A functional polymorphism affecting the APOA5 gene expression is causally associated with plasma triglyceride levels conferring coronary atherosclerosis risk in Han Chinese Population

Weihua Shou, Ying Wang, Fang Xie, Beilan Wang, Lin Yang, Hong Wu, Yi Wang, Zhimin Wang, Jinxiu Shi, Wei Huang

* Shanghai-MOST Key Laboratory of Health and Disease Genomics, Department of Genetics, Chinese National Human Genome Center, Shanghai, China
b Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

c Department of Cardiology, Shanghai Hospital, The Second Military Medical University, Shanghai, China

ARTICLE INFO

Article history:
Received 31 March 2014
Received in revised form 29 July 2014
Accepted 13 August 2014
Available online 20 August 2014

Keywords:
APOA5 gene
Triglyceride
Coronary artery disease
Single nucleotide polymorphism
Gene expression

ABSTRACT

Apolipoprotein A5 (APOA5) gene plays a key role in plasma triglyceride (TG) metabolism, and shows the involvement in coronary artery disease (CAD). A set of single nucleotide polymorphisms around the APOA5 gene was identified to be associated with plasma TG levels. It is of biological and clinical importance to discern the genuine genetic determinants. A polymorphism in 3' untranslated region of the APOA5 gene, rs2266788, is deserving of investigation for suggestive clues from the association in multiple independent studies. In this study, rs2266788 was genotyped in 3222 unrelated subjects consisting of 2062 CAD cases and 1160 controls. The statistical analyses indicated that the minor C allele of rs2266788 was significantly associated with elevated plasma TG levels and higher CAD risk. In normal human liver tissues, comparison of global APOA5 mRNA levels among genotypes and allelic expression imbalance analysis showed the decreased gene expression for the C allele. Luciferase assays confirmed a concordant result that transcriptional activity was lowered for the C allele compared with the T allele in four cell lines. Multiple lines of evidence in our study supported that rs2266788 was causally associated with plasma TG levels conferring CAD risk in Han Chinese population owing to a cis-acting effect to the APOA5 gene expression.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Plasma triglyceride (TG) levels are usually related to the incidence of multiple cardiovascular diseases, type 2 diabetes mellitus, concomitant metabolic abnormalities of these disorders, and even the efficacy of anti-hyperlipidemic treatment [1]. Elevated plasma TG levels had been recognized as an epidemiological risk to premature coronary artery disease (CAD) [2–5]. Comprehensive understanding of the factors responsible for variation in plasma TG levels is of clinical and biological importance.

Plasma TG concentration is a complex polygenic trait. The studies for genetic determinants of TG metabolism have been extensively carried out and found many genomic loci implicated in regulation of TG concentrations. In total, nearly 40 loci were identified significantly associated with TG levels by large scale genome-wide association studies (GWAS) of blood lipids, candidate target studies and animal model studies [6–8]. Notably, a set of associations were identified in the 11q23.3 region harboring the APOA1/C3/A4/A5 gene cluster, which has been well documented for its involvement in blood lipid metabolism [9]. This locus was evaluated to account for the largest share of genetic contribution to plasma TG variance within human population [10].

Apolipoprotein A5 (APOA5) gene (MIM 606368), encodes an exchangeable apoprotein APOAV, which is exclusively expressed in human liver tissue, and secretes into plasma to play a key role in plasma TG metabolism [11]. The APOA5 gene was a more recently discovered member in the gene cluster through orthologous sequence comparison between human and mouse [12]. Genetically modified mice gave the compelling evidence that a strikingly inverse relationship between apoAV content and plasma TG levels was identified in transgenic and knockout mouse models [12,13]. The effect of sequence variations around the APOA5 gene in humans was also extensively investigated in diverse ethnic groups. Numerous single nucleotide polymorphisms (SNPs) along the APOA1/C3/A4/A5 gene cluster were found with strong evidence of association with blood lipids and cardiovascular diseases.
However, the variants bona fide producing functional consequences have been ill-defined against the proxy SNPs, and the underlying mechanisms to accounting for the statistic associations are much less clear. Hence, it is challenging to recognize causative variants modulating TG levels, and even the potential implication in CAD development.

