

ORIGINAL ARTICLE

A common variant in the *CDKN2B* gene on chromosome 9p21 protects against coronary artery disease in Americans of African ancestry

Brian G Kral^{1,8}, Rasika A Mathias^{1,8}, Bhoom Suktitipat², Ingo Ruczinski², Dhananjay Vaidya¹, Lisa R Yanek¹, Arshed A Quyyumi³, Riyaz S Patel³, A Maziar Zafari^{3,4}, Viola Vaccarino^{3,5}, Elizabeth R Hauser⁶, William E Kraus⁷, Lewis C Becker¹ and Diane M Becker¹

A 58 kb region on chromosome 9p21.3 has consistently shown strong association with coronary artery disease (CAD) in multiple genome-wide association studies in populations of European and East Asian ancestry. In this study, we sought to further characterize the role of genetic variants in 9p21.3 in African American individuals. Apparently healthy African American siblings ($n=548$) of patients with documented CAD <60 years of age were genotyped and followed for incident CAD for up to 17 years. Tests of association for 86 single-nucleotide polymorphisms (SNPs) across the 9p21.3 region in a generalized estimating equation logistic framework under an additive model adjusting for traditional risk factors, family, follow-up time and population stratification were performed. A single SNP within the *CDKN2B* gene met stringent criteria for statistical significance, including permutation-based evaluations. This variant, rs3217989, was common (minor allele (G) frequency 0.242), conveyed protection against CAD (odds ratio (OR)=0.19, 95% confidence interval (CI): 0.07 to 0.50, $P=0.0008$) and was replicated in a combined analysis of two additional case/control studies of prevalent CAD/MI in African Americans ($n=990$, $P=0.024$, OR=0.779, 95% CI: 0.626–0.968). This is the first report of a CAD association signal in a population of African ancestry with a common variant within the *CDKN2B* gene, independent from previous findings in European and East Asian ancestry populations. The findings demonstrate a significant protective effect against incident CAD in African American siblings of persons with premature CAD, with replication in a combination of two additional African American cohorts.

Journal of Human Genetics (2011) 56, 224–229; doi:10.1038/jhg.2010.171; published online 27 January 2011

Keywords: African American; CDKN2B; coronary artery disease; genetics; 9p21

INTRODUCTION

Although coronary artery disease (CAD) is the leading cause of death in African American men and women,^{1,2} there is a paucity of data regarding genetic variants related to CAD in persons of African compared with European ancestry. Consistently implicated in several large genome-wide association studies (GWAS) in primarily white, North American or northern European populations,^{3–7} the 9p21.3 chromosomal region seems to harbor a locus associated with CAD risk. Significant associations of specific single-nucleotide polymorphisms (SNPs) in this region have since been replicated in Chinese,^{8,9} Italian,^{10,11} South Korean,^{12,13} Japanese^{13,14} and American Hispanic¹⁵ populations but not in African Americans.^{4,15} The strongest associations are all within a 58 kb region (chromosome 9: 22 062 301–22 120 3894) containing multiple SNPs in tight linkage disequilibrium

(LD) in European-derived populations, but no known protein-coding genes. Nonetheless, the 9p21 locus seems to consistently convey risk independent of conventional CAD risk factors in general population studies.^{4,11} The strength of association of the 9p21 locus with CAD was recently demonstrated in a systematic review of 47 distinct data sets, including eight Asian studies, with 35 872 cases and 95 837 controls, but none with African Americans.¹⁶

Two cyclin-dependent kinase inhibitor genes, *CDKN2A* and *CDKN2B*, located ~115 kb away from the peak association signal in the 9p21.3 locus, have been suggested as the likely susceptibility genes in this region, and their translated proteins are thought to have important roles in cell cycle regulation.¹⁷ More recently, a large non-coding antisense transcribed RNA named ANRIL (*CDKN2BAS*) has been implicated as a possible regulatory element.^{11,18} The physical

¹Department of Medicine, The Johns Hopkins GeneSTAR Research Program, The Johns Hopkins University School of Medicine, Baltimore, MD, USA; ²The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; ³Department of Medicine, Division of Cardiology, Emory University School of Medicine, Atlanta, GA, USA; ⁴Atlanta Veterans Affairs Medical Center, Decatur, GA, USA; ⁵Rollins School of Public Health, Emory University, Atlanta, GA, USA and ⁶Center for Human Genetics, Division of Cardiology, Duke University Medical Center, Durham, NC, USA; ⁷Department of Medicine, Division of Cardiology, Duke University Medical Center, Durham, NC, USA

⁸These authors contributions are equal to this work.

Correspondence: Dr. BG Kral, Department of Medicine, The Johns Hopkins GeneSTAR Research Program, The Johns Hopkins University School of Medicine, 1830 East Monument Street, Room 8023, Baltimore, MD 21287, USA.

E-mail: bkral@jhmi.edu

Received 4 November 2010; revised 7 December 2010; accepted 20 December 2010; published online 27 January 2011

location of ANRIL is notable because it includes the SNPs previously strongly associated with CAD; it also overlaps (14763bp) with the physical gene locus *CDKN2B*.

