Genetic analysis of the 9p21.3 CAD risk locus in Asian Indians

Jayashree Shanker1*; Prathima Arvind1*; Srikarthika Jambunathan1; Jiny Nair2; Vijay V. Kakkar2,3

1Mary and Garry Weston Functional Genomics Unit, Thrombosis Research Institute, Bangalore, India; 2Chairman and Managing Trustee, Thrombosis Research Institute, Bangalore, India; 3Thrombosis Research Institute, London, UK

Summary

The 9p21.3 locus is the best replicated region to date for coronary artery disease (CAD). We investigated the association of 9p21.3 common variants with CAD, candidate gene expression including ANRIL, a non-coding RNA, followed by in vitro validation. Five variants, rs10757278, rs10757274, rs2383206, rs1333049 and rs4977574 were genotyped in 1,034 cases and 1,034 controls. Gene expression of C9orf5, MTAP 1, MTAP 2, p16INK4a, p14ARF, p15INK4b and two ANRIL splice variants, NR_003529 and EU741058, were measured in 100 cases and 100 controls. Human aortic smooth muscle cells (HuAoSMCs) were transfected with siRNA targeting ANRIL exon 19 (siRNA-1) or exon 2 (siRNA-2) and consequent effect determined. rs2383206 showed the highest association with CAD (odds ratio [OR] 2.02, 95% confidence interval [CI] 1.56–2.62) and an adjusted OR of 2.55, 1.33–2.88 along with rs10757278. Conventional risk factors (conventional RFs), rs2383206 and rs10757278 variants together yielded a higher c index (OR 0.790, 95% CI 0.770–0.810) as compared to conventional RFs (OR 0.783, 95% CI 0.763–0.803) or genetic variants (OR 0.561, 95% CI 0.536–0.586) alone. GAAAA haplotype showed significant protective association with CAD compared to CGGGG risk haplotype (OR 0.45, 95% CI 0.27–0.77). Expression of p16INK4a, p14ARF and p15INK4b as well as plasma CDKN2A levels were lower in cases than controls. GG genotype was associated with higher EU741058 expression and lower p16INK4a expression. HuAoSMCs transfected with siRNA-1 showed lower NR_003529, p16INK4a and p14ARF expression. Our study provides further evidence on the significance of 9p21.3 locus for CAD wherein the risk allele regulate the expression of ANRIL and adjacent tumour suppressor genes which in turn alter smooth muscle proliferation, a fundamental process in atherosclerosis.

Keywords

Coronary artery disease, 9p21.3 locus, association study, gene expression, Asian Indians

Introduction

The 9p21.3 region, termed as the ‘9p21.3 CAD risk interval’, encompasses multiple genetic variants in tight linkage disequilibrium, spanning a 58-kb region, and considered to be the most significant locus for coronary artery disease (CAD) (1). The unequivocal association of 9p21.3 common variants with CAD was discovered independently through genome wide association studies (2–4) and subsequently validated in different ethnic populations (5–11). The 9p21 variant is the only consistent genetic marker to date carrying an allelic odds ratio (OR) of 1.2–1.3 for CAD risk (12), and a population attributable risk of 20% for myocardial infarction (13), which is independent of the well-established risk factors including family history (14). In vitro experimental studies (15, 16) and in vivo studies on mice models (17) have shown that although the 9p21.3 locus per se is bereft of any protein coding genes, the variants are located within a 58-kb region designated as ANRIL (antisense non-coding ribonucleic acid of INK locus) and exerts its effect on the neighboring tumour suppressor genes, CDKN2A (p16INK4a / p14ARF) and CDKN2B (p15INK4b) through epigenetic mechanisms (18). The ANRIL appears to regulate cell cycle and proliferation process through suppression of CDKN2A and CDKN2B gene activity (15). Again, our knowledge on the functional relationship between ANRIL and the adjacent candidate genes has been attained through in vitro studies on vascular smooth muscle cells, where selective blocking of specific exons of ANRIL affect CDKN2A/2B expression (15). Considering that proliferation of smooth muscle cells is a fundamental process in atherosclerosis development (19), the full potential of ANRIL’s role in CAD pathogenesis is yet to be realised. The 9p21.3 risk allele is also associated with various other vascular phenotypes such as stroke, peripheral artery disease, aortic abdominal and intracranial aneurysm etc. (20), implying that it may play a fundamental role in vascular physiology. In fact, it is said to act as a vas-
cular growth regulator element, exhibiting dose effect between the number of risk alleles and disease severity (21).