A SNP in the 3′ untranslated region (UTR) of APOA5, rs2266788, was early found to be associated with plasma TG concentrations across multiple ethnic groups [14,15]. This finding was substantiated by strong evidence from a GWAS of European cohorts including type 2 diabetes patients [16]. Whereas, this SNP was usually less appreciated for a true causative variant. In addition, we appraised the etiologic role of rs2266788 in CAD incidence. It is helpful and necessary to pinpoint the actual genetic determinants for the prediction of dyslipidemia risk, and the prevention against atherosclerosis to CAD endpoint.

2. Material and methods

2.1. Ethics statement

All participants gave written informed consent. The approval to undertake this study was granted by the Ethics Review Committee of the Chinese National Human Genome Center at Shanghai, and it was conducted according to the Declaration of Helsinki Principles. The samples and the collected information were made anonymous prior to testing.

2.2. Population samples and SNP genotyping

Blood samples were collected from 3222 unrelated individuals with self-reported origin of central Han Chinese (Shanghai and the neighboring provinces), consisting of 2062 CAD cases and 1160 controls free of atherosclerosis. The demographic and clinic characteristics of the subjects are summarized in Table 1. Coronary atherosclerosis was diagnosed on the basis of coronary angiography. A consensus diagnosis was presented as the ratio of coronary stenosis and myocardial infarction. Blood lipid profile was determined by using FlexiGene DNA Kit (Qiagen, Valencia, CA, USA). The polymorphism rs2266788 was genotyped by TaqMan method on ABI PRISM 7900HT Sequence Detection System (Applied Biosystems, Inc., Foster City, CA), and allelic discrimination was performed with Sequence Detection Systems (SDS) 2.0 software (Applied Biosystems). Primers are listed in Table S1.

2.3. Liver tissue, gene expression and allelic expression imbalance

A total of 89 normal human liver biopsy samples were collected from unrelated Han Chinese transplant donors and stored in RNAlater (Sigma-Aldrich) at −70 °C. Subjects, ranging from 30 to 45 years old, were morphologically normal and negative for human immunodeficiency virus and hepatitis. No drugs, alcohol and cigarettes were consumed by subjects for at least 30 days prior to surgery. Genomic DNA and total RNA were simultaneously extracted from the liver tissue samples using QIAGEN AllPrep DNA/RNA Mini Kit according to the manufacturer’s instructions. The quality of each RNA sample was confirmed by the presence of strong 18S and 28S bands following agarose gel electrophoresis. Total RNA was treated with RNase-free DNase to remove residual genomic DNA. First-strand cDNA was synthesized using M-MLV Reverse Transcriptase (Promega) in a sterile 25 μl volume containing 0.5 μg of total RNA, 0.5 μM dNTPs, 1 mM Oligo(dT)15, 5 μM random nonamer primers and 20 U ribonuclease inhibitor (Takara).

APOA5 mRNA expression level was quantified by real-time PCR using SYBR Green RT-PCR kit (Takara) on ABI PRISM 7900HT Sequence Detection System. The house-keeping gene, Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene, was used as an internal control. Primers are listed in Table S1. Relative expression levels were calculated using the standard curve method in SDS 2.0 software, and all assays were performed in triplicate for each sample. Quantification of gene expression was presented as the ratio of APOA5 to GAPDH.

Allelic expression imbalance (AEI) was measured by parallel quantitative genotyping of heterozygous SNPs at the gene transcription level, and genomic DNA was used as a comparative reference to calibrate inherent measuring bias in favor of certain allele. All liver samples were genotyped, and heterozygotes were selected to determine AEI levels. SNaPshot method was used to discriminate alleles for both genomic DNA and mRNA. Primers are listed in Table S1. AEI was measured as the ratio of the fluorescence signals from the two alleles. The procedure was described in more detail by Shou et al. [22].