Early-onset CAD is more heritable than CAD occurring at older ages.¹⁹ CAD at particularly young ages aggregates strongly in families, accounts for 60% of all CAD occurring before the of age 65 years,²⁰ and independently conveys an excess risk of CAD in first degree relatives two to five times that of the general population.^{19,21,22} We have previously shown a high prevalence of CAD risk factors²³ and subclinical CAD²⁴ as well as a higher than expected incidence of CAD events²⁵ in African Americans with a family history of premature CAD. In this study, we characterize the association of 9p21.3 genetic variants with incident CAD in initially healthy African American siblings of persons with documented early-onset CAD (<60 years of age). We test for replication of findings in two additional African American populations.

MATERIALS AND METHODS

Study populations

Primary study

GeneSTAR: Between 1990 and 2002, GeneSTAR enrolled 548 asymptomatic, apparently healthy young African American siblings (<60 years of age) of 278 patients with premature CAD in a prospective study of families with premature CAD, GeneSTAR (Genetic Study of Atherosclerosis Risk). Mean sibship size was 1.97 ± 1.2 . All siblings had DNA isolated and stored at the time of enrollment.

Siblings were identified from probands hospitalized with documented CAD including acute myocardial infarction (MI; $n=40$), coronary artery bypass surgery ($n=81$), percutaneous coronary intervention ($n=109$), angina with angiographic evidence of flow-limiting coronary stenosis ($n=40$) or following sudden cardiac death ($n=8$). Their siblings were eligible if they were <60 years of age and had no known history of CAD. Siblings were excluded if they had autoimmune disease, life-threatening co-morbidity (that is, acquired immunodeficiency syndrome, cancer) or were receiving chronic glucocorticosteroid therapy as previously described.²⁶ The study was approved by the Johns Hopkins Medicine Institutional Review Board and all study participants gave informed consent.

All eligible siblings underwent a baseline comprehensive risk factor screening following a 12-h overnight fast. A physical examination was carried out, blood was taken for lipid and glucose levels, and a complete medical history was elicited. Cardiac risk factors were defined using standard thresholds and methods as previously described.²⁵ Participants were followed at 5-year intervals for up to 17 years after baseline screening for incident CAD events. A trained telephone interviewer elicited a history of any cardiac-related procedures or symptoms, the use of any coronary disease or risk factor-related medications, and any history of CAD. Data was also recorded on all CAD diagnostic tests and results. In the event of a death, the closest family member was interviewed as a proxy. All reported CAD events and potential CAD events based on diagnostic procedures and therapies were reviewed from physician records, hospital records, death records and autopsy records using standardized methods as defined in the Framingham Heart Study, using the same classifications and definitions.²⁷ The first CAD event during the follow-up period was recorded and included sudden cardiac death, MI, unstable angina with coronary artery bypass surgery, unstable angina with percutaneous coronary intervention, stable angina with coronary artery bypass surgery or percutaneous coronary intervention, and medically treated angina with no intervention. Each person could only enter CAD event modeling once. All records were independently reviewed by three investigators, and any single discordant classification was referred to external cardiology reviewers as previously described using a standardized coding schema.²⁵

African American replication populations

Emory Genebank Study. The Emory Genebank Study was designed to investigate the relationship between biochemical and genetic factors with CAD in subjects undergoing cardiac catheterization. Study participants were enrolled at the Emory University Hospital, Emory University Hospital Midtown, Emory

Clinic and Grady Memorial Hospital in Atlanta, Georgia. Subjects with $\geq 50\%$ stenosis in one or more coronary arteries or with previous history of MI, coronary artery bypass surgery or percutaneous transluminal coronary angioplasty were defined as cases ($n=321$). Subjects with no evidence of CAD on cardiac catheterization and no previous history of MI or CAD were defined as controls ($n=146$). The appropriate Institutional Review Board approved the study and all subjects provided written informed consent. Information on ethnicity was self-reported.

CATHGEN. The study participants were enrolled at Duke University Medical Center (Durham, NC, USA) through the CATHGEN biorepository, consisting of subjects >18 years of age, recruited sequentially through the cardiac catheterization laboratories from 2001 to 2005. Biological samples and extensive clinical, angiographic and longitudinal follow-up data were collected on all subjects consenting to participation. Cases of MI ($n=280$) were defined as those having a history of MI (by self-report and corroborated by review of medical records using standardized criteria), or having suffered an MI during the study follow-up period using the same standardized methods to classify events. Controls ($n=243$) were defined as those with no previous history or evidence of significant CAD before or subsequent to cardiac catheterization, including MI, coronary revascularization, cardiomyopathy with an ejection fraction on left ventriculography <40%, or significant CAD on coronary angiography defined as a CAD index²⁸ ≤ 23 and no coronary vessel with a stenosis >50%. This study was approved by the Duke University Medical Center Institutional Review Board on Human Subjects and all subjects gave written informed consent.