There are limited studies on the 9p21.3 locus in Asian Indians (7, 22, 23) where the primary focus has been on genotype/haplo-type association of 9p21.3 common variants with CAD. The first ever report published on Asian Indians showed a high frequency of the G allele (0.56) in rs10757278 variant (0.56) and was associated with >2-fold risk for CAD (OR 2.15, 95% confidence interval [CI] 1.02 – 4.60) and a high population attributable risk of up to 46% in men (7). The risk attribute of 9p21.3 common allele is said to be modulated by the flanking SNPs in this region (22). In the present study, we have investigated the 9p21.3 locus with respect to the association of five common genetic variants with CAD, compared the relative expression of ANRIL splice variants and other candidate genes in the 9p21.3 locus in a representative case-control cohort of Asian Indians, and finally observing the consequent effect of using short interfering RNA (siRNA) targeted against specific exons in ANRIL, on the neighbouring candidate genes. Using such a multi-pronged approach, we have been able to demonstrate a close association between ANRIL and its neighbouring genes which may collectively contribute to the development of CAD.

Material and methods

Ethics statement

The IARS protocol has been approved by the ethics committee of the Thrombosis Research Institute, and all clinical investigations were conducted according to the guidelines of the Indian Council of Medical Research on bioethics (24) and as per the principles expressed in the Declaration of Helsinki.

Study population

The study cohort was selected from the Indian Atherosclerosis Research Study (IARS) which is a large ongoing epidemiological study established in January 2003 with the objective of understanding the contribution of traditional and emerging risk factors of premature CAD in Asian Indians living in their home country. Overview of the IARS design has been previously published (25).

Recruitment of cases and controls was based on predefined inclusion/exclusion criteria. Age at onset was ≤60 years for males and ≤65 years for females. CAD patients (cases) showed clinical evidence of stable angina, unstable angina or myocardial infarction, diagnosed by coronary angiography and electrocardiogram (ECG) and treated with standard medication, coronary intervention or bypass surgery. Control subjects were without family history of cardiovascular disease, clinically asymptomatic for coronary heart disease as demonstrated by normal ECG readings and were enrolled from the community. All participants provided voluntary informed signed consent. A cohort of 1,034 cases and 1,034 age- and gender-matched controls were selected from among the IARS cohort for the genetic association study. An independent cohort of 100 cases with positive family history and 100 healthy controls without family history of cardiovascular disease were included for the candidate gene expression studies.

Sample preparation

Overnight fasting blood samples were collected from all study participants. Aliquots of serum and plasma were stored at –80°C and used for biochemical analysis. Genomic DNA was extracted by a modified salting out procedure (26) and quantitated using NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). Total RNA was extracted from 3 ml of fresh EDTA whole blood sample using QIamp RNA blood mini kit (Qiagen, Valencia, CA, USA) following manufacturer’s instructions and total RNA was quantified by NanoDrop ND-1000 Spectrophotometer (NanoDrop, Wilmington, DE, USA). RNA samples were further digested with DNase I. First strand synthesis of cDNA was performed with cDNA archive kit (Applied Biosystems, Foster City, CA, USA).

Laboratory assays

Lipid markers namely serum total cholesterol (TC) and triglycerides (TG) (Randox Laboratories, Antrim, UK) were estimated by standard enzymatic analysis on Cobas-Fara II Clinical Chemistry Auto analyzer (F. Hoffman La Roche Ltd, Basel, Switzerland). High-density lipoprotein-cholesterol (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, Gujarat, India). Plasma Low Density Lipoprotein-cholesterol (LDL-c) was calculated using Friedewald’s formula (27) The inter-assay co-efficient of variation for the commercial controls and normal serum pool ranged from 4.9% to 7.0% for total cholesterol, 6.1% to 7.7% for triglycerides and 7.1% to 12.2% for HDL-cholesterol.

Genotyping of 9p21.3 variants

A total of five genetic variants in the 9p21.3 CAD risk interval namely rs10757278, rs10757274, rs2383206, rs1333049 and rs4977574, were selected based on published evidence of significant association with CAD. Relative positions of the SNPs on ANRIL are shown in Suppl. Figure 1A (available online at www.thrombosis-online.com). These variants were genotyped by TaqMan allelic discrimination assay on a 7900HT Fast Real Time PCR instrument (Applied Biosystems). Positive in-house controls containing samples with known genotypes along with ‘no template’ control were run with every assay batch.