2.4. Plasmid constructs, host cell culture, and luciferase reporter assays

To construct luciferase reporter plasmids, fragments of the entire APOA5 3′ UTR of different rs2266788 alleles were cloned into the pGL3-Promoter luciferase reporter vector (Promega, Madison, WI). The target fragments were amplified by PCR from genomic DNA templates with known genotype, and then inserted into the Xbal restriction site of the vector, downstream of the luc+ transcriptional unit (Fig. 3A). The PCR primers are listed in Table S1. The recombinant constructs were verified by direct sequencing.

Human hepatoma cell lines BEL-7402, BEL-7405, Huh7 and HepG2 were used for transient transfection to assay luciferase reporter activity. RPM1640 medium for BEL-7402 and BEL-7405, Dulbecco’s modified Eagle’s medium (DMEM) for Huh7, and minimal essential Eagle’s medium (MEM) for HepG2 were used in cell culture, supplemented with 10% fetal bovine serum (FBS),

<table>
<thead>
<tr>
<th>Table 1: Clinical characteristics of the study subjects.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics</td>
</tr>
<tr>
<td>Male/female</td>
</tr>
<tr>
<td>Age (years)</td>
</tr>
<tr>
<td>Cases/controls</td>
</tr>
<tr>
<td>TC (mmol/L)</td>
</tr>
<tr>
<td>HDL (mmol/L)</td>
</tr>
<tr>
<td>LDL (mmol/L)</td>
</tr>
<tr>
<td>TG (mmol/L)</td>
</tr>
</tbody>
</table>

Note: TG, triglyceride; TC, total cholesterol; HDL, high-density lipoprotein-cholesterol; LDL, low-density lipoprotein-cholesterol. Data are shown as mean ± standard deviation.
100 IU/ml penicillin, and 100 μg/ml streptomycin. Cells were seeded into 24-well plates at a density of 1 × 10^5 per well. After overnight culture, cells were co-transfected with 0.4 μg of each allele reporter constructs and 0.03 μg Renilla luciferase pRL-SV40 plasmid as an internal control per well, using jetPEI transfection reagent. Cells were harvested 48 h after transfection and lysed with Passive Lysis Buffer (Promega) for subsequent reporter gene assays. Each assay was performed in quadruplicate. The lysate was used to detect luciferase activity on Microlumat Plus LB 96 V luminometer (Berthold Technologies, Bad Wildbach, Germany) with Dual-Luciferase Reporter Assay System (Promega).

2.5. Statistical analysis

Deviation from Hardy–Weinberg equilibrium (HWE) was assessed by the chi-square test (1 df). Actual measures of plasma lipid traits were statistically transformed (TG and HDLC: log transformed; LDL-C: square root transformed; TC: natural log transformed) to approximate a normal distribution and obtain equality of variance.[23] (Fig. S1). Statistical association analyses were performed using program PLINK v1.07 (http://pngu.mgh.harvard.edu/~purcell/plink/). A linear regression model was applied to assess the association with quantitative measures of blood lipid traits under an additive genetic model, with adjustments for age, sex, and CAD status. The genotypic distribution between cases and controls was compared by Cochran–Armitage trend test. The associations between the genotypes and CAD risk, odds ratios (ORs) and 95% confidence intervals (CIs) were evaluated using binary logistic regression model with adjustments for gender, age, and TG level. The Haploview program [24] was used to analyze pairwise linkage disequilibrium (LD) with data extracted from 1000 Genomes datasets (http://www.1000genomes.org/home). The gene expression difference was compared by one-way ANOVA with Tukey's post hoc test using SPSS 13.0 package (SPSS Inc., Chicago, IL). Values are expressed as mean ± standard deviation (SD), and P < 0.05 was considered significant.