Genotyping and SNP Selection in the primary study cohort (GeneSTAR)

In GeneSTAR, SNP genotyping was performed at deCODE Genetics (Reykjavik, Iceland) using the Human 1Mv1_C array from Illumina (San Diego, CA, USA), where 1 044 094 markers were released with an average call rate per sample of 99.65% and an overall missing data rate of 0.35%. The Illumina 1M array did not include rs10757278, shown previously to be associated with CAD,⁵ and this SNP was genotyped separately using a Taqman assay on an ABI Prism 7900HT Sequence Detection System (Applied Biosystem, Carlsbad, CA, USA) at deCODE Genetics, so as to avoid reliance on *in silico* genotyping methodologies, such as imputations. A total of 100 SNPs that mapped to the 9p21.3 region (chromosome 9: 21920505–22128762) were available for these analyses, of which 86 were non-monomorphic and informative for tests for association. We used PLINK²⁹ to detect and remove Mendelian errors. Hardy–Weinberg equilibrium and minor allele frequency (MAF) for each SNP was tested in a defined set of independent subjects ($n=326$) representing the founders of the pedigrees. We detected no deviation from Hardy–Weinberg equilibrium for any of the SNPs at our threshold of 0.0005 (that is, $P=0.05$ per number of SNPs tested). Admixture estimates were obtained using a subset of 18 982 SNPs from the GWAS array that were selected on the basis of low SNP correlation and high genetic distance between populations (F_{ST}) optimal for the differentiation of CEPH, Yoruban, and Chinese+Japanese HapMap (<http://www.hapmap.org>) samples selected as ancestral reference populations. Using STRUCTURE (v2.2; <http://pritch.bsd.uchicago.edu/software>) the mean estimated Yoruban ancestry in these 548 African American siblings was 79.29% (range 41.41–99.98%).

Genotyping in the replication populations (Genebank and CATHGEN)

The SNP rs3217989 was genotyped at deCODE Genetics using a Taqman assay on an ABI Prism 7900HT Sequence Detection System.

Statistical analyses

In the GeneSTAR population, logistic regression models were used to test for association between each individual SNP and incident CAD under the log-additive model (that is, a linear recoding of the SNP as 0/1/2 for 0, 1 and 2 copies of the minor allele). Regression models were implemented in the generalized estimating equation framework with an exchangeable covariance matrix to correct for familial correlation.³⁰ Data were analyzed adjusting for

traditional CAD covariates including age, gender, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, triglycerides, systolic blood pressure, body mass index, fasting-plasma glucose, smoking status and study follow-up time. Analyses were carried out using SAS v 9.1.3 (SAS Institute, Cary, NC, USA) and SUDAAN v 10.0 (Research Triangle Institute, Research Triangle Park, NC, USA). Given the range of Yoruban ancestry observed in the GeneSTAR African American siblings noted above, principal components-based estimates of admixture were obtained using the smartpca program in EIGENSOFT (v3.0; <http://genepath.med.harvard.edu/~reich/Software.htm>).³¹ Analyses were repeated with and without the inclusion of the first two eigenvectors in the logistic regression models. Because the correction for admixture did not affect the results, the models we present do not include admixture. Estimates of LD and LD block definition were calculated using Haploview 4.2 (<http://www.broad.mit.edu/mpg/haploview/>).³²

Permutation testing

We performed permutation tests for all SNPs with a MAF >4% to assess the genotype-specific null distributions in R (v. 2.9.0), which allowed for valid inferences in comparison with the observed test statistics. Event outcomes were shuffled 10 000 times across families of the same size, and then within families (thereby, keeping the number of affected individuals in families the same, and preserving the LD structure between SNPs). For each SNP, a *P*-value was recorded in each of the shuffles. To evaluate the significance of the association, we obtained an overall *P*-value for each genotype by counting how many of the 10 000 *P*-values obtained by shuffling the data were below the *P*-value observed in the original data. We considered the association of SNPs with CAD significant if this permutation *P*-value withstood a Bonferroni correction, controlling the family-wise error rate at 5%.

Replication analysis

SNPs that met our stringent permutation-based criteria for association with incident CAD in the discovery GeneSTAR population were genotyped for replication in a total of 990 individuals (601 cases and 389 controls) using data from the Emory Genebank Study and CATHGEN cohorts. Each study first separately performed an age- and sex-adjusted log-additive logistic regression analysis for each SNP to test for association with CAD outcomes. Using the age and sex adjusted calculated odds ratios (ORs), 95% confidence intervals (CIs), and the *P*-values from these two populations, we performed a SNP-based meta-analysis using a Mantel–Haenszel fixed effect model to calculate the overall combined OR, 95% CI and *P*-value, taking the direction of allelic effect into account.

RESULTS

GeneSTAR population

The GeneSTAR population had a baseline mean age of 46.9 ± 7.0 years (range 26–60 years of age). Baseline population characteristics are shown in Table 1. Traditional CAD risk factors are prevalent, as we have previously reported.²³ During a mean follow-up time of

8 ± 3 years (range 5–17 years), there were 35 CAD events (77% acute coronary syndromes, which included acute MI and unstable angina with revascularization, and 23% stable symptomatic CAD with angiographic evidence of >50% stenosis in at least one epicardial coronary vessel, with and without revascularization).

