCPHERSON R & TYBJAERG-HANSEN A 2016 Genetics of Coronary Artery Disease. *Circ Res*, 118, 564–78. [PubMed: 26892958]
- MUSUNURU K, STRONG A, FRANK-KAMENETSKY M, LEE NE, AHFELDT T, SACHS KV, LI X, LI H, KUPERWASSER N, RUDA VM, PIRRUCCELLO JP, MUCHMORE B, PROKUNINA-OLSSON L, HALL JL, SCHADT EE, MORALES CR, LUND-KATZ S, PHILLIPS MC, WONG J, CANTLEY W, RACIE T, EJEBE KG, ORHO-MELANDER M, MELANDER O, KOTELIANSKY V, FITZGERALD K, KRAUSS RM, COWAN CA, KATHIRESAN S & RADER DJ 2010 From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature*, 466, 714–9. [PubMed: 20686566]
- NILSSON SK, CHRISTENSEN S, RAARUP MK, RYAN RO, NIELSEN MS & OLIVECRONA G 2008 Endocytosis of apolipoprotein A-V by members of the low density lipoprotein receptor and the VPS10p domain receptor families. *J Biol Chem*, 283, 25920–7. [PubMed: 18603531]
- PARÉ G, SERRE D, BRISSON D, ANAND SS, MONTPETIT A, TREMBLAY G, ENGERT JC, HUDSON TJ & GAUDET D 2007 Genetic analysis of 103 candidate genes for coronary artery disease and associated phenotypes in a founder population reveals a new association between endothelin-1 and high-density lipoprotein cholesterol. *The American journal of human genetics*, 80, 673–682. [PubMed: 17357073]
- PENNACCHIO LA, OLIVIER M, HUBACEK JA, COHEN JC, COX DR, FRUCHART JC, KRAUSS RM & RUBIN EM 2001 An apolipoprotein influencing triglycerides in humans and mice revealed by comparative sequencing. *Science*, 294, 169–73. [PubMed: 11588264]