3. Results

3.1. Association of the polymorphism with plasma TG levels and coronary atherosclerosis

The SNP rs2266788 (C/T) was genotyped in all 3222 Han Chinese individuals. No deviation from HWE was shown for this SNP in controls or all samples (P > 0.05) (Table S2). The frequency of the minor allele C in our samples was 23.5%, which is fully concordant with the frequency (22.1%) in the CHB (Han Chinese in Beijing) population from HapMap III (http://hapmap.ncbi.nlm.nih.gov/). Through multiple linear regression analysis, strong evidence was identified for a significant association between polymorphism rs2266788 (P = 1.02 × 10^{-13}) and plasma TG levels in the Han Chinese population (Table 2). The copy number of the C allele was associated with increased TG concentration (Fig. S2), considering the significant difference in TG levels (Table S3) between the case and control groups (P < 0.001) and between two gender cohorts (P < 0.01), age, sex, and CAD status were included as covariates in multiple linear regression models. The significance of association between the variant and TG levels remained strong (P = 5.17 × 10^{-13}), exceeding the genome-wide significance cut-off (P < 5 × 10^{-8}). In stratified analyses (Table 2), the association of this polymorphism with plasma TG concentration was overall significant in controls (P = 1.01 × 10^{-7}), cases (P = 3.48 × 10^{-7}), females (P = 4.68 × 10^{-6}), and males (P = 6.60 × 10^{-5}). The association between rs2266788 and other blood lipids was also tested (Table S4). The SNP was associated with plasma HDLC levels (P = 0.005) with a modest effect size, but the direction of the effect was inverse to that of the effect on plasma TG levels. This result is consistent with the previous observation that plasma APOAV levels are positively correlated with HDLC and negatively correlated with TG concentration[25].

The relevance of rs2266788 to CAD incidence was also investigated. Cochran–Armitage trend tests were performed in the total samples, in females, and in males (Table 3). The association was nearly significant (OR = 1.119, 95% CI 0.992–1.263, Prand = 0.066) in the total samples, suggesting the susceptibility of this variant to CAD development, and a possible risk factor from TG-raising C allele. The minor C allele was significantly associated with an increased CAD risk in the female cohort (OR = 1.267, 95% CI 1.044–1.523, Prand = 0.005), but not in the male cohort. In logistic regression analyses (Table 4), the association of rs2266788 with CAD became significant after adjustment for age and sex in the total samples (OR = 1.151, 95% CI 1.012–1.309, P = 0.033). The association became stronger and the minor C allele conferred a larger risk in females (OR = 1.397, 95% CI 1.139–1.714, P = 0.001) after including age as a covariate in the logistic model. Noticeably, TG levels were significantly higher in females than in males among our subjects (Table S3). For a stronger predisposition of elevated TG to CAD, the genetic effect seemed more significant in TG-higher subjects. Similarly, in an investigation of rs2075291 (c.553G>T; p.185Gly>Cys) among Asian-American hypertriglyceridemia patients, differences in plasma lipid levels across genotypes were greater in the high-TG group than in the low-TG group[26]. Furthermore, after adjustment for TG levels, the association became nonsignificant in the total samples and was attenuated in female subjects (Table 4). The contrastive results with and without adjustment for TG provide a statistic evidence of the effects of rs2266788 in plasma TG regulation and CAD incidence, and show that rs2266788 is a genetic risk factor for CAD mediated through TG levels.

3.2. LD analyses to dissect nearby associations

Numerous SNPs in the 11q23.3 region have been reported to be associated with TG levels. To gain insight into the LD pattern of rs2266788 and other SNPs with strong evidence of association, the data for these SNPs (Table S5) were extracted from 1000 Genomes Phase I datasets for LD analyses. Strikingly, rs2266788 and multiple SNPs were in strong LD in CHB population (Fig. 1). A battery of significant SNPs, rs1558861, rs10790162, rs964184, rs8589566, rs2075290, rs2266788, rs2072560, rs651821, and rs662799, virtually represented the same statistical association signal. In particular, rs2266788 was in perfect LD with rs2072560, rs2075290, rs6589566, and rs964184.