Association analyses in the GeneSTAR population

Logistic regression analyses of 86 individual informative SNPs with incident CAD, adjusting for age, sex, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, triglycerides, fasting blood glucose, systolic blood pressure, current smoking status, body mass index and years of follow-up time were performed using 35 incident CAD cases and 513 controls. Considering the LD structure between these SNPs, we considered a simple Bonferroni's correction for multiple testing of all SNPs to be overly conservative. Thus, we used 10 000 permutation tests to assess the significance of the individual SNPs (Figure 1). One SNP, rs3217989, met the threshold for significance and was strongly associated with incident CAD, independent of traditional CAD risk factors. This SNP, rs3217989, is located in the 3' untranslated region of the *CDKN2B* gene and its minor allele (MAF=0.242) is protective against incident CAD, with an OR=0.19 (95% CI: 0.07 to 0.50, *P*=0.0008).

Figure 2 shows the chromosomal region of 9p21.3 (chromosome 9: 21 920 505–22 128 762) illustrating the physical location and level of significance of SNPs associated with CAD and the LD structure between the 86 SNPs in the GeneSTAR African American siblings.

Of the nine SNPs that passed a nominal threshold at the asymptotic level, a single SNP passed the permutation-based approach (rs3217989, Table 2) and is independent of the previously published region on 9p21 (rs10757278, Table 3). Table 3 demonstrates very low correlations in the nine GeneSTAR SNPs with the two most significant GeneSTAR SNPs (rs3217989 and rs17761446), and with the previously published lead SNP (rs10757278).

Because of the lack of representation on the GWAS array used, additional genotyping was carried out on rs10757278, one of the most published SNPs in the 9p21 region associated with MI and CAD in populations other than African ancestry.⁵ We found no significant association between this SNP and CAD (*P*=0.709, MAF=21%). Furthermore, rs10757278 seems to be uncorrelated with the peak SNP, rs3217989, in our data (Table 3). Of note, we observed a *P*-value of 0.002 at rs17761446 (MAF=5.4%), a SNP that is 6.4 kb away from rs10757278. However, this SNP did not pass our permutation-based criteria for statistical significance. In addition, there was no correlation between rs17761446 and rs10757278 ($r^2=0.017$ in our data), suggesting that rs17761446 is not likely to be representative of the peak signal previously published at rs10757278.

Replication of association

The MAFs of rs3217989 in the Emory Genebank Study and CATHGEN African Americans were similar to that found in the GeneSTAR population (0.264 and 0.280, respectively). Genebank and CATHGEN each separately showed a trend for a protective effect of the minor allele of rs3217989 (*P*=0.15 and 0.08, respectively), with an overall significant combined meta-analysis result (meta-analysis OR=0.78, 95% CI: 0.63–0.97, *P*=0.02). The population characteristics of Genebank and CATHGEN can be found in Supplementary Table 1.

DISCUSSION

This is the first and only study in African Americans to date to show a significant association of CAD with any gene variant within the chromosome 9p21. 3 locus. This SNP, rs3217989, is located within the 3' untranslated region of *CDKN2B* and is independent from the

Table 1 Baseline population characteristics of the GeneSTAR African American siblings (*n*=548)^a

Age (years)	46.9 ± 7.0
Incident CAD (%)	6.4
Age at CAD years	55.6 ± 7.1
Male sex (%)	34.5
LDL cholesterol (mmol l ⁻¹)	3.52 ± 1.12
HDL cholesterol (mmol l ⁻¹)	1.40 ± 0.42
Triglycerides (mmol l ⁻¹)	1.37 ± 0.86
Systolic blood pressure (mm Hg)	135.3 ± 16.2
Fasting glucose (mmol l ⁻¹)	5.72 ± 2.35
Current smoking (%)	32.1
BMI (kg m ⁻²)	31.3 ± 6.6

Abbreviations: BMI, body mass index; CAD, coronary artery disease; HDL, high-density lipoprotein; LDL, low-density lipoprotein.

^aContinuous variables presented as mean ± 1 s.d.

