Original article

CELSR2–PSRC1–SORT1 gene expression and association with coronary artery disease and plasma lipid levels in an Asian Indian cohort

Prathima Arvind (MSc)\textsuperscript{a}, Jiny Nair (MSc)\textsuperscript{a}, Srikarthika Jambunathan (PhD)\textsuperscript{a}, Vijay V. Kakkar (DSc)\textsuperscript{b, c}, Jayashree Shanker (PhD)\textsuperscript{a, *}

\textsuperscript{a} Mary and Garry Weston Functional Genomics Unit, Thrombosis Research Unit, Bangalore 560099, India
\textsuperscript{b} Thrombosis Research Institute, Bangalore 560099, India
\textsuperscript{c} Thrombosis Research Institute, London, UK

\textbf{A B S T R A C T}

\textbf{Background:} Genetic regulation of plasma lipids has been shown to influence the risk of coronary artery disease (CAD). We analyzed the relationship between rs599839 and rs646776 single nucleotide polymorphisms (SNPs) present in the CELSR2–PSRC1–SORT1 gene cluster, candidate gene expression, and their association with CAD and circulating lipid levels in a representative cohort of Asian Indians selected from the Indian Atherosclerosis Research Study.

\textbf{Methods:} SNPs rs599839 and rs646776 were genotyped by Taqman assay in 1034 CAD patients (cases) and 1034 age- and gender-matched controls. Expression of CELSR2, PSRC1, and SORT1 genes was measured in 100 cases and 100 controls. Plasma levels of total cholesterol (TC), triglycerides, high-density lipoprotein-cholesterol, and low-density lipoprotein-cholesterol (LDL-c) were measured by enzymatic assay.

\textbf{Results:} Both rs646776 and rs599839 were in strong linkage disequilibrium (r² = 0.98) and showed significant protective association with CAD (OR = 0.315, 95% CI 0.136–0.728, p = 0.007 and OR = 0.422, 95% CI 0.181–0.981, p = 0.045, respectively). Haplotype TA showed 72% frequency and was associated with CAD (OR 0.77, 95% CI 0.67–0.88, p = 0.0002). PSRC1 gene expression was lower in the cases than in the controls (0.75 ± 0.405 versus 1.04 ± 0.622, p = 2.26 × 10⁻⁴). The homozygous variant and heterozygous genotypes showed 30% and 15% higher PSRC1 expression, respectively. Correspondingly, the minor alleles were associated with lower plasma TC and LDL-c levels.

\textbf{Conclusion:} PSRC1 in the cholesterol gene cluster shows a significant association with CAD by virtue of the two SNPs, rs646776 and rs599839 that also regulate plasma cholesterol levels.

\textcopyright 2014 Japanese College of Cardiology. Published by Elsevier Ltd. All rights reserved.

\section*{Introduction}

Coronary artery disease (CAD) is reported to be one of the leading causes of death in young Indians in the 30–50 year age group [1]. Identification of predisposing genetic factors that promote the development of premature CAD should eventually help in reducing the overall mortality rates by promoting targeted therapy. In this regard, genome wide association studies (GWAS) in diverse ethnic populations have been instrumental in the discovery of numerous genetic variants for CAD [2–4] in addition to confirming previously identified markers like those in the 9p21 locus [3,4]. Apart from traditional risk factors such as age, hypertension, and diabetes, heritable factors such as genes that regulate plasma lipid levels underscore the significant role of lipids in the development of CAD [5,6]. Studies have identified single nucleotide polymorphisms (SNPs) in novel genetic regions that affect plasma lipid levels [7–9] and some of those SNPs reside within putative candidate genes that are known to regulate the lipid metabolism [10,11].

The gene cluster, located on chromosome 1p13.3 region (in \textit{Homo sapiens}) comprises three distinct genes, CELSR2 (cadherin, EGF LAG seven-pass G-type receptor 2), PSRC1 (proline/serine rich coiled coil 1), and SORT1 (sortilin 1). The CELSR2 gene encodes a non-classic type of cadherin involved in contact-mediated cell adhesion and receptor–ligand interactions [12]. The PSRC1 gene product plays a role in microtubule destabilization [13] while the SORT1 gene encodes for the sortilin protein that plays an important role in the uptake of lipids [14]. A variant, rs599839 was initially identified through GWAS study in the intergenic region

\textsuperscript{*} Corresponding author at: Mary and Garry Weston Functional Genomics Unit, Thrombosis Research Institute, Anekal Taluk, Bangalore 560099, Karnataka, India. Tel.: +91 80 27835303; fax: +91 80 27835302.
\textit{E-mail address:} jayashreeshanker@triindia.org.in (J. Shanker).