Table 2

<table>
<thead>
<tr>
<th>Sample</th>
<th>N</th>
<th>Beta (95% CI)</th>
<th>P value</th>
<th>Betaadj (95% CI)^2</th>
<th>Prand value^3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>3222</td>
<td>0.051 (0.037–0.064)</td>
<td>1.02 × 10^{-13}</td>
<td>0.049 (0.036–0.062)</td>
<td>5.17 × 10^{-13}</td>
</tr>
<tr>
<td>Controls</td>
<td>1160</td>
<td>0.059 (0.037–0.081)</td>
<td>1.08 × 10^{-07}</td>
<td>0.059 (0.038–0.081)</td>
<td>1.01 × 10^{-07}</td>
</tr>
<tr>
<td>Cases</td>
<td>2062</td>
<td>0.045 (0.028–0.062)</td>
<td>1.54 × 10^{-07}</td>
<td>0.043 (0.027–0.06)</td>
<td>3.48 × 10^{-07}</td>
</tr>
<tr>
<td>Females</td>
<td>1236</td>
<td>0.052 (0.031–0.073)</td>
<td>1.83 × 10^{-06}</td>
<td>0.05 (0.029–0.071)</td>
<td>4.68 × 10^{-06}</td>
</tr>
<tr>
<td>Males</td>
<td>1986</td>
<td>0.051 (0.033–0.068)</td>
<td>6.85 × 10^{-05}</td>
<td>0.05 (0.033–0.067)</td>
<td>6.60 × 10^{-05}</td>
</tr>
</tbody>
</table>

CI, confidence interval; CAD, coronary artery disease.

^2 Statistics in multiple linear regression models with adjustments for age, sex, and CAD status (adjustment for age and sex in control and case groups; adjustment for age and CAD status in female and male cohorts).

^3 Prand value.
different and more complex LD pattern (Fig. S3) was demonstrated in CEU (Utah residents with Northern and Western European ancestry from the CEPH collection) population, and these associations were poorly dissected in previous studies of Europeans. The polymorphism rs3135506 (c.553G>T; p.195Ser>Try) and other SNPs in LD was absent in CHB (Fig. 1), and rs3135506 was not in LD with rs2266788 in CEU (Fig. S3). Thus, rs3135506 might be an independent association in populations with European ancestry. Likewise, rs2075291 (c.553G>T; p.185Gly>Cys) was not in LD with other SNPs in CHB and was monomorphic in CEU.

3.3. Reduced gene expression of the C allele in vivo and in vitro

We examined the APOA5 mRNA expression level in normal human liver samples, and compared the expression among each genotype of rs2266788. The results showed that the APOA5 gene expression was lower in samples carrying the C allele, and there was 60% reduction in the gene expression for CC homozygotes (Fig. 2A). Therefore, the evidence based on transcriptional activity between the two alleles of the same source, prevents interference from trans-acting regulators or environmental confounders, and improves the sensitivity to identify cis-acting effects even with limited sample size [27]. To survey whether relative expression imbalance occurred between the two alleles of rs2266788 in human liver samples, we inquired into heterozygous subjects. Almost a uniform deviation of allelic expression ratios in mRNA from the ratios in genomic DNA was observed, demonstrating unequal transcript abundance between the two alleles (Fig. 2B). The APOA5 expression level from the C allele was significantly lower than that from the T allele \((P = 5.15 \times 10^{-5})\), and the C allele showed about 70% of the transcriptional activity of the T allele. The differential allelic expression strongly suggested a cis-acting effect of this polymorphism.

In order to further examine whether rs2266788 is a functional variant leading to alteration of the APOA5 mRNA expression, in vitro luciferase reporter assays were performed. Luciferase reporter vectors carrying different alleles of the APOA5 3’ UTR were constructed (Fig. 3A), and pairwise comparison of their relative transcriptional activity was performed separately in BEL-7402, BEL-7405, HuH7 and HepG2 cell lines from human hepatocellular carcinoma. As expected, the plasmid with the C allele displayed significantly lower luciferase expression level than the plasmid with the T allele in all cell lines (Fig. 3B). The C allele also had about 70% of the transcriptional activity of the T allele, which was completely in agreement with the results of the AEI analysis.

Table 3
Allelic analyses of the association between rs2266788 and CAD by Cochran-Armitage trend tests.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Cases (C/T)*</th>
<th>Controls (C/T)*</th>
<th>P \text{trend}</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1000/3124</td>
<td>516/1804</td>
<td>3.378</td>
<td>1.119 (0.992–1.263)</td>
</tr>
<tr>
<td>Females</td>
<td>320/968</td>
<td>246/938</td>
<td>5.94</td>
<td>0.015 (1.044–1.523)</td>
</tr>
<tr>
<td>Males</td>
<td>680/2156</td>
<td>270/866</td>
<td>0.02</td>
<td>1.012 (0.861–1.189)</td>
</tr>
</tbody>
</table>

OR, odds ratio; CI, confidence interval. P value < 0.05 is shown in bold.
* Allelic counts in cases or controls.