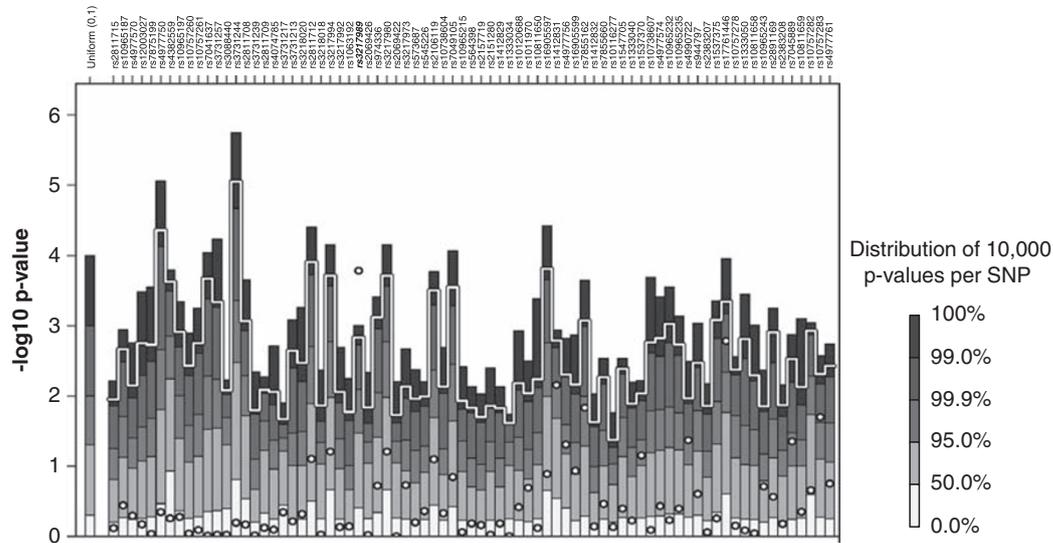


Figure 1 Permutation-based P -values per SNP for 77 SNPs having MAF > 4%. Shaded regions represent the 100th, 99.9th, 99th, 95th and 50th percent of the distribution of permutation P -values per SNP as indicated in the legend. On the far left, the uniform (0, 1) distribution highlights the departure of the SNP finite sample null distributions from this nominal (asymptotic) distribution. The white line indicates the P -value cutoff for each SNP necessary to achieve significance after a Bonferroni's correction, derived from the permutation distribution and white dots indicate the observed P -value. Significant SNPs are those where the white dot is above the white line and are indicated by a bold SNP name. From Table 2, rs1329881 had an asymptotic P -value < 0.05 but a MAF < 4% and could not be permuted (asymptotic P -value=0.024).

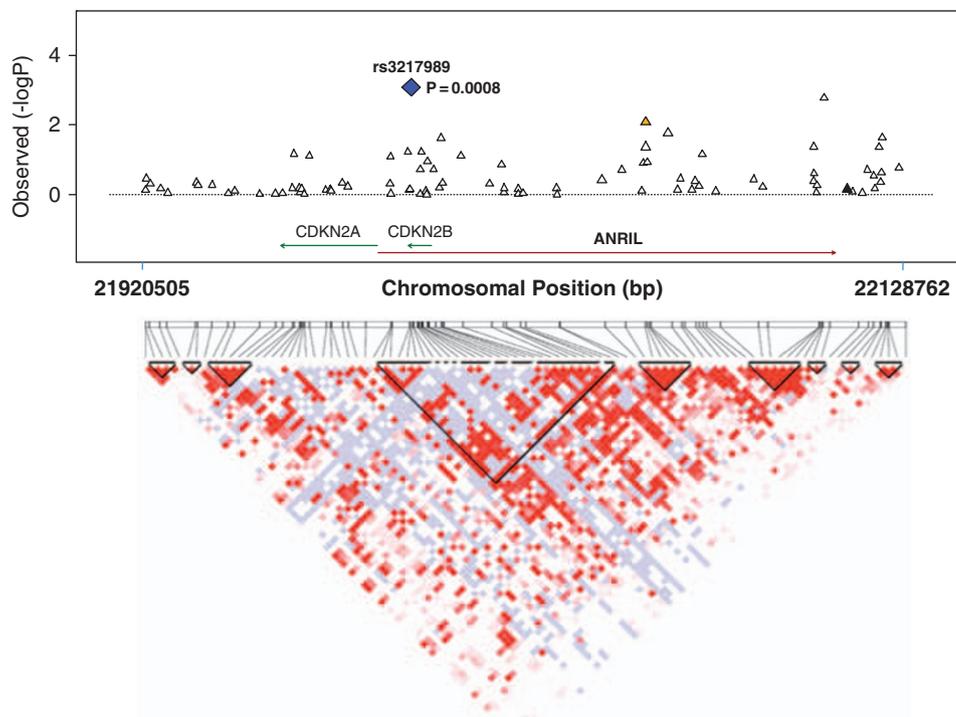


Figure 2 Graphical overview of the 9p21.3 locus displaying location and asymptotic P -values of 86 SNPs and LD structure between the 86 SNPs in the GeneSTAR African American siblings along with annotated genes and transcribed mRNA in the region. The blue diamond represents rs3217989, the peak SNP in the GeneSTAR African American siblings. The black triangle represents rs10757278, previously reported by Helgadottir *et al.*,⁵ the orange triangle represents rs1412831, the only SNP with correlation > 0.7 as shown in Table 3 with rs3217989. Location of *CDKN2A*, *CDKN2B*, and the large non-coding RNA ANRIL are depicted in the green and red lines, respectively. The lower portion of the figure shows the LD structure in this region, presenting strength of D' ranging from high (red) to low (white).

LD block previously associated with CAD in non-African ancestry populations. In our GeneSTAR cohort the protective effect of the minor allele is potent, with almost a fivefold decrease in incident CAD

risk. The direction of effect was replicated and also found to be significant in a meta-analysis of two populations of Americans with similar African ancestry in North Carolina and Georgia.