\textit{http://dx.doi.org/10.1016/j.jjcc.2014.02.012}

0914-5087/© 2014 Japanese College of Cardiology. Published by Elsevier Ltd. All rights reserved.
between CELSR2 and PSRC1 and was associated with increased risk of CAD [15]. Subsequently, variants in this cholesterol gene cluster (rs646776 and rs599839) have been associated with low low-density lipoprotein cholesterol (LDL-c) levels [8,16]. Further, a recent replication study conducted on Asian Indians from the Sikh Diabetes Study (SDS) cohort reported a significant association between variants in the lipid regulating gene loci and the lipid profiles including high-density lipoprotein cholesterol (HDL-c) and triglyceride (TG) levels [17]. These studies highlight the significant role played by these SNPs and emphasize the need to replicate these findings in other populations with the possibility of translating these findings into clinically meaningful outcomes. We have examined the association of two SNPs, rs599839 and rs646776 with CAD, the expression of CELSR2, PSRC1, and SORT1 genes and fasting lipid levels in a representative cohort of Asian Indians.

Materials and methods

Study population

The study subjects (cases and controls) were selected from the ongoing Indian Atherosclerosis Research Study (IARS), an epidemiological study that was initiated in January 2004 with the objective of understanding the contributions of traditional and emerging risk factors of premature CAD in Asian Indians residing in the Indian sub-continent. The design and overview of the IARS have been previously published [18]. Briefly, CAD patients were enrolled from two major cardiology specialty hospitals, Narayana Health, Bangalore and the Asian Heart Institute and Research Center, Mumbai as well as from the local clinics/hospitals in Bangalore between 2004 and 2011. CAD patients had a strong family history of cardiovascular disease, with males having age at CAD onset <60 years and females <65 years. Presence of CAD was based on the report of cardiologists’ diagnosis and included any of the following criteria – angiographically confirmed presence of disease, with/without history of myocardial infarction and treated with standard medications, percutaneous coronary intervention (PCI), or coronary angiography bypass graft (CABG). Controls were matched for gender and age (±3 years) and were enrolled from the same geographical area as that of the cases. They were healthy, clinically asymptomatic for CAD, showed normal electrocardiography (ECG) readings and did not have a family history of cardiovascular disease. All the participants were above 18 years, with family living in the Indian sub-continent for at least two generations, did not have any other major illness such as cancer, liver disease, or primordial heart diseases such as cardiomyopathies or congenital heart disease, and were free of concomitant infections such as cold and cough at the time of enrolment. Written informed consent was obtained from all the participants. The IARS protocol has been approved by the Thrombosis Research Institute Ethics committee and is based on the bioethics guidelines of the Indian Council of Medical Research [19]. A total of 1034 cases and 1034 age- and gender-matched controls, were selected for the association study, while an additional 100 cases and 100 controls were included for the gene expression study.

Sample preparation

Venous blood samples were collected after an overnight fast from all the study participants. Aliquots of serum and plasma were stored at −80 °C. Genomic DNA was extracted by a modified salting out procedure [20]. Total RNA was extracted from fresh 3 ml ethylenediaminetetraacetic acid whole blood sample using the QIAmp RNA blood mini kit (Qiagen, Valencia, CA, USA), following manufacturer’s instructions. Both genomic DNA and total RNA were quantitated using NanoDrop ND-1000 Spectrophotometer (NanoDrop, Wilmington, DE, USA). RNA samples were further digested with DNase I followed by first strand synthesis of cDNA using the cDNA Archive kit (Applied Biosystems, Foster City, CA, USA).