4. Discussion

Results from nuclear families [28] and twin studies [29] indicated a considerable genetic contribution to the variation in plasma TG levels. Undoubtedly, the APOA5 gene accounts for a pronounced part of the contribution. Previous studies offered us a suggestive clue to inquire into how the genetic effects of the APOA5 gene result in variable plasma TG levels. The SNP rs2266788 became the promising target to investigate. An understanding of the genetic determinants underlying TG regulation and the risk of CAD development could prompt the clinic translation to early prevention and novel therapeutic intervention to dyslipidemia and CAD.

4.1. The association of rs2266788

In this study, we have provided strong evidence that rs2266788, a functional variant in 3’ UTR of the APOA5 gene, is significantly associated with plasma TG levels conferring CAD risk in the Han Chinese population owing to its regulatory effect on transcription of the APOA5 gene. We appreciated the potential involvement of the plasma TG-regulating genetic variant in CAD pathogenesis, and found that the TG-raising allele was associated with increased predisposition to CAD. The results reinforce the understanding that genetic determinants of TG concentration may be relevant to CAD susceptibility.

Polymorphism rs2266788 was significantly associated with plasma TG levels, not only in the total samples but also in CAD patients, normal controls, female and male cohort. The minor C allele was strongly associated with increased TG concentration. The association and the direction of the effect have been observed consistently across multiple population samples of diverse ethnicities [14,15]. The replicable results suggested that this variant was probably a genetic cause of plasma TG variation. Thus, we explored the mechanism by which the variant modulates TG levels. An association between rs2266788 and the APOA5 gene expression was established. In normal liver tissues, global APOA5 mRNA levels were lower in subjects carrying the C allele and were the lowest for CC genotype. In subjects heterozygous for rs2266788, obvious AEI between the two alleles of the APOA5 gene was detected, in that the gene expression from the same source was lower for the C allele relative to the T allele. Luciferase assays confirmed a similar result that transcriptional activity was reduced for the C allele in comparison with the T allele. Therefore, the evidence based on in vivo and in vitro experiments supported rs2266788 might regulate in cis the APOA5 mRNA expression. An association between rs2266788 and CAD incidence was also identified in our samples, especially in female cohort with higher TG levels. The genetic structure in Han Chinese populations across extensive geographic locations using genome-wide SNPs has been well constructed, and revealed that geographic matching might be a good proxy for genetic matching [30]. The subjects in the present study could be clustered into the central Han subgroup with a high degree of resemblance according to geographic origin. Therefore, the association with CAD was less likely to be spurious by population stratification. High plasma TG level was supposed to be a predisposing factor for coronary atherosclerosis. An equal increment in TG levels conferred a similar increase in CAD risk. This point was reinforced by the stronger association with CAD in females with higher TG levels and
the loss of the significant association after adjustment for TG levels in logistic regression analyses (Table 4). The most recent result from a study of large-scale meta-analysis on lipid loci illustrated that the strength of a SNP effect on TG levels was correlated with the magnitude of its effect on CAD risk, which genetically suggests that TG levels causally influence the risk of CAD [31]. Multiple lines of evidence support a causal association of the cis-acting SNP rs2266788 with plasma TG levels and CAD incidence for the TG-raising allele conferring a higher CAD risk.

Fig. 1. LD plot of SNPs with strong evidence of association within the 11q23.3 region in CHB of 1000 Genome Phase I. The strength of pairwise LD is in correspondence with $r^2$ metrics.

