



Published in final edited form as:

*J Hum Genet.* 2011 March ; 56(3): 224–229. doi:10.1038/jhg.2010.171.

## A Common Variant in the *CDKN2B* Gene on Chromosome 9p21 Protects Against Coronary Artery Disease in Americans of African Ancestry:

### *CDKN2B* Gene Variant and CAD in African Americans

Brian G. Kral<sup>1,\*</sup>, Rasika A. Mathias<sup>1,\*</sup>, Bhoom Suktitipat<sup>2</sup>, Ingo