- PI Y, ZHANG L, YANG Q, LI B, GUO L, FANG C, GAO C, WANG J, XIANG J & LI J 2012 Apolipoprotein A5 gene promoter region-1131T/C polymorphism is associated with risk of ischemic stroke and elevated triglyceride levels: a meta-analysis. *Cerebrovasc Dis*, 33, 558–65. [PubMed: 22688093]
- POLLARD KS, HUBISZ MJ, ROSENBLOOM KR & SIEPEL A 2010 Detection of nonneutral substitution rates on mammalian phylogenies. *Genome research*, 20, 110–121. [PubMed: 19858363]
- PRICE AL, ZAITLEN NA, REICH D & PATTERSON N 2010 New approaches to population stratification in genome-wide association studies. *Nat Rev Genet*, 11, 459–63. [PubMed: 20548291]
- PRIORE OLIVA C, CARUBBI F, SCHAAP FG, BERTOLINI S & CALANDRA S 2008 Hypertriglyceridaemia and low plasma HDL in a patient with apolipoprotein A-V deficiency due to a novel mutation in the APOA5 gene. *J Intern Med*, 263, 450–8. [PubMed: 18324930]
- PROCHASKA CL, PICHETH G, ANGHEBEM-OLIVEIRA MI, COSTANTINI CO, DE SOUZA EM, PEDROSA FO & SCARTEZINI M 2010 The polymorphisms –1131T>C and the S19W of the APOA5 gene are not associated with coronary artery disease in a Brazilian population. *Clin Chem Lab Med*, 48, 419–22. [PubMed: 20170397]
- QUANG D, CHEN Y & XIE X 2014 DANN: a deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics*, 31, 761–763. [PubMed: 25338716]
- SHAHID SU, SHABANA NA, COOPER JA, REHMAN A & HUMPHRIES SE 2017 Common variants in the genes of triglyceride and HDL-C metabolism lack association with coronary artery disease in the Pakistani subjects. *Lipids Health Dis*, 16, 24. [PubMed: 28143480]
- SHEN GQ, GIRELLI D, LI L, RAO S, ARCHACKI S, OLIVIERI O, MARTINELLI N, PARK JE, CHEN Q, TOPOL EJ & WANG QK 2014 A novel molecular diagnostic marker for familial and early-onset coronary artery disease and myocardial infarction in the LRP8 gene. *Circ Cardiovasc Genet*, 7, 514–20. [PubMed: 24867879]
- SHEN GQ, LI L, GIRELLI D, SEIDELMANN SB, RAO S, FAN C, PARK JE, XI Q, LI J, HU Y, OLIVIERI O, MARCHANT K, BARNARD J, CORROCHER R, ELSTON R, CASSANO J, HENDERSON S, HAZEN SL, PLOW EF, TOPOL EJ & WANG QK 2007 An LRP8 variant is associated with familial and premature coronary artery disease and myocardial infarction. *Am J Hum Genet*, 81, 780–91. [PubMed: 17847002]
- SHIRTS BH, HOWARD MT, HASSTEDT SJ, NANJEE MN, KNIGHT S, CARLQUIST JF, ANDERSON JL, HOPKINS PN & HUNT SC 2012 Vitamin D dependent effects of APOA5 polymorphisms on HDL cholesterol. *Atherosclerosis*, 222, 167–74. [PubMed: 22425169]
- SIEPEL A, BEJERANO G, PEDERSEN JS, HINRICHS AS, HOU M, ROSENBLOOM K, CLAWSON H, SPIETH J, HILLIER LW, RICHARDS S, WEINSTOCK GM, WILSON RK, GIBBS RA, KENT WJ, MILLER W & HAUSSLER D 2005 Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res*, 15, 1034–50. [PubMed: 16024819]
- SUDMANT PH, RAUSCH T, GARDNER EJ, HANDSAKER RE, ABYZOV A, HUDDLESTON J, ZHANG Y, YE K, JUN G, HSI-YANG FRITZ M, KONKEL MK, MALHOTRA A, STUTZ AM, SHI X, PAOLO CASALE F, CHEN J, HORMOZDIARI F, DAYAMA G, CHEN K, MALIG M, CHAISSON MJ, WALTER K, MEIERS S, KASHIN S, GARRISON E, AUTON A, LAM HY, JASMINE MU X, ALKAN C, ANTAKI D, BAE T, CERVEIRA E, CHINES P, CHONG Z, CLARKE L, DAL E, DING L, EMERY S, FAN X, GUJRAL M, KAHVECI F, KIDD JM, KONG Y, LAMEIJER EW, MCCARTHY S, FLICEK P, GIBBS RA, MARTH G, MASON CE, MENELAOU A, MUZNY DM, NELSON BJ, NOOR A, PARRISH NF, PENDLETON M, QUITADAMO A, RAEDER B, SCHADT EE, ROMANOVITCH M, SCHLATT L A, SEBRA R, SHABALIN AA, UNTERGASSER A, WALKER JA, WANG M, YU F, ZHANG C, ZHANG J, ZHENG-BRADLEY X, ZHOU W, ZICHNER T, SEBAT J, BATZER MA, MCCARROLL SA, GENOMES PROJECT C, MILLS RE, GERSTEIN MB, BASHIR A, STEGLE O, DEVINE SE, LEE C, EICHLER EE & KORBEL JO 2015 An integrated map of structural variation in 2,504 human genomes. *Nature*, 526, 75–81. [PubMed: 26432246]
- TALMUD PJ, PALMEN J, PUTT W, LINS L & HUMPHRIES SE 2005 Determination of the functionality of common APOA5 polymorphisms. *J Biol Chem*, 280, 28215–20. [PubMed: 15941721]