Gene expression assay

The following candidate genes and their splice variants in the 9p21.3 region were selected for the study – C9orf53 (C9 open reading frame 53), MTAP (Methylthioadenosine phosphorylases, transcript1 (T1) and transcript 2 (T2)), Cyclin-dependent kinase inhibitors, CDKN2A (p16INK4a, p14ARF), CDKN2B (p15INK4b), full-
length ANRIL transcript, _NR_003529_ and a short splice variant, _EU741058_. Suppl. Figure 1A shows the relative locations of candidate genes in the 9p21 locus while Suppl. Figure 1B shows a diagrammatic representation of ANRIL transcripts with corresponding PCR primer positions (both available online at www.thrombosis-online.com). Relative gene expression was measured by Taqman real-time assay. FAM and TAMRA dye labelled probe-primer mix (20X) and Universal Master Mix (2X) were purchased from ABI (Applied Biosystems). β-actin was used as endogenous control and run with every sample. All assays were set up in duplicates and performed as per manufacturer recommended protocol. Data was processed using sequence detection software version 2.3 (Applied Biosystems). Relative changes in mRNA expression levels between cases and controls were calculated with RQ manager. Assays were repeated in duplicates for those samples showing skewed expression levels. Persistent outliers were excluded from further analysis using the Outlier function in SPSS v 17.0 software.

**Measurement of CDKN2A and CDKN2B protein levels**

Protein levels of CDKN2A and CDKN2B were measured in EDTA plasma using standard ELISA kits (CUSABIO, Wuhan, China) following the manufacturer’s recommended protocol. We measured plasma CDKN2A levels in 39 cases and 43 controls and CDKN2B levels in 41 cases and 43 controls, respectively.

**In vitro functional studies**

Primary human aortic smooth muscle cells (HuAoSMCs) were obtained from GIBCO (Life Technologies, Carlsbad, CA, USA) and cultured in Dulbecco’s Minimal Essential Media (Lonza, Walkersville, MD, USA), supplemented with 10% Fetal Bovine Serum (GIBCO), 1% Penicillin-Streptomycin (GIBCO) and Smooth Muscle Growth Supplement (GIBCO). For all cell based assays, HuAoSMCs lesser than seven passages were used. Silencer® FAM-labelled negative control #1 (Ambion, Foster City, CA, USA) was used as control siRNA in all transfection experiments. Custom Silencer® siRNA-1 (targeting ANRIL exon 19) was the same as used by Congrains et al. (15) (GAACCAG GACTGGAACATT). For Custom Silencer® siRNA-2, a pool of three siRNAs targeting exon 2 were used –5’GACTATTTTGCCACGACATTTCAAAAGGATTTCAAGAGAAGATATTGCTG3’; 5’CCATGCTGTTAGATTCTCAGCTCTCTCATCTGATCTCGTCTGGCCCATGACTT3’; 5’TCTTTGTGGTAGTTAGGCTGTGTTATGCTGAGCCACTGACACCAAATTGTG3’. For the ANRIL silencing assay, 10⁵ cells were transfected with Lipofectamine RNAmax (Life Technologies) as per manufacturer recommended protocol. Briefly, 50 nM of siRNA was transfected with 9 µl of Lipofectamine RNAmax (Life Technologies), and cells were harvested after 48 hours (h). There were duplicate assays set up for each treatment with a total of three rounds of experiment. Total RNA was extracted from harvested cells using TRIzol® method. Briefly, 1 ml of TRIzol was directly added to the cells after removing the growth medium, contents transferred to an Eppendorf tube and incubated for 5 minutes (min) at room temperature. A volume of 200 µl chloroform was then added, shaken vigorously for 15 seconds (sec) and incubated for 2–3 min at room temperature. The mixture was centrifuged at 12,000 X g for 15 sec at 4°C, aqueous phase was collected into a fresh tube, 500 µl of isopropanol was added and further incubated at room temperature for 10 min. Following centrifugation for 10 min at 4°C, the pellet was washed with 75% ethanol and centrifuged at 7,500 X g for 5 min at 4°C. Air dried pellet was finally resuspended in 20 µl of RNase free water. RNA was quantified on NanoDrop ND-1000 Spectrophotometer (NanoDrop). Relative gene expression was measured by Taqman real-time PCR (Applied Biosystems).

**Cell proliferation assay**

HuAoSMCs were grown in 96-well plate and the extent of cellular proliferation was measured with Presto Blue Cell Viability reagent (Life Technologies). Initially, HuAoSMCs (2,500 cells/well) were transfected with siRNA as mentioned above and 72 h post-transfection, cell proliferation assay was done as per manufacturer’s protocol. Presto Blue Cell Viability reagent was added to each well, incubated for 2 h at 37°C and absorbance was read at 570 nm and 600 nm using a microplate reader (BioTek Power wave X, BioTek, Winooski, VT, USA).