Table 2 Asymptotic results on tests for association for nine SNPs with P -value < 0.05 in the GeneSTAR African American siblings

SNP	+Strand	–Strand	Position	OR (95% CI)	P-value	MAF
rs3217989	ANRIL	CDKN2B	21993790	0.18 (0.06–0.49)	0.0008	0.222
rs13298881 ^a	ANRIL		22002051	0.21 (0.05–0.81)	0.024	0.036
rs1412831	ANRIL		22058646	0.33 (0.14–0.75)	0.008	0.218
rs4977756	ANRIL		22058652	0.43 (0.19–0.97)	0.044	0.354
rs7855162	ANRIL		22064793	0.50 (0.29–0.88)	0.017	0.340
rs4990722	ANRIL		22105217	0.36 (0.13–0.96)	0.043	0.137
rs17761446	ANRIL		22108102	4.84 (1.81–12.9)	0.002	0.054
rs7045889			22123251	0.55 (0.31–0.98)	0.044	0.472
rs10757283			22124172	0.52 (0.30–0.91)	0.023	0.454

Abbreviations: CI, confidence interval; MAF, minor allele frequency; OR, odds ratio; SNP, single-nucleotide polymorphism.

^ars13298881 (MAF < 4%) excluded from Figure 1.

Table 3 Correlation (r^2) of nine SNPs with P -value < 0.05 with peak SNPs (rs3217989, rs17761446) and previously published SNP (rs10757278)^a in GeneSTAR African American siblings; one pair (rs3217989 and rs1412831) has an $r^2 > 0.7$

GeneSTAR SNPs	rs3217989	rs17761446	rs10757278 ^a
rs3217989	—	0.017	0.054
rs13298881	0.011	0.000	0.004
rs1412831	0.743	0.002	0.047
rs4977756	0.442	0.008	0.036
rs7855162	0.160	0.011	0.092
rs4990722	0.059	0.009	0.038
rs17761446	0.017	—	0.016
rs7045889	0.033	0.041	0.017
rs10757283	0.059	0.038	0.004

Abbreviation: SNP, single-nucleotide polymorphism.

^aFrom Helgadottir *et al.*⁵

We postulate that the heterogeneity in the magnitude of the OR estimates in the discovery and replication samples is a function of study design. The cumulative CAD incidence in GeneSTAR was 6% and by design, GeneSTAR CAD events occurred in persons who were likely close in age-range, genetic susceptibility, and shared environment to the proband. The GeneSTAR sibling CAD events were thus likely causally more homogeneous as compared with all possible causes of CAD in the general population. In contrast, the population-based case–control studies, which are 50% cases and 50% controls by design, likely have greater variety of genetic and environmental factors among those subjects with CAD events. Given greater probable causal heterogeneity, CAD cases in these would demonstrate an OR closer to null than that seen in the discovery cohort. The observation of a consistent protective effect between the discovery and replication samples supports our finding in the discovery population. Although it would have been desirable to have more replication populations of greater size, very few African American are fully phenotyped for CAD or have a GWAS. The fact that our finding remains robust in small samples supports the likelihood that it is real.

No other study has published any associations with any gene variants, identified by GWAS, and CAD in African American populations. In the only GWAS reporting CAD results in African Americans, the 9p21 polymorphisms studied were not significantly associated

with CAD.⁴ However, in studies published to date, the statistical power has been limited.⁴ One study in African Americans was limited to three significant SNPs identified from a GWAS in populations of European ancestry and there was again no detectable significance.¹⁵ Our results in African American families show that rs3217989 is in a different LD block than the lead SNPs previously found in the non-protein coding region in 9p21 in other populations. However, we also acknowledge that our study had limited statistical power to detect the previously reported association of these lead SNPs, including rs10757278, with CAD. Nonetheless, we were able to show that rs3217989 was independent of rs10757278, and therefore the lack of association of rs10757278 does not diminish the significance of our primary findings.

SNP rs3217989 appears to be monomorphic in populations of European ancestry in the Human Genome Diversity Panel data (<http://hgdp.uchicago.edu/cgi-bin/gbrowse/HGDP/>). Given more fragmented LD structure in African Americans, the finding of a risk allele at rs3217989 suggests a functional variant located closer to rs3217989 in persons of African ancestry than the traditional previously published 9p21 intergenic locus.

The 9p21 locus previously identified in most studies has very few gene candidates. Previous GWAS studies and subsequent replications have reported that most associations at genome-wide significance occur in a large haplotype block upstream and independent from two major genes, *CDKN2A* and *CDKN2B*.⁴ These genes encode protein inhibitors of cyclin-dependent kinases, p16^{INK4a} and p15^{INK4b}, respectively, expressed at high levels in endothelial and inflammatory cells, and are also thought to help regulate cell proliferation, cell aging and apoptosis.^{33–35}