- TINGLEY D, YAMAMOTO T, HIROSE K, KEELE L & IMAI K 2014 Mediation: R package for causal mediation analysis.
- TRIGLYCERIDE CORONARY DISEASE GENETICS, C., EMERGING RISK FACTORS, C., SARWAR N, SANDHU MS, RICKETTS SL, BUTTERWORTH AS, DI ANGELANTONIO E, BOEKHOLDT SM, OUWEHAND W, WATKINS H, SAMANI NJ, SALEHEEN D, LAWLOR D, REILLY MP, HINGORANI AD, TALMUD PJ & DANESH J 2010 Triglyceride-mediated pathways and coronary disease: collaborative analysis of 101 studies. *Lancet*, 375, 1634–9. [PubMed: 20452521]
- WANG AZ, LI L, ZHANG B, SHEN GQ & WANG QK 2011a. Association of SNP rs17465637 on chromosome 1q41 and rs599839 on 1p13.3 with myocardial infarction in an American caucasian population. *Ann Hum Genet*, 75, 475–82. [PubMed: 21463265]
- WANG F, XU CQ, HE Q, CAI JP, LI XC, WANG D, XIONG X, LIAO YH, ZENG QT, YANG YZ, CHENG X, LI C, YANG R, WANG CC, WU G, LU QL, BAI Y, HUANG YF, YIN D, YANG Q, WANG XJ, DAI DP, ZHANG RF, WAN J, REN JH, LI SS, ZHAO YY, FU FF, HUANG Y, LI QX, SHI SW, LIN N, PAN ZW, LI Y, YU B, WU YX, KE YH, LEI J, WANG N, LUO CY, JI LY, GAO LJ, LI L, LIU H, HUANG EW, CUI J, JIA N, REN X, LI H, KE T, ZHANG XQ, LIU JY, LIU MG, XIA H, YANG B, SHI LS, XIA YL, TU X & WANG QK 2011b Genome-wide association identifies a susceptibility locus for coronary artery disease in the Chinese Han population. *Nat Genet*, 43, 345–9. [PubMed: 21378986]
- WEINBERG RB, COOK VR, BECKSTEAD JA, MARTIN DD, GALLAGHER JW, SHELNESS GS & RYAN RO 2003 Structure and interfacial properties of human apolipoprotein A-V. *J Biol Chem*, 278, 34438–44. [PubMed: 12810715]
- WILLER CJ, SCHMIDT EM, SENGUPTA S, PELOSO GM, GUSTAFSSON S, KANONI S, GANNA A, CHEN J, BUCHKOVICH ML, MORA S, BECKMANN JS, BRAGG-GRESHAM JL, CHANG HY, DEMIRKAN A, DEN HERTOOG HM, DO R, DONNELLY LA, EHRET GB, ESKO T, FEITOSA MF, FERREIRA T, FISCHER K, FONTANILLAS P, FRASER RM, FREITAG DF, GURDASANI D, HEIKKILA K, HYPONEN E, ISAACS A, JACKSON AU, JOHANSSON A, JOHNSON T, KAAKINEN M, KETTUNEN J, KLEBER ME, LI X, LUAN J, LYYTIKAINEN LP, MAGNUSSON PKE, MANGINO M, MIHAILOV E, MONTASSER ME, MULLER-NURASYID M, NOLTE IM, O'CONNELL JR, PALMER CD, PEROLA M, PETERSEN AK, SANNA S, SAXENA R, SERVICE SK, SHAH S, SHUNGIN D, SIDORE C, SONG C, STRAWBRIDGE RJ, SURAKKA I, TANAKA T, TESLOVICH TM, THORLEIFSSON G, VAN DEN HERIK EG, VOIGHT BF, VOLCIK KA, WAITE LL, WONG A, WU Y, ZHANG W, ABSHER D, ASIKI G, BARROSO I, BEEN LF, BOLTON JL, BONNYCASTLE LL, BRAMBILLA P, BURNETT MS, CESANA G, DIMITRIOU M, DONEY ASF, DORING A, ELLIOTT P, EPSTEIN SE, INGI EYJOLFSSON G, GIGANTE B, GOODARZI MO, GRALLERT H, GRAVITO ML, GROVES CJ, HALLMANS G, HARTIKAINEN AL, HAYWARD C, HERNANDEZ D, HICKS AA, HOLM H, HUNG YJ, ILLIG T, JONES MR, KALEEBU P, KASTELEIN JJP, KHAW KT, KIM E, et al. 2013 Discovery and refinement of loci associated with lipid levels. *Nat Genet*, 45, 1274–1283. [PubMed: 24097068]
- XU C, YANG Q, XIONG H, WANG L, CAI J, WANG F, LI S, CHEN J, WANG C, WANG D, XIONG X, WANG P, ZHAO Y, WANG X, HUANG Y, CHEN S, YIN D, LI X, LIU Y, LIU J, WANG J, LI H, KE T, REN X, WU Y, WU G, WAN J, ZHANG R, WU T, WANG J, XIA Y, YANG Y, CHENG X, LIAO Y, CHEN Q, ZHOU Y, HE Q, TU X & WANG QK 2014 Candidate pathway-based genome-wide association studies identify novel associations of genomic variants in the complement system associated with coronary artery disease. *Circ Cardiovasc Genet*, 7, 887–94. [PubMed: 25249547]
- ZHOU J, XU L, HUANG RS, HUANG Y, LE Y, JIANG D, YANG X, XU W, HUANG X, DONG C, YE M, LIAN J & DUAN S 2013 Apolipoprotein A5 gene variants and the risk of coronary heart disease: a casecontrol study and metaanalysis. *Mol Med Rep*, 8, 1175–82. [PubMed: 23970179]