Lipid assays

Serum total cholesterol (TC) and TG were estimated by standard enzymatic analysis on the Cobas-Fara II Clinical Chemistry Auto analyzer (F. Hoffman La Roche Ltd., Basel, Switzerland). HDL-c concentrations were estimated after precipitating the non-HDL fractions with a mixture of 2.4-mmol/l phosphotungstic acid and 39 mmol/l magnesium chloride (Bayer Diagnostics, India). Plasma LDL-c was calculated using the Friedewald formula [21]. The inter-assay coefficient of variation for the commercial controls and normal serum pool ranged from 4.9% to 7.0% for TC, 6.1% to 7.7% for TG, and 7.1% to 12.2% for HDL-c.

The atherogenic index of plasma (AIP) was estimated using the formula [AIP = log(TG/HDL-c)] [22]. Subjects were classified into three groups based on their AIP scores as follows: Group 1 (low atherogenic risk): <0.011; Group 2 (intermediate risk): 0.11–0.21; Group 3 (high risk): >0.21. Further, all the participants in the association cohort were classified into those with or without metabolic syndrome based on standard Adult Treatment Panel (ATP) III guidelines (National Cholesterol Education Program-ATP-III, 2001) [23] as well as modified ATPIII criteria [24]. The primary difference between the two methods of classification is that abdominal obesity defined by waist circumference of WC >102 cm (40 in.) in men and >88 cm (35 in.) in women was revised to >90 cm for men and >80 cm for women based on publications that have shown that the revised cut-offs represent abdominal obesity better for the Asian Indian population due to their unique pattern of fat distribution [25].

Genotyping of CELSR2–PSRC1–SORT1 variants

Two SNPs in the CELSR2–PSRC1–SORT1 gene loci namely rs646776 and rs599839 were genotyped by TaqMan allelic discrimination assay on a 7900HT Fast Real Time PCR instrument (Applied Biosystems). In-house controls with known genotype were used as positive control while a no-template reaction was assigned as the negative control and run with every batch of Taqman assay.

Gene expression assay

Expression levels of CELSR2, PSRC1, and SORT1 genes were measured using Taqman real-time PCR assay. The FAM and TAMRA dye labeled probe-primer mix (20 ×) and the Universal Master Mix (2 ×) were purchased from ABI (Applied Biosystems). β-Actin was used as an endogenous control and run along with every sample. All assays were done in duplicates and according to manufacturer’s instructions. Data were processed using SDS v2.3 (sequence detection software) (Applied Biosystems). RQ Manager was used to calculate the relative changes in mRNA expression levels between the cases and the controls. The assay was repeated in duplicates for samples showing skewed expression pattern while the persistent outliers were excluded from further analysis.

Network analysis

We generated a protein interaction network with the PSRC1 as seed gene using the software Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) version 9.05 (http://string-db.org) [26,27] in order to identify its interacting partners.
Statistical methods

Statistical analysis was done using SPSS v17.0 software (SPSS Inc., Chicago, IL, USA). Continuous variables were expressed as mean ± standard error of the mean (SEM). Chi-square test was used for testing the association of CELSR2–PSRC1–SORT1 variants with CAD and for estimating the odds ratios (ORs) and 95% confidence intervals (CIs), while Student’s t-test and univariate analysis were used to test for the mean differences in the expression levels and other quantitative traits between the cases and the controls. Age, gender, diabetes, hypertension, plasma lipids, body mass index (BMI) and smoking were considered as covariates and appropriately adjusted for during multivariate analysis to test for the association of the two SNPs with plasma lipid levels. SNPstats online software was also used to calculate allele and genotype frequencies, Hardy–Weinberg equilibrium and genotype/haplotype associations with CAD (http://bioinfo.iconcologia.net/SNPstats) [28]. Estimation of linkage disequilibrium (LD) was performed using the Haplovew v3.32 software (http://www.broad.mit.edu/mpg/haplovew) [29].

Results

Clinical characteristics of study participants

The clinical profile of subjects enrolled in (a) genotype–phenotype association study and (b) gene expression study is summarized in Table 1. In the gene expression study cohort, after removal of outliers and individuals with missing information, data on 98 cases and 97 controls were included in the final analysis. Males showed higher representation (76–92%) than females (24–8%). The average age of the cases and the controls and the age at onset of CAD were around 50 years. Classical risk factors such as diabetes and hypertension were more prevalent in the cases than in the controls. Measurement of waist–hip ratio and measurement of waist circumference were comparable between the two groups whereas the mean BMI (kg/m²) was significantly higher in cases as compared to the controls in the association study cohort (25.97 ± 0.13 kg/m² versus 23.47 ± 0.13 kg/m², p = 0.008). TC and LDL-c levels were found to be lower in the cases than in the controls, which might be due to the usage of statins in the former group.