**Statistical methods**

Routine statistical analysis was performed using SPSS v17.0 software (SPSS Inc, Chicago, IL, USA). Continuous variables were expressed as mean ± standard error of the mean (SEM). Chi-square analysis and binary logistic regression was used for testing the association of 9p21.3 variants with CAD and for estimating the ORs and 95% CIs, while Student’s t-test and univariate analysis were used to test for mean differences in levels of gene expression and other quantitative traits between cases and controls. Age, gender, diabetes, hypertension, total cholesterol and HDL-C were treated as covariates and appropriately adjusted for during analysis. SNPstats online software was also used to calculate allele and genotype frequencies, Hardy-Weinberg equilibrium and genotype/haplotype associations with CAD (28). Estimation of linkage disequilibrium (LD) was performed using Haplovie v3.32 software (29). Measure of discriminative accuracy was assessed based on area under the receiver-operating characteristic (ROC) curve (AUC or c index) for three different risk prediction models: model 1, conventional risk factors (RFs) alone; model 2, 9p21.3 variants alone; and model 3, conventional RFs and genetic variants.

**Results**

**Clinical characteristics of study participants**

Clinical profile of participants for the association study and gene expression study is shown in Table 1. Average age of cases and controls was 50 years in both the study groups. Males were better represented (76% – 90%) than females. Classical risk factors such
as diabetes and hypertension were highly prevalent among the cases. Lipids, particularly total cholesterol and low-density lipoprotein (LDL)-cholesterol, were lower among the cases than in the controls, which can be attributed to the high usage of statins in this group. Average age at CAD onset was less than 50 years.

Association of 9p21 variants and haplotypes with CAD

All five variants in the 9p21 region were in Hardy-Weinberg equilibrium in the control group (p > 0.05). Allele and genotype frequencies and corresponding ORs are shown in Table 2. Frequency of the common risk allele varied between 0.54–0.57. All five variants showed significant independent association with CAD. rs2383206 showed the highest individual association with CAD (OR 2.02, 95% CI 1.56–2.62, p < 0.0001) and in combination with rs10757278, yielded an OR of 2.64 (95% CI 2.61–2.67) (p = 0.004) after adjusting for classical risk factors. In all five 9p21 variants were in strong LD (r = 0.90–0.98) (see Suppl. Figure 2, available online at www.thrombosis-online.com) and were represented as two distinct blocks. Block 1 included three variants (rs10757274, rs4977574, rs2383206) and block 2 comprised of two variants (rs10757278, rs1333049). Interestingly, rs2383206 exhibited lesser correlation with rs10757278 (r = 0.91) as compared to the other variants (r = 0.93 to 0.97).

Expression of candidate genes and their splice variants in 9p21.3 locus

Candidate gene expression data was obtained for 100 cases and 100 controls of which there were four data outliers in the control group that were removed prior to analysis. We observed ~1.5-fold lower expression of p14ARF (0.934 ± 0.058 vs 1.381), p16INK4a (0.840 ± 0.058 vs 1.232 ± 0.065) and p16INK4a (0.761 ± 0.080 vs 1.351 ± 0.128) (p < 0.0001) genes in cases as compared to controls (Figure 1). Further, C9orf53 gene also showed nearly 1.6 fold lower expression in the CAD samples (p < 0.001) while the ANRIL transcripts as well as MTAP T1 and MTAP T2 did not exhibit any significant difference. No significant differential gene expression was seen in an independent analysis of myocardial infarction and stable angina samples.

Analysis of correlation among the various genes and splice variants showed an interesting trend. NR_003529 expression corre-

<table>
<thead>
<tr>
<th>Clinical factors</th>
<th>Genetic association cohort</th>
<th>Gene expression cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases (n=1,034)</td>
<td>Control (n=1,034)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>50.03 ± 0.252</td>
<td>50.17 ± 0.261</td>
</tr>
<tr>
<td>Age at CAD onset (years)</td>
<td>47.81 ± 0.246</td>
<td>-</td>
</tr>
<tr>
<td>Males N (%)</td>
<td>801 (77.5)</td>
<td>785 (75.9)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>25.96 ± 0.127</td>
<td>25.46 ± 0.137</td>
</tr>
<tr>
<td>Waist/Hip Ratio (cm)</td>
<td>0.94 ± 0.002</td>
<td>0.953 ± 0.002</td>
</tr>
<tr>
<td>TC (mg/dl)</td>
<td>152.42 ± 1.34</td>
<td>176.71 ± 1.21</td>
</tr>
<tr>
<td>TG (mg/dl)</td>
<td>155.55 ± 2.41</td>
<td>167.04 ± 3.88</td>
</tr>
<tr>
<td>HDL-c (mg/dl)</td>
<td>37.52 ± 0.27</td>
<td>40.30 ± 0.32</td>
</tr>
<tr>
<td>LDL-c (mg/dl)</td>
<td>81.79 ± 1.15</td>
<td>104.62 ± 1.012</td>
</tr>
<tr>
<td>Smoking N (%)</td>
<td>406 (39.3)</td>
<td>232 (22.5)</td>
</tr>
<tr>
<td>Hypertension N (%)</td>
<td>557 (54)</td>
<td>156 (15.3)</td>
</tr>
<tr>
<td>Diabetes mellitus N (%)</td>
<td>450 (43.6)</td>
<td>164 (16.1)</td>
</tr>
<tr>
<td>Statin N (%)</td>
<td>414 (73.4)</td>
<td>2 (0.5)</td>
</tr>
</tbody>
</table>