ANRIL (*CDKNA2BAS*), which encodes a large antisense non-coding RNA, is another candidate locus in this region. *ANRIL* spans 126.3 kb and consists of 20 exons subjected to alternate splicing including the first two exons which appear to overlap two exons of *CDKN2B*.³⁶ *ANRIL* expression has been documented in atherosclerotic vessels, vascular endothelial cells, monocyte-derived macrophages and coronary smooth muscle cells.¹¹ In a subset of individuals in the Ottawa Heart Study,¹⁸ the previously identified 9p21 risk alleles were associated with *ANRIL* mRNA of differential lengths. Furthermore, expression of *CDKN2B* was correlated with the long variant of *ANRIL*, suggesting that the 9p21 risk alleles may be biologically tied to atherosclerosis risk through *CDKN2B* expression. Given these important findings, it is possible that the association of the *CDKN2B* gene variant with CAD in GeneSTAR African Americans is related to an alteration in *ANRIL*. However, the GeneSTAR *CDKN2B* SNP is located at the 5' end of *ANRIL*, whereas the aforementioned SNPs are located more at the 3' end of *ANRIL*, from exon 13 to 20, where *ANRIL* splicing variation has been shown.¹⁸ Alternatively, the variant we report in the 3' untranslated region of *CDKN2B* may contribute to an alteration in *CDKN2B* expression and/or function independently of *ANRIL*. For example, the 3' untranslated region could contain a regulatory binding site for transcription or stability of the coded protein.

CONCLUSION

We have found a novel variant in *CDKN2B* located in an LD block distinct from that of the previous European-derived signals in 9p21. This variant appears to be protective against CAD in African Americans. Further investigation of the *CDKN2B* gene and protein as well as the non-coding RNA *ANRIL* in persons of African ancestry will be necessary to provide additional clues as to the biological mechanisms of association between chromosome 9p21.3 and CAD.

CONFLICT OF INTEREST

The author declares no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by grants HL072518 from the National Heart, Lung and Blood Institute, and M01-RR00052 from the National Center for Research Resources, National Institutes of Health, Bethesda, MD. Genotyping of the Emory and Duke samples at deCODE genetics was supported by NIH grant R01HL089650-02 from the National Heart, Lung and Blood Institute.