Fig. 2. APOA5 mRNA expression difference in normal human liver samples. (A) The APOA5 expression levels among genotypes of rs2266788. The relative APOA5 expression levels were normalized to GAPDH. (B) Allele-specific expression of APOA5 in liver samples. Allelic ratios of rs2266788 were measured in mRNA and genomic DNA for heterozygous subjects, and ratios in mRNA were normalized to those in DNA.
4.2. The functionality of rs2266788

The chromosomal region 11q23.3, harboring the APOA1/C3/A4/A5 gene cluster, produces the strongest association signals in GWAS of plasma TG levels. In particular, a dense set of SNPs, spanning ~60 kb APOA1/C3/A4/A5 gene cluster, was reported to be associated with TG levels. However, strong and complex patterns of LD extension in the gene cluster bring about obstacles to elucidate the true underlying variants and the corresponding genes responsible for these associations in the lack of experimental evidence for functional analysis.

Lipoprotein lipase (LPL) was early recognized as the fulcrum of plasma TG metabolism for hydrolyzing TG-rich lipoprotein particles [32]. The liver-expressed APOAV enhances LPL activity, facilitates LPL-mediated TG clearance, and inhibits the hepatic secretion of very low-density lipoprotein, the major carrier of TGs [33–35]. On the other hand, the results from genetically modified mouse models showed that apoCIII inhibited LPL-mediated TG hydrolysis [36]. The gene APOC3 and APOA5, in the same APOA1/C3/A4/A5 lipid gene cluster in the human genome, play separate but opposing roles in determining plasma TG concentrations [37]. Common variants of the APOC3 gene have also been associated with TG levels in humans [9]. Thus, an important problem is that whether the APOA5 SNPs simply reflect LD with functional variants of APOC3, or whether they stand for independent associations. Previous reports proposed a strong LD between APOA5 rs662799 and APOC3 rs2854117 (−482C>T) [38,39] which influenced TG levels by disrupting the insulin response element of APOC3 [40]. Nevertheless, in the present LD analysis (Fig. 1; Fig. S3), there was no strong correlation between the APOA5 SNPs with association and rs2854117, suggesting that APOA5 SNPs rs662799, rs651821 and rs2266788 could not mark the functional effect of APOC3 −482C>T.

The SNP rs964184, in perfect LD with rs2266788, produced a replicable top-ranked association with plasma TG levels, certain component traits of metabolic syndrome and CAD risk in large-scale GWAS [6,17,41,42]. The variant resides in the downstream of a zinc finger protein gene ZNF259, which is involved in cell proliferation and signal transduction [43]. Because of the location and a lack of persuasive evidence from functional exploration, rs964184 was hardly legitimated as the actual causative variant. Although an association was found between rs964184 and serum APOAV protein levels in Mexicans [42], no molecular mechanism of the variant was provided, but still a statistic association with the quantity of TG-regulating protein. Two statistically significant SNPs in the APOA5 gene promoter, rs662799 (−1131T→C) and rs651821 (−3A→G), were in strong LD with rs2266788. Their position in the gene instinctively makes them relevant to the gene transcription and translation. A functional E-box binding upstream stimulatory factors modulated by insulin via the phosphatidylinositol 3-kinase pathway [44], and farnesoid X receptor and peroxisome proliferator-activated receptor α response elements were characterized in the upstream of the gene [45]. Unfortunately, the two variants did not reside within or in proximity to any obvious transcription factor binding sites. Although a large number of the statistical associations of rs662799 and rs651821 were reported, whether these variants are causative remains uncertain to come to a definite conclusion. In assays of individual variants, rs662799 and rs651821 did not significantly alter luciferase activity, and rs651821 also did not influence in vitro translation efficiency despite a base change in the Kozak sequence [20]. A more detailed examination of the diversely combined effects of rs662799, rs651821, and rs2266788 on luciferase activity showed an interaction among them, and suggested a driving effect from the combination of rs651821 and rs2266788 [21]. After resequencing the exons plus exon/intron boundaries of the APOA5 gene, no other potential common variants were found that could account for the association of rs662799, rs651821, or rs2266788 [46]. Therefore, it is reasonable to approve rs2266788 to be a variant with functional consequence, though interaction with other SNPs cannot be completely excluded. Excitingly, a latest study uncovered the molecular mechanism for the alteration of the APOA5 gene expression by this variant. The gene expression was lowered by the minor C allele creating a potential miRNA binding site for liver-expressed miR-485-5p [47]. The identification of rs2266788-mediated posttranscriptional regulation of the APOA5 gene supported the present results.