In the association study cohort, there were 464 cases of chronic stable angina (CSA), 61 with unstable angina (UA), and 482 cases with myocardial infarction (MI). Information was not available for 27 cases. Severity of CAD based on the number of diseased vessels showed the following distribution: 1-vessel disease: 186; 2-vessel disease: 278; 3– or more vessel disease: 501; data missing: 69. With regard to medical management of CAD, 169 patients underwent PCI, 692 patients underwent CABG, while 173 were only on standard drug therapy. Medication details are provided in Table 1 and primarily include statins, beta-blockers, angiotensin-converting enzyme inhibitors, calcium-channel blockers, nitrates, antiplatelets, fenofibates, and hypoglycemic agents.

In the gene expression study cohort, there were 51 cases of CSA and 46 cases of MI. Information was missing for 4 cases. Frequency distribution values based on the number of diseased vessels were given as follows: 1-vessel disease: 10; 2-vessel disease: 33; 3– or more vessel disease: 53; data not available: 5 cases. 88 patients underwent CABG; three had PCI while nine patients were treated with standard medication only. As there is a restricted access to the old patient records at Narayana Health due to the ongoing process of digitization of hospital records, we have been unable to retrieve medical information for all the participants. Further, as

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Clinical characteristics of study participants included in the association and expression study.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Association study cohort</td>
</tr>
<tr>
<td>Clinical factors</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>50.13 ± 0.261</td>
</tr>
<tr>
<td>Age at onset (years)</td>
<td>48.56 ± 0.740</td>
</tr>
<tr>
<td>Males, N (%)</td>
<td>801 (77.5%)</td>
</tr>
<tr>
<td>Females, N (%)</td>
<td>233 (22.5%)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>25.97 ± 0.13</td>
</tr>
<tr>
<td>Waist circumference (cm)</td>
<td>89.90 ± 0.327</td>
</tr>
<tr>
<td>Waist/hip ratio (cm²)</td>
<td>0.95 ± 0.003</td>
</tr>
<tr>
<td>Laboratory studies</td>
<td></td>
</tr>
<tr>
<td>TC (mg/dl)</td>
<td>152.42 ± 1.33</td>
</tr>
<tr>
<td>TG (mg/dl)</td>
<td>165.55 ± 2.41</td>
</tr>
<tr>
<td>HDL-c (mg/dl)</td>
<td>37.52 ± 0.27</td>
</tr>
<tr>
<td>LDL-c (mg/dl)</td>
<td>81.79 ± 1.146</td>
</tr>
<tr>
<td>Medical history</td>
<td></td>
</tr>
<tr>
<td>Smoking, N (%)</td>
<td>406 (39.3%)</td>
</tr>
<tr>
<td>Hypertension, N (%)</td>
<td>557 (54%)</td>
</tr>
<tr>
<td>Diabetes mellitus, N (%)</td>
<td>450 (43.6%)</td>
</tr>
<tr>
<td>Metabolic syndrome</td>
<td></td>
</tr>
<tr>
<td>MS-ATPIII</td>
<td>205 (19.4%)</td>
</tr>
<tr>
<td>MS-modified ATPIII</td>
<td>556 (53.8%)</td>
</tr>
<tr>
<td>Medications</td>
<td></td>
</tr>
<tr>
<td>Statin, N (%)</td>
<td>775 (75%)</td>
</tr>
<tr>
<td>Fenofibrates, N (%)</td>
<td>20 (1.93)</td>
</tr>
<tr>
<td>Beta blockers</td>
<td>660 (63.83)</td>
</tr>
<tr>
<td>ACE-inhibitors</td>
<td>329 (31.82)</td>
</tr>
<tr>
<td>Calcium-channel blockers</td>
<td>285 (27.56)</td>
</tr>
<tr>
<td>Nitrates</td>
<td>377 (36.46)</td>
</tr>
<tr>
<td>Antiplatelets</td>
<td>790 (76.40)</td>
</tr>
<tr>
<td>Hypoglycemic agents</td>
<td>201 (19.44)</td>
</tr>
</tbody>
</table>