Table 1.

Demographic and Clinical Features of Study Subjects

Items	Panel 1	Panel 2	
	Premature CAD	Premature CAD	Non-CAD controls
N	128	342	537
Gender, Male, n (%)	83 (64.8%)	222 (64.9%)	260 (48.4%)
Ethnicity	Caucasian	Caucasian	Caucasian
Age, years [†]	38.9 ± 4.5	39.8 ± 5.1	53.4 ± 12.1
Hypertension, n (%)	60 (46.9%)	157 (45.9%)	224 (41.7%)
SBP, mmHg [†]	128.9 ± 18.2	129.0 ± 18.6	130.5 ± 19.4
DBP, mmHg [†]	79.0 ± 12.0	78.0 ± 12.0	74.9 ± 11.0
Tch, mg/dL [†]	227.1 ± 76.8	221.8 ± 60.3	189.0 ± 4.9
HDL-C, mg/dL [†]	38.5 ± 12.2	38.8 ± 11.4	49.7 ± 16.2
LDL-C, mg/dL [†]	131.8 ± 44.3	135.9 ± 43.0	110.7 ± 34.3
Triglyceride (TG), mg/dL [†]	302.5 ± 380.6	248.7 ± 265.8	147.6 ± 97.9
BMI, kg/m ² [†]	30.1 ± 5.1	29.7 ± 5.5	29.0 ± 7.0

SBP, Systolic blood pressure; DBP, Diastolic blood pressure; Tch, total cholesterol; HDL-C, high density lipoprotein cholesterol ; LDL-C, lower density lipoprotein cholesterol. TG, triglyceride; BMI, body mass index.

[†]Data are shown as mean ± SD

Table 2.

11 *APOA5* Variants Identified by Sanger Sequencing Analysis

Variants	Ref/Alt	Effect	HGVSp	Sanger	GRASP	SIFT	Polyphen2	RegulomeDB	CADD	DANN	phyloP46way/ phastCons46way
rs651821	C/T	UTR5	.	0.083	MobCH3 (3.1×10 ⁻¹¹)	.	1f		0.84	0.77	0.05/0.74
rs41338746	C/A	Intronic	.	0.008	.	.	4		0.83	0.87	0.22/0.61
rs41416350	C/T	Intronic	.	0.008	.	.	4		0.75	0.68	0.22/0.53
rs3135506	G/C	Missense	p.S19W	0.047	TG (7.5×10⁻¹⁴)	D	P		1.00	1.00	0.93/0.82
rs144178633	G/T	Synonymous	p.R24R	0.004	.	.	4		0.86	0.84	0.04/0.90
rs34282181	G/T	Missense	p.D37E	0.004	HDL-C (1.0×10 ⁻⁶)	T	B		0.51	0.99	0.24/0.87
rs12287066	G/T	Synonymous	p.I44I	0.054	TG (3.7×10 ⁻⁸⁹)	.	4		0.58	0.68	0.78/0.83
rs138033117	C/T	Missense	p.E99K	0.021	.	D	D		1.00	1.00	0.97/0.99
rs3135507	C/T	Missense	p.V153M	0.061	TG (2.5×10 ⁻⁷)	T	B		0.96	0.99	0.23/0.96
rs146781249	G/C	Synonymous	p.S232S	0.011	.	.	2b		0.02	0.79	0.07/0.92
rs619054	G/A	UTR3	.	0.077	CAD (6.7×10 ⁻³)	.	5		0.87	0.74	0.93/0.85