- 1 Ferdinand, K. C. Coronary artery disease in minority racial and ethnic groups in the United States. *Am. J. Cardiol.* **97**, 12A–19A (2006).
- 2 Carnethon, M. R., Lynch, E. B., Dyer, A. R., Lloyd-Jones, D. M., Wang, R., Garside, D. B. *et al.* Comparison of risk factors for cardiovascular mortality in black and white adults. *Arch. Intern. Med.* **166**, 1196–1202 (2006).
- 3 Samani, N. J., Erdmann, J., Hall, A. S., Hengstenberg, C., Mangino, M., Mayer, B. *et al.* For WTCCC and the cardiogenics consortium. Genomewide association analysis of coronary artery disease. *N. Engl. J. Med.* **357**, 443–453 (2007).
- 4 McPherson, R., Pertsemlidis, A., Kavavlar, N., Stewart, A., Roberts, R., Cox, D. R. *et al.* A common allele on chromosome 9 associated with coronary heart disease. *Science* **316**, 1488–1491 (2007).
- 5 Helgadóttir, A., Thorleifsson, G., Manolescu, A., Gretarsdóttir, S., Blondal, T., Jonasdóttir, A. *et al.* A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* **316**, 1491–1493 (2007).
- 6 Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3000 shared controls. *Nature* **47**, 661–678 (2007).
- 7 Kathiresan, S., Voight, B. F., Purcell, S., Musunuru, K., Ardisino, D., Mannucci, P. M. *et al.* Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat. Genet.* **41**, 334–341 (2009).
- 8 Zhang, Q., Wang, X. F., Cheng, S. S., Wan, X. H., Cao, F. F., Li, L. *et al.* Three SNPs on chromosome 9p21 confer increased risk of myocardial infarction in Chinese subjects. *Atherosclerosis* **207**, 26–28 (2009).
- 9 Ding, H., Xu, Y., Wang, X., Wang, Q., Zhang, L., Tu, Y. *et al.* 9p21 is a shared susceptibility locus strongly for coronary artery disease and weakly for ischemic stroke in Chinese Han population. *Circ. Cardiovasc. Genet.* **2**, 338–346 (2009).
- 10 Shen, G. Q., Rao, S., Martinelli, N., Li, L., Olivieri, O., Corrocher, R. *et al.* Association between four SNPs on chromosome 9p21 and myocardial infarction is replicated in an Italian population. *J. Hum. Genet.* **53**, 144–150 (2008).
- 11 Broadbent, H. M., Peden, J. F., Lorkowski, S., Goel, A., Ongen, H., Green, F. *et al.* For the PROCARDIS consortium. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p21. *Hum. Mol. Genet.* **17**, 806–814 (2008).
- 12 Shen, G. Q., Li, L., Rao, S., Abdullah, K. G., Ban, J. M., Lee, B. S. *et al.* Four SNPs on chromosome 9p21 in a South Korean population implicate a genetic locus that confers high cross-race risk for development of coronary artery disease. *Arterioscler. Thromb. Vasc. Biol.* **28**, 360–365 (2008).
- 13 Hinohara, K., Nakajima, T., Takahashi, M., Hohda, S., Sasaoka, T., Nakahara, K. *et al.* Replication of the association between a chromosome 9p21 polymorphism and coronary artery disease in Japanese and Korean populations. *J. Hum. Genet.* **53**, 357–359 (2008).
- 14 Hiura, Y., Fukushima, Y., Yuno, M., Sawamura, H., Kokubo, Y., Okamura, T. *et al.* Validation of the association of genetic variants on chromosome 9p21 and 1q41 with myocardial infarction in a Japanese population. *Circ. J.* **72**, 1213–1217 (2008).
- 15 Assimes, T. L., Knowles, J. W., Basu, A., Iribarren, C., Southwick, A., Tang, H. *et al.* Susceptibility locus for clinical and subclinical coronary artery disease at chromosome 9p21 in the multi-ethnic ADVANCE study. *Hum. Mol. Genet.* **17**, 2320–2328 (2008).
- 16 Palomaki, G. E., Melillo, S. & Bradley, L. A. Association between 9p21 genomic markers and heart disease: a meta-analysis. *JAMA* **303**, 648–656 (2010).
- 17 Kim, W. Y. & Sharpless, N. E. The regulation of INK4/ARF in cancer and aging. *Cell* **127**, 265–275 (2006).
- 18 Jarinova, O., Stewart, A. F., Roberts, R., Wells, G., Lau, P., Naing, T. *et al.* Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus. *Arterioscler. Thromb. Vasc. Biol.* **29**, 1671–1677 (2009).
- 19 Nora, J. J., Lortscher, R. H., Spangler, R. D., Nora, A. H. & Kimberling, W. J. Genetic—epidemiologic study of early-onset ischemic heart disease. *Circulation* **61**, 503–508 (1980).
- 20 Williams, R. R. Understanding genetic and environmental risk factors in susceptible persons. *West. J. Med.* **141**, 799–806 (1984).
- 21 Deutscher, S., Ostrander, L. D. & Epstein, F. H. Familial factors in premature coronary heart disease—a preliminary report from the Tecumseh Community Health Study. *Am. J. Epidemiol.* **91**, 233–237 (1970).
- 22 Rissanen, A. M. Familial occurrence of coronary heart disease: effect of age at diagnosis. *Am. J. Cardiol.* **44**, 60–66 (1979).
- 23 Becker, D. M., Yook, R. M., Moy, T. F., Blumenthal, R. S. & Becker, L. C. Markedly high prevalence of coronary risk factors in apparently healthy African-American and white siblings of persons with premature coronary heart disease. *Am. J. Cardiol.* **82**, 1046–1051 (1998).
- 24 Blumenthal, R. S., Becker, D. M., Yanek, L. R., Moy, T. F., Michos, E. D., Fishman, E. K. *et al.* Comparison of coronary calcium and stress myocardial perfusion imaging in apparently healthy siblings of individuals with premature coronary artery disease. *Am. J. Cardiol.* **97**, 328–333 (2006).
- 25 Vaidya, D., Yanek, L. R., Moy, T. F., Pearson, T. A., Becker, L. C. & Becker, D. M. Incidence of coronary artery disease in siblings of patients with premature coronary artery disease: 10 years of follow-up. *Am. J. Cardiol.* **100**, 1410–1415 (2007).
- 26 Blumenthal, R. S., Becker, D. M., Moy, T. F., Coresh, J., Wilder, L. B. & Becker, L. C. Exercise thallium tomography predicts future clinically manifest coronary heart disease in a high-risk asymptomatic population. *Circulation* **93**, 915–923 (1996).
- 27 Wilson, P. W., D'Agostino, R. B., Levy, D., Belanger, A. M., Silbershatz, H. & Kannel, W. B. Prediction of coronary heart disease using risk factor categories. *Circulation* **97**, 1837–1847 (1998).
- 28 Smith, L. R., Harrell, F. E. Jr, Rankin, J. S., Califf, R. M., Pryor, D. B., Muhlbaier, L. H. *et al.* Determinants of early versus late cardiac death in patients undergoing coronary artery bypass graft surgery. *Circulation* **84** (5: Suppl), III245–III253 (1991).
- 29 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
- 30 Zeger, S. L. & Liang, K. Y. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* **42**, 121–130 (1986).
- 31 Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A. & Reich, D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
- 32 Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
- 33 Kamb, A., Gruis, N. A., Weaver-Feldhaus, J., Liu, Q., Harshman, K., Tavtigian, S. V. *et al.* A cell cycle regulator potentially involved in genesis of many tumor types. *Science* **264**, 436–440 (1994).
- 34 Pomerantz, J., Schreiber-Agus, N., Liegeois, N. J., Silverman, A., Alland, L., Chin, L. *et al.* The Ink4a tumor suppressor gene product, p19Arf, interacts with MDM2 and neutralizes MDM2's inhibition of p53. *Cell* **92**, 713–723 (1998).
- 35 Hannon, G. J. & Beach, D. p15INK4B is a potential effector of TGF-beta-induced cell cycle arrest. *Nature* **371**, 257–261 (1994).
- 36 Pasmant, E., Laurendeau, I., Heron, D., Vidaud, M., Vidaud, D. & Bieche, I. Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res.* **67**, 3963–3969 (2007).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)