Continuous variables are expressed as mean ± standard error. BMI, body mass index; HDL-c, high-density-lipoprotein cholesterol; LDL-c, low-density-lipoprotein cholesterol; TC, total cholesterol; TG, triglycerides.
lated positively with EU741058 (r=0.25) and p16\(^{INK4a}\) (r=0.31). EU741058 showed significant correlation with p16\(^{INK4a}\) and p14\(^{ARF}\) genes (r=0.22, p=0.003) while the correlation with p15\(^{INK4b}\) was marginal (r=0.14, p=0.06). There was strong correlation among p16\(^{INK4}\), p14\(^{ARF}\), p15\(^{INK4b}\) genes (r=0.50–0.60) (p<0.0001) after adjustment for age, gender and statins.

### Association between candidate gene expression and CAD

There was significant association of p15\(^{INK4b}\) (OR 0.348, 95% CI 0.197–0.613, p<0.0001) and C9orf53 (OR 0.493, 95% CI 0.293–0.830, p=0.008) with CAD. However, only p15\(^{INK4b}\) retained statistical significance following adjustment for confounders. In a comparison of the top vs bottom quartile of gene expression, both p15\(^{INK4b}\) (OR 0.16, 95% CI 0.034–0.71) (p=0.016) and p16\(^{INK4a}\) (OR 0.16, 95% CI 0.042–0.62) (p=0.008) showed significant protection against CAD, even after adjusting for the conventional risk factors. However, ANRIL splice variants, MTAP T1 and T2 transcripts and C9orf53 transcripts did not show any significant association.

### Association between 9p21.3 genotypes and candidate gene expression

Subjects carrying the homozygous GG risk genotype (0.862 ± 0.060) or heterozygous AG genotype (0.674 ± 0.047) showed higher expression of EU741058 as compared to AA homozygote (0.507 ± 0.044) (p=0.001) (Figure 2A). On the other hand, GG

---

**Table 2: Allele and genotype frequency of the five 9p21 common variants.**

<table>
<thead>
<tr>
<th>SNP ID</th>
<th>Gene name</th>
<th>SNP*</th>
<th>Allele frequency (%)</th>
<th>Genotype frequency (%)</th>
<th>OR (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1330049</td>
<td>ANRIL</td>
<td>C&gt;G</td>
<td>0.57</td>
<td>0.43</td>
<td>1.60</td>
<td>4.03 x 10(^{-5})</td>
</tr>
<tr>
<td>rs2383206</td>
<td>ANRIL</td>
<td>G&gt;A</td>
<td>0.56</td>
<td>0.44</td>
<td>2.02</td>
<td>1.01 x 10(^{-7})</td>
</tr>
<tr>
<td>rs10757278</td>
<td>ANRIL</td>
<td>G&gt;A</td>
<td>0.55</td>
<td>0.45</td>
<td>1.73</td>
<td>2.75 x 10(^{-5})</td>
</tr>
<tr>
<td>rs10757274</td>
<td>ANRIL</td>
<td>G&gt;A</td>
<td>0.54</td>
<td>0.46</td>
<td>1.85</td>
<td>3.36 x 10(^{-6})</td>
</tr>
<tr>
<td>rs4977574</td>
<td>CDKN2A/2B, ANRIL</td>
<td>G&gt;A</td>
<td>0.54</td>
<td>0.46</td>
<td>1.82</td>
<td>3.36 x 10(^{-6})</td>
</tr>
</tbody>
</table>

*All SNPs were in HWE in the control data set. allele 1 – major allele, allele 2 – minor allele, genotype 1,1– homozygous normal genotype, genotype 1,2 – heterogeneous genotype, genotype 2,2 – homozygous variant genotype; OR- odds ratio, CI-confidence interval.

---

**Figure 1:** Mean blood expression levels of candidate genes and ANRIL transcripts in the 9p21.3 locus in cases and controls.
genotype was associated with lower expression of tumour suppressor gene $p16^{INK4a}$ as compared with AA genotype ($GG=1.059 \pm 0.139$, $AG=0.902 \pm 0.082$, $AA=1.528 \pm 0.279$; $p<0.01$) (Figure 2B). This pattern was consistent across all the five 9p21.3 variants.