BMI body mass index; FBS, fasting blood sugar; HDL-c, high-density lipoprotein cholesterol; LDL-c, low-density lipoprotein cholesterol; MS, metabolic syndrome; ATP, Adult Treatment Panel; TC, total cholesterol; TG, triglycerides.

a Continuous variables are expressed as mean ± standard error.

b Information on medication was missing for 1 case in the association study cohort and 1 case in the gene expression study cohort.
The AIP was estimated for all study participants. In the association study cohort, the mean AIP was significantly higher in the cases (0.614) than in the controls (0.557) ($p = 1.26 \times 10^{-6}$). Mean AIP was higher in the males ($n = 1572$) (0.608) than in the females ($n = 478$) (0.608 versus 0.515, $p = 4.18 \times 10^{-11}$). There was a significant association between three AIP groups and CAD ($p = 1.29 \times 10^{-12}$) and the percent frequencies in the cases and controls were as follows: low risk: 1.3% and 4.9%; intermediate risk: 1.6% and 6.5%; high risk: 97.0% and 88.6%. Logistic regression showed an OR of 3.988 (95% CI 2.011–7.909) for CAD when the 1st and 3rd groups were compared after adjusting for age, gender, diabetes, and hypertension. Significant higher mean AIP values were noted in the cases in spite of 73.4% of the cases being on statin medication while only 0.5% of the controls were on any lipid-lowering drug.

The overall trend was similar in the gene expression cohort; however the differences or association was not statistically significant possibly due to the small sample size.

In the association study cohort of 2068 subjects, there was no significant difference in the proportion of subjects with metabolic syndrome (MS) in the cases ($n = 205$, 19.8%) and the controls ($n = 221$, 21.4%) ($p = 0.415$) when standard ATP III guidelines were applied. However, the pattern changed with the application of the revised WC cut-offs where the cases had higher proportion of MS ($n = 556$, 53.8%) than among the controls ($n = 499$, 48.3%) ($p = 0.014$).

In the gene expression study cohort, MS information was available for 100 cases and 96 controls. While the standard ATP III classification did not show any significant difference in the prevalence of MS in the cases ($n = 33$, 32.7%) and the controls ($n = 39$, 40.6%) ($p = 0.300$), with the revised ATP III, MS prevalence was interestingly higher in the controls ($n = 55$, 57.3%) than in the cases ($n = 42$, 41.6%) ($p = 0.033$).

Association of SNPs with CAD

rs646776 and rs599839 SNPs in the CELSR2–PSRC1–SORT1 gene loci were in Hardy–Weinberg equilibrium in the control group ($p > 0.05$). rs599839 (OR = 0.315, 95% CI 0.136–0.728, $p = 0.007$) and rs646776 (OR = 0.216, 95% CI 0.087–0.535, $p = 0.001$) showed significant protective association with CAD and were in strong LD with each other ($r = 0.98$) (Fig. 1). The common ‘TA’ haplotype for the two variants, rs646776 and rs599839 showed a frequency of 0.73. The alternate ‘GC’ haplotype (0.27) showed protection against CAD (OR 0.77, 95% CI 0.67–0.88, $p = 0.0002$).

Association of SNPs with plasma lipid levels

We examined the effect of rs646776 and rs599839 on plasma lipid levels. Both variants showed a significant association with TC ($p = 0.007$) and LDL-c ($p = 0.01$), with the minor alleles being associated with lower lipid levels. This association remained significant even after adjusting for age, gender, diabetes, hypertension, lipids, BMI, and smoking ($p < 0.002$). Homozygotes for the minor allele (CC or GG) showed an overall reduction of approximately 11 mg/dl (0.29 mmol/l) in TC levels and around 8 mg/dl (0.21 mmol/l) in LDL-c levels (Table 2).