Effect, location of variant in gene *APOA5*;

HGVSp, amino acid change predicted by program SnpEff;

Sanger, frequency of minor allele of each variant by Sanger sequencing analysis;

GRASP, phenotype and associated p values reported in GRASP (Genome-Wide Repository of Associations between SNPs and Phenotypes)

SIFT, putative functions predicted by SIFT, D = Damaging, T = Tolerated;

Polyphen2, putative functions predicted by Polyphen2, D = Probably Damaging, P=Possibly Damaging, B = Benign;

RegulomeDB, putative regulatory score predicted by RegulomeDB, smaller score means stronger evidence of regulatory roles;

CADD, ranked score of function prediction by CADD, the larger the number the more likely damaging;

DANN, ranked score of function prediction by DANN, the larger the number the more likely damaging;

PhyloP46way, a ranked conservation score, the higher score the more conservative;

PhastCons46way, a ranked conservation score, the higher score the more conservative;

“.”, no available data.

Table 3.Genetic Association analysis of *APOA5* p.S19W with premature CAD

Model	Genotype	Case	Control	OR (95%CI) [†]	p [‡]
Allelic	C	27	58	2.10 (1.17–3.76)	0.013
G	657	1016			
Additive	CC	1	0	2.14 (1.18–3.86)	0.012
CG	25	58			
GG	316	479			
Recessive	CC+CG	26	58	2.30 (1.24–4.25)	0.008
GG	316	479			

[†]Odds ratio and 95% confidence interval estimated by logistic regression model with adjustment for covariates;

[‡]Adjusted p values estimated by logistic regression model with adjustment for covariates.

Table 4.Association analysis of *APOA5* variant p.S19W with lipid traits or BMI

Mediator	Additive model			Recessive model		
	β [†]	<i>s.e.</i> [†]	<i>p</i> [‡]	β [†]	<i>s.e.</i> [†]	<i>p</i> [‡]
Tch	-0.003	0.038	0.935	0.001	0.039	0.982
HDL-C	0.092	0.048	0.059	0.098	0.049	0.048
LDL-C	-0.038	0.057	0.507	-0.032	0.058	0.578
TG	-0.216	0.098	0.027	-0.223	0.010	0.025
BMI	-0.011	0.034	0.755	-0.001	0.034	0.787

Tch, total cholesterol; HDL-C, high-density lipoprotein cholesterol; LDL-C, lower-density lipoprotein cholesterol; TG, triglyceride; BMI, body mass index.

[†]Effect sizes and standard errors estimated by linear regression model;

[‡]Adjusted *p* values estimated by linear regression model with adjustment of covariates (e.g., sex, age, and hypertension).

Table 5.Mediation analysis of *APOA5* variant p.S19W with CAD

Mediator	Effect	Additive Model		Recessive Model	
		OR (95%CI) [†]	<i>p</i> [‡]	OR (95%CI) [†]	<i>p</i> [‡]
HDL-C	Mediation effect	0.98 (0.96–1.00)	0.048	0.98 (0.96–1.00)	0.040
	Direct effect	1.05 (0.97–1.12)	0.250	1.06 (0.97–1.15)	0.230
TG	Mediation effect	0.98 (0.96–1.00)	0.040	0.98 (0.95–1.00)	0.042
	Direct effect	1.07 (0.99–1.14)	0.100	1.08 (0.99–1.17)	0.096

HDL-C, high-density lipoprotein cholesterol; TG, triglyceride.

[†]Odds ratio and 95% confidence interval calculated by mediation analysis with 1,000 times of bootstrapping;[‡]*p* values calculated by mediation analysis with 1,000 times of bootstrapping with adjustment of covariates (e.g., sex, age, and hypertension).