Association of plasma levels of CDKN2A and CDKN2B with CAD

After eliminating outliers using Q-Q plot, we had data on plasma levels of CDKN2A for 39 cases and 42 controls and CDKN2B levels for 35 cases and 37 controls, respectively. Univariate analysis showed significantly lower CDKN2A levels in cases as compared to controls (11.239 ± 0.934 vs 15.085 ± 0.877, $p=0.004$), after adjusting for age and gender (Figure 3). On the other hand, CDKN2B did not show significant differences between cases and controls (32.291 ± 2.203 vs 30.012 ± 2.111, $p=0.460$).

In vitro functional studies on ANRIL and neighbouring genes in the 9p21.3 locus

To study the effect of selective ANRIL silencing on genes in the 9p21.3 locus, we used HuAoSMCs, which were transfected with siRNA specific to exon 19 (siRNA-1) and exon 2 (siRNA-2). Expression of NR_003529 and EU741058 were analysed and compared with un-transfected controls. NR_003529 showed mean lower expression in siRNA-1 treated cells as compared to controls (0.660 ± 0.076 vs 1.192 ± 0.293), although the difference was not significant ($p=0.12$) (Figure 4A). Further, there was lower expression of $p16^{INK4a}$ (0.707 ± 0.060 vs 1.046 ± 0.097) ($p=0.014$) and $p14^{ARF}$ (0.749 ± 0.054 vs 1.06 ± 0.101) ($p=0.022$) in cases relative to controls. As expected, EU741058 lacking exons 17–18 did not show significant difference between the treated and untreated cultures (0.910 ± 0.141 vs 1.074 ± 0.093). In contrast, in cultures treated with siRNA-2, there was no significant difference in expression for NR_003529 between treated and untreated control cultures (1.055 ± 0.169 vs 1.192 ± 0.293). Interestingly, both $p16^{INK4a}$ (0.707 ± 0.065 vs 1.046 ± 0.097) ($p=0.015$) and $p14^{ARF}$ (0.777 ± 0.079 vs 1.064 ± 0.101) showed lower expression in siRNA2 treated cultures as compared to untreated controls (Figure 4B). This pattern was comparable to siRNA-1 treatment. Both siRNA-1 and siRNA-2 did not show any significant effect on expression of neighbouring genes, MTAP, and C9orf53 expression.

Cellular proliferation in ANRIL knock-down cultures

Given the role of CDKN2A/2B in cell proliferation and their proximity to the ANRIL locus, we also checked the effect of ANRIL knockdown on proliferation of HuAoSMCs. There was marginal reduction in cell proliferation in cultures transfected with siRNA 1 as compared to siRNA-2 (Figure 4C).

Discussion

We have described a multi-layered analysis of the 9p21.3 region in a representative cohort from the Indian Atherosclerosis Research Study. We have shown strong association between common variants in the 9p21.3 locus and CAD thus replicating previously published findings. By comparing the expression profile of number of genes and ANRIL transcript in the INK4/ARF locus, we have shown that the mean expression of tumour suppressor/ cell cycle regulator gene transcripts, $p16^{INK4a}$, $p14^{ARF}$ and $p15^{INK4b}$ were significantly lower in cases than in controls, with the top quartiles of both $p16^{INK4a}$ and $p15^{INK4b}$ transcripts showing significant protec-
tion against CAD. Interestingly, the 9p21.3 risk allele was associated with higher expression of the short transcript, EU741058 rather than the full-length transcript, NR_003529. Further, blocking specific exon targets of ANRIL with siRNA in aortic smooth muscle cells cultures showed reduced the expression of CDKN2A transcripts, \textit{p16}^{\textit{INK4a}} and \textit{p14}^{\textit{ARF}}, both of which have been ascribed with regulation of cell cycle G1 progression. From these findings it appears that the 9p21 risk alleles located in the ANRIL locus modulate CAD risk by exerting their effect on the adjacent \textit{CDKN2A/2B} genes.

By virtue of its consistent performance in various genome wide association studies and following unprecedented evidence of strong association in meta-analysis, the common variants in 9p21.3 chromosomal region have emerged as the single most robust genetic locus for CAD (4, 5, 10, 30, 31). These variants are located within a 58-kb genomic region and present as a tight cluster, showing high inter-SNP correlation (r>0.90) (5, 32). The most redeeming features of these variants have been the portability of information across different ethnic global populations and their independent association with CAD (3, 20). In fact, the conferred increase in CAD risk for the carriers of either one or two risk alleles is reported to be \textasciitilde30\% and \textasciitilde60\%, respectively, which is the highest effect size for any genetic marker recorded to date with any adequate evidence (18).