Association between genotype and CELSR2–PSRC1–SORT1 gene expression

As compared to the common TT genotype in rs646776 or AA genotype in rs599839, the homozygous variant genotype (CC or GG) showed 30% increase in PSRC1 mRNA expression while the heterozygote genotypes showed 15% higher expression (Fig. 2). There was however no significant change in the expression levels of CELSR2 and SORT1 genes across all the three genotypes.

Association of CELSR2–PSRC1–SORT1 gene expression with CAD and lipids

Next, we looked at the association of CELSR2–PSRC1–SORT1 gene expression in CAD. Expression levels of PSRC1 transcript were significantly lower in the cases than in the controls (0.75 ± 0.405 versus 1.04 ± 0.622, $p = 2.255 \times 10^{-4}$). On the other hand, the expression of CELSR2 and SORT1 genes did not show any significant difference between the cases and the controls. Further, the PSRC1 gene showed significant protective association with CAD (OR = 0.216, 95% CI 0.087–0.535, $p = 0.001$). We observed strong correlation of PSRC1 mRNA expression with SORT1 ($r = 0.69$) and CELSR2 ($r = 0.46$). PSRC1 expression also showed modest correlation with TC ($r = 0.22$, $p = 0.003$), LDL-c ($r = 0.17$, $p = 0.024$), and HDL-c ($r = 0.16$, $p = 0.032$) levels while CELSR2 showed correlation with TC ($r = 0.22$, $p = 0.003$), LDL-c ($r = 0.18$, $p = 0.011$), and TG levels ($r = 0.18$, $p = 0.013$).

Table 2

<table>
<thead>
<tr>
<th></th>
<th>CC</th>
<th>TC</th>
<th>TT</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TC (mg/dl)</td>
<td>155.68</td>
<td>163.35</td>
<td>166.60</td>
<td>0.007</td>
</tr>
<tr>
<td>TG</td>
<td>150.63</td>
<td>160.03</td>
<td>166.50</td>
<td>0.131</td>
</tr>
<tr>
<td>HDL-c</td>
<td>38.52</td>
<td>39.41</td>
<td>38.6</td>
<td>0.172</td>
</tr>
<tr>
<td>LDL-c</td>
<td>87.38</td>
<td>90.99</td>
<td>95.21</td>
<td>0.006</td>
</tr>
</tbody>
</table>

rs599839

<table>
<thead>
<tr>
<th></th>
<th>AA</th>
<th>AG</th>
<th>GG</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TC (mg/dl)</td>
<td>165.75</td>
<td>163.15</td>
<td>156.13</td>
<td>0.007</td>
</tr>
<tr>
<td>TG</td>
<td>166.91</td>
<td>168.74</td>
<td>149.94</td>
<td>0.106</td>
</tr>
<tr>
<td>HDL-c</td>
<td>38.67</td>
<td>39.27</td>
<td>38.74</td>
<td>0.387</td>
</tr>
<tr>
<td>LDL-c</td>
<td>95.09</td>
<td>91.15</td>
<td>87.74</td>
<td>0.01</td>
</tr>
</tbody>
</table>

TC, total cholesterol; TG, triglycerides; HDL-c, high-density lipoprotein cholesterol; LDL-c, low-density lipoprotein cholesterol.

![Fig. 1. Linkage disequilibrium plot for rs646776 and rs599839 SNPs in the 1p13.3 gene loci. Haploview linkage disequilibrium plot with D' value of 98 between the two SNPs.](image.png)
Fig. 2. Mean mRNA expression levels of CELSR2, PSRC1 and SORT1 genes across rs646776 and rs599839 genotypes. mRNA expression level of CELSR2, PSRC1 and SORT1 in the presence of homozygote (major and minor) and heterozygote alleles for rs646776 and rs599839 SNPs. Note: for rs646776, TT: major allele, CC: minor allele; for rs599839, AA: major allele, GG: minor allele.

Network analysis

A bionetwork was generated using STRING to understand the relationship of the PSRC1 gene with other interactors where PSRC1 was used as the seed gene. This gene showed interactions with various partners including genes involved in lipoprotein uptake and transport (SORT1 and SORL1), mitosis (CDC3 and CDC8), and Golgi transport (KIF20A) while PSRC1 itself is responsible for microtubule destabilization and spindle assembly (Fig. 3).