The coding SNP rs3135506 introduces a nonsynonymous change from polar amino acid serine to aromatic tryptophan at position −5 from the cleavage site of the APOAV signal peptide. This residue change

Fig. 3. Luciferase assay for two alleles of rs2266788 in four cell lines. (A) Schematic diagram of luciferase reporter constructs containing different alleles of the APOA5 3′ UTR. (B) Paired comparison of relative luciferase activity between two alleles in four cell lines. The luciferase activity of the T allele was set as the reference in each cell line, and relative values for the C allele are obtained by normalization to the T allele. The results represent mean ± SD in each cell line.

The SNP rs964184, in perfect LD with rs2266788, produced a replicateable top-ranked association with plasma TG levels, certain component traits of metabolic syndrome and CAD risk in large-scale GWAS [6,17,41,42]. The variant resides in the downstream of a zinc finger protein gene ZNF259, which is involved in cell proliferation and signal transduction [43]. Because of the location and a lack of persuasive evidence from functional exploration, rs964184 was hardly legitimated as the actual causative variant. Although an association was found between rs964184 and serum APOAV protein levels in Mexicans [42], no molecular mechanism of the variant was provided, but still a statistic association with the quantity of TG-regulating protein. Two statistically significant SNPs in the APOA5 gene promoter, rs662799 (−1131T→C) and rs651821 (−3A→G), were in strong LD with rs2266788. Their position in the gene instinctively makes them relevant to the gene transcription and translation. A functional E-box binding upstream stimulatory factors modulated by insulin via the phosphatidylinositol 3-kinase pathway [44], and farnesoid X receptor and peroxisome proliferator-activated receptor α response elements were characterized in the upstream of the gene [45]. Unfortunately, the two variants did not reside within or in proximity to any obvious transcription factor binding sites. Although a large number of the statistical associations of rs662799 and rs651821 were reported, whether these variants are causative remains uncertain to come to a definite conclusion. In assays of individual variants, rs662799 and rs651821 did not significantly alter luciferase activity, and rs651821 also did not influence in vitro translation efficiency despite a base change in the Kozak sequence [20]. A more detailed examination of the diversely combined effects of rs662799, rs651821, and rs2266788 on luciferase activity showed an interaction among them, and suggested a driving effect from the combination of rs651821 and rs2266788 [21]. After resequencing the exons plus exon/intron boundaries of the APOA5 gene, no other potential common variants were found that could account for the association of rs662799, rs651821, or rs2266788 [46]. Therefore, it is reasonable to approve rs2266788 to be a variant with functional consequence, though interaction with other SNPs cannot be completely excluded. Excitingly, a latest study uncovered the molecular mechanism for the alteration of the APOA5 gene expression by this variant. The gene expression was lowered by the minor C allele creating a potential miRNA binding site for liver-expressed miR-485-5p [47]. The identification of rs2266788-mediated posttranscriptional regulation of the APOA5 gene supported the present results.

The coding SNP rs3135506 introduces a nonsynonymous change from polar amino acid serine to aromatic tryptophan at position −5 from the cleavage site of the APOAV signal peptide. This residue change
variants or related consequence. One urgent challenge is to pinpoint actual causative levels in Han Chinese population, which was consistent with previous through 11q23.3 region, for which experimental evidence of function is lacking. Consequently, rs2266788 is significantly associated with plasma TG levels conferring CAD risk in Han Chinese population owing to its cis-acting effect to the APOA5 gene expression. Accordingly, this work reasonably prioritized rs2266788 as a causative variant.

**Funding**

This work was supported by Shanghai Postdoctoral Scientific Program (12R21421100) and the National Key Talent Project (130041).

**Conflict of interest**

There was no conflict of interest for the authors in the present study.

**Acknowledgements**

We would like to thank all the contributors that made this research possible. We are grateful to all of the donors participating in this study. We also thank all the clinicians for their assistance to recruiting participants to the study.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.abb.2014.08.006.

**References**