In a pilot study on a representative subset of subjects from the IARS, we had previously reported 2.2-fold increased risk for CAD for the rs10757278 common variant (7). In the present study, we have replicated these findings and shown significant independent association of five 9p21.3 common variants. rs2383206 exhibited the highest independent association with an OR of 2.02, 95\% CI 1.56–2.62, \textit{p}<0.0001, which further increased to OR 2.55 with the addition of rs10757278 to the model. This, to our knowledge, is the highest risk estimate by any genetic marker for CAD. This could be partly attributed to the inclusion of a highly selected cohort of angiographically proven CAD patients with strong family history from a hospital-based setting that were compared to age and gender matched asymptomatic controls enrolled from the local community. In addition, the ‘GGGG’ haplotype was highly represented in our study. In comparison, the haplotype constituted by the alternate alleles, ‘GAAAA’ showed significant protection against CAD. In a comprehensive analysis of the 9p21.3 haplotypes, it has been shown that SNPs flanking previously reported SNPs have the ability to modify CAD risk in this population (22).

The 9p21.3 common variants are located within ANRIL, a long intergenic non coding RNA (33). ANRIL has been shown to exert its effect on CAD through epigenetic regulation of neighbouring tumour suppressor genes, \textit{CDKN2A} and \textit{CDKN2B}, implicated in cell cycle progression, cell proliferation and senescence. One of the early functional studies demonstrated that the risk allele increases gene expression in HuAoSMCs and homozygous carriers of the risk allele show 2.2-fold higher expression of the shorter splice variant, EU741058 and 1.2-fold lower expression of the long ANRIL variant in peripheral blood cells (16). In the same year, Liu et al reported reduced expression of all INK4A/ARF transcripts in the peripheral blood from healthy donors (34). In a subsequent study, Holdt et al. have shown higher expression of EU741058 and

![Figure 3: Plasma levels of CDKN2A and CDKN2B in cases and controls.](image)

![Figure 4: Results of in vitro functional studies. A) Comparison of expression levels of NR_003529 in siRNA-1 treated and untreated cell culture; B) Mean expression levels of tumour suppressor genes (\textit{p16}^{\textit{INK4a}}, \textit{p14}^{\textit{ARF}} and \textit{p15}^{\textit{INK4b}}) in siRNA-1 treated and untreated cell culture. C) Scatter plot showing the profile of cell proliferation rates in cultures with media alone, with scrambled siRNA and target specific siRNA-1 and siRNA-2.](image)
NR_003529 among carriers of risk haplotype both in peripheral blood and atherosclerotic plaque specimens (35). Interestingly, a recent paper revealed lower expression of ANRIL splice variant, targeting exon 1–2, among risk allele carriers (15). Thus it becomes apparent that ANRIL exists in many intermediary forms (splice variants) and the relationship between the risk allele and ANRIL expression may be determined by the specific exon/s being targeted in a given study. In fact, in the present study, we did not observe any significant difference in the two ANRIL transcripts between CAD samples and age and gender matched controls. However, we noted lower expression of the adjacent tumour suppressor genes in cases relative to the controls.

It is worthwhile at this juncture to reflect on the possible explanations for the contrasting results being reported on the differential pattern of expression of ANRIL and the neighboring tumour suppressor genes. Despite large amount of published work, the information remains rather incomplete. While the long non-coding RNAs (lncRNAs) such as ANRIL are fast evolving as a new class of regulatory RNAs, ANRIL per se is regulated by transcription factors (E2F1) (36), transcription factor binding sites (STAT 1) and a dense array of enhancers (37), variable response to inflammatory signalling molecules like interferon (IFN)-gamma (38), differential methylation pattern (39) etc. Presence of numerous splice variants further complicates matters (42).

Genes in the 9p21.3 region, CDKN2A/2B in particular, exist in a tightly linked cluster with ANRIL and may be co-regulated. The enhancers present in the 9p21.3 locus physically interact with CDKN2A/2B locus (37), while IFN-gamma has been recently shown to modulate the expression of p16INK4a and p15INK4b genes in diverse cell types, irrespective of the 9p21.3 risk genotype (38). Studies on orthologous human 9p21.3 locus in mice models has shown extensive, tissue-specific compensatory regulation of the Cdkn2a and Cdkn2b genes in the various types of knockout mice investigated (43). Thus, it appears that as multiple factors control the expression of CDKN2A/2B, it is understandable that there are contrasting reports from studies that have examined only one aspect i.e. candidate gene expression, in whole blood, PBMCs and/or cell lines derived from CAD patients and healthy donors but have not looked at various key inter-related aspects in totality.