Discussion

Based on the literature search from GWAS, we identified two SNPs, rs599839 and rs646776 present in the CELSR2–PSRC1–SORT1 cholesterol gene cluster on chromosome 1p13.3 and examined their effect on gene expression, association with CAD and circulating lipid levels in a representative cohort of Asian Indians. We observed a strong independent association of both the variants with the disease, lower expression of PSRC1 gene transcript in the cases and the minor alleles of both SNPs being associated with lower level of circulating lipids, particularly TC and LDL-c. These two variants that were initially identified through GWAS in the Caucasian population showed an association with CAD and the lipid phenotype [15,30–32] and have been subsequently validated in other populations [33] including the Sikh Diabetes Study cohort [34]. Thus, the most plausible conclusion derived from these reported studies including the present study is that the CELSR2–PSRC1–SORT1 locus exerts its effect on CAD through a well-established risk factor, namely plasma LDL-c.

It is interesting to note that of the three genes in the above-mentioned cluster, SORT1 has been reported to be directly involved in lipid regulatory metabolism [14,35,36] while the primary role of CELSR2 and PSRC1 is cell adhesion and cytoskeleton stabilization, respectively [13,37]. To the contrary, we observed a significant association of PSRC1 mRNA expression with CAD as well as the association of the two SNPs with TC and LDL-c levels in the present study. This was comparable to a report by Wild et al. [38] who showed a significant association of rs599839 SNP with expression of PSRC1 transcripts in the monocytes and plasma LDL-c levels. A number of studies have also shown an association of the variant alleles of both rs599839 and rs646776 with low plasma LDL-c levels [10,39,40]. In our preliminary analysis of the protein interaction network generated using STRING v9.05, with PSRC1 as the seed gene, we observed that the PSRC1 appeared to be a key protein linking CELSR2 and SORT1 on one side and factors affecting mitosis and cell division on the other side (Fig. 3). Further, the link of PSRC1

Fig. 3. Network analysis with PSRC1 gene as the seed gene. Network constructed using STRING, a web-based tool with PSRC1 as the protein of interest shows interactions (edges) of PSRC1 with other proteins.
with SORL1 is interesting in that the latter encodes for a sortilin-related receptor protein, which is located on chromosome 11q23, yet another important locus regulating plasma lipids [34]. From this it appears that the relationship among the candidate genes in the 1p13.3 locus is complex and remains to be fully elucidated.

The fact that lipid is an important factor for the development of atherosclerosis is understood from the fact that there was a strong association of atherogenic index and CAD. Further, MS is considered to be a precursor for CAD. Asian Indians have a greater predisposition to MS and the need for ethnicity-specific cut-offs has been extensively discussed [25,41,42]. This was evident from the fact that there was high MS prevalence both among the cases and controls and the frequency differences in the cases and controls were statistically significant when the revised cut-off was used for waist circumference. Thus both lipid and visceral adiposity appear to be significant factors contributing to the development of metabolic diseases and CAD in this population.

Both rs646776 and rs599839 are located in an intergenic region between CELSR2 and PSRC1 genes. SNP rs599839 was identified first to be located in the 3′-untranslated region (UTR) of PSRC1, and showed association with CAD in the British and German cohort study [43]. Subsequently, a meta-analysis carried out on the British population demonstrated that the rs646776 and rs599839 SNPs were significantly associated with lower levels of circulating LDL-c [16]. In the present study, which includes a representative cohort of Asian Indians living in the Indian sub-continent, we observed a similar pattern, wherein the minor alleles of both the SNPs were associated with reduced LDL-c levels. This demonstrates a protective role of both the variant alleles. We also noted higher expression of PSRC1 mRNA in the carriers of the minor alleles, C and G. Contrary to this, in an earlier study, higher mRNA expression levels of SORT1 were shown to induce a significant cellular uptake of LDL-c [10] and the homozygous minor allele of rs646776 SNP was associated with 12-fold higher SORT1 and PSRC1 expression in human liver samples [44]. In fact, this was the first study on a series of human cohorts and human-derived hepatocytes to provide functional evidence on the association of CELSR2–PSRC1–SORT1 locus with CAD.

A non-coding SNP located in the 3′-UTR of angiotensin II type 1 receptor gene has been shown to affect the miRNA-mediated regulation of the mRNA [45], while in the case of human LDL receptor gene, a SNP present in the 3′-UTR was associated with plasma lipid levels in Caucasian males [46]. This provides an insight into the post-transcriptional regulation brought about by these non-coding SNPs. rs599839 is also a non-coding SNP located in the 3′-UTR of the PSRC1 gene which could possibly play a regulatory role at the mRNA level. Given that the increased expression level of PSRC1 mRNA transcripts and the reduced levels of LDL-c and TC were detected in the presence of minor alleles of both rs599839 and rs646776 SNPs, it might indicate a possible role of the PSRC1 gene in the direct or indirect regulation of plasma lipid levels. The PSRC1 and SORT1 genes lay in close proximity in the 1p13.3 genomic locus, which was reflected by the presence of strong correlation in the gene expression between the microtubule destabilizing PSRC1 and plasma lipid regulating SORT1 genes in our study. Furthermore, in the light of rs646776 and rs599839 SNPs being present as a tight intergenic cluster, it would be interesting to understand the underlying mechanisms, the functional relationship, and the manner in which the PSRC1 and SORT1 genes might complement each other in modulating CAD risk through their combined effect on cholesterol metabolism.

A number of GWAS studies have confirmed the consistent yet modest association of the 9p21.3 common variants with incident CAD. There is however an urgent need to identify additional such variants that can improve risk prediction. Given the published evidence on the role of these SNPs in modulating PSRC1 expression, their protective association with CAD and circulating cholesterol, it is a matter of time before rs599839 and rs646776 can be considered as novel variants with potential clinical utility. However, much work needs to be done to understand the underlying mechanism by which these SNPs actually contribute to the clinical manifestation of the disease.

We acknowledge some of the limitations of this study. To begin with, the controls were selected from among individuals who were clinically asymptomatic for the disease and showed normal ECG readings. This does not exclude the presence of an underlying silent disease. Nonetheless, as the only fool-proof method to rule out the presence/absence of disease is through a coronary angiogram, which is not ethically feasible in an experimental setting, this could to some extent have an impact on the significance of our study findings. Importantly, our initial findings on the association of the two SNPs with CAD have not been validated in an independent cohort, considering that the contribution of genetic factors to CAD risk is rather small and is primarily based on differences in the risk allele frequency in cases versus controls. In this context it is worthwhile to recall that this locus retained genome-wide significance (>10−8) even after stringent Bonferroni correction in the Caucasian [47] and other Asian [48] populations demonstrating the robustness of their association with CAD, primarily mediated through the lipids [8]. Further, only one locus, namely CELSR2–PSRC1–SORT1, was considered for this study while a recent study has identified around 157 loci associated with CAD, primarily mediated through the lipids [8].

In conclusion, this pilot study reiterates the involvement of the CELSR2–PSRC1–SORT1 locus in modulating the risk of CAD through hitherto unknown mechanisms that regulate plasma lipid levels. While the association of the two SNPs, rs646776 and rs599839, with CAD and lipids corroborate with other published studies, PSRC1 and SORT1 appear to be potential gene candidates in the lipid pathway. However, until the functional relationship among the members of this cholesterol locus and other reported lipid genes is established, the true clinical utility of the CELSR2–PSRC1–SORT1 locus will not be fully realized.

Conflict of interest

The authors have no conflicts of interest to disclose.

Acknowledgments

We gratefully acknowledge the support extended by the trustees of the Thrombosis Research Institutes in London and Bangalore, the W. Garfield Weston Foundation, UK, Foundation Bay, Switzerland, the Tata Social Welfare Trust, India (TSWT/IG/SNB/JP/Pdm) and the Department of Biotechnology, Ministry of Science and Technology, Government of India (BT/01/CDE/08/07). We thank the study participants and the doctors; without whom this work would not have been possible. The sponsors did not participate in the design, conduct, sample collection analysis, and interpretation of the data or in the preparation, review, or approval of the manuscript.
References


[34] Arvind P. Arvind et al. / Journal of Cardiology 64 (2014) 339–346


[38] Arvind P. Arvind et al. / Journal of Cardiology 64 (2014) 339–346