The homozygous risk genotype was associated with higher expression of EU741058 and lower expression of p16INK4a gene and this pattern was similar across all the five SNPs tested. The inverse effect between ANRIL and p16INK4a gene expression indicate an antisense regulation of p16INK4a transcription. Interestingly, an inverse effect of risk allele has been previously reported for ANRIL with p15INK4b expression (40). Other genes such as p15INK4b, C9orf53, MTAP T1 and T2 did not show significant changes in expression with reference to the risk alleles. The underlying mechanism might be that in a diseased state, the risk allele induces higher expression of ANRIL, which then suppresses the expression of the tumour suppressor genes, indirectly promoting cell proliferation. The C9orf53 gene also showed nearly 1.6-fold lower expression in the CAD samples while, MTAP1 and MTAP 2, and other INK4b/ARF splice variants did not show differential expression. Presence of strong correlation in expression among the genes spanning the INK4/ARF locus both in the present study and other reported studies (40) imply co-regulation. The potential implications of the above findings are yet to be fully realised.

By using an in vitro HuAoSMC culture model, we have shown that treating cells with siRNA targeted against exon 19 resulted in lower expression of NR_003529 with parallel suppression of expression of p16INK4a and p14INK4b. To the contrary, siRNA targeting exon 2 only affected p16INK4a and p14INK4b expression as shown in Figure 4A and B. In a similar study, Congrains et al. have also shown reduction in CDKN2A expression as also a marked increase in CDKN2B expression which was not seen in our study (15). In both instances, siRNA knock-down efficiency was experimentally confirmed. Although we summarised our findings based on four rounds of experiments, additional validation might bring parity between outcomes in the two studies. Furthermore, siRNA knock-down of ANRIL exon 19 brought down the cellular proliferation rates in HuAoSMCs whereas knock-down of exon 2 did not bring about any significant change and was comparable to control cultures. This finding is compatible with the reported observation of Congrains et al. (15). In fact, another study has also shown that the

What is known about the topic?

- Genetic variants in the 9p21.3 chromosomal region show strong independent and consistent association with coronary artery disease (CAD) and other vascular phenotypes.
- The genetic variants are located within a 58 kb region containing a non-coding RNA called ANRIL and exhibit tight linkage disequilibrium among them.
- There is close interaction between ANRIL and adjacent tumour suppressor genes CDKN2A/2B that are important for cellular proliferation and cell cycle regulation, a fundamental process in atherosclerosis.
- There are conflicting reports on the pattern of association between the 9p21.3 risk alleles, the various ANRIL transcripts and adjacent tumour suppressor genes in CAD.

What does this paper add?

- The common risk allele shows a high frequency in this representative cohort of Asian Indians and a combination of rs2383206 and rs10757278 showed the best risk association with CAD.
- The association between ANRIL and tumour suppressor genes appears to be far stronger than with other neighbouring genes such as MTAP and C9orf. This was also evident from the strong association between the 9p21.3 risk alleles, ANRIL and expression of CDKN2A in particular as well as circulating CDKN2A levels.
- Transfection studies reveal strong co-regulation of ANRIL and CDKN locus that in turn affect proliferation of smooth muscle cells in vitro.
- This study demonstrates that the 9p21.3 risk alleles regulate smooth muscle proliferation, a fundamental process in the development of atherosclerosis, through their effect on ANRIL and CDKN genes.
risk allele influences VSMC proliferation (41). Such studies highlight important mechanisms that link the 9p21.3 genomic locus with CAD development through the regulation of cellular proliferation mechanism.

Plasma levels of CDKN2A were significantly lower in CAD samples as compared to controls while CDKN2B levels did not show a significant difference. This finding is in tune with other relevant observations recorded in this study; i.e. reduced expression of \( p16^{INK4a} \) and \( p14^{ARF} \) transcript in cases, in presence of the GG risk genotype and in response to siRNA targeting specific exons in ANRIL. It is well known that mRNA levels may not correlate with the corresponding protein levels because of post-transcriptional modifications. In line with this, we did not find significant correlation of \( p16^{INK4a} \) and \( p14^{ARF} \) transcript expression and CDKN2A levels although the overall trend in the mean levels was comparable. The small sample size used for protein estimation may have contributed to this discrepancy in our study.

In conclusion, the present study provides further evidence on the strong association between 9p21.3 genetic variants and CAD. Modulation of this risk appears to be mediated by ANRIL through its interactions with the adjacent cyclin-dependent kinase genes under the influence of the 9p21.3 risk alleles. Data gathered from peripheral blood expression analysis of the different splice variants in \( INK4a/ARF \) locus as well as the impact of knock-down of specific exon targets of ANRIL on adjacent genes and its end effect on cellular proliferation underline the singularly complex regulatory mechanisms in operation at the 9p21 locus. While our findings are concurrent with many other published studies, it is important to assimilate and realize the potential contribution of 9p21.3 genetic markers as independent factors in risk prediction, that can lead to correct reclassification of subjects into their appropriate risk groups, and which should eventually provide the necessary guidance for improving management of cardiovascular disease in a clinical setting.

Acknowledgements

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.

Conflicts of interest

None declared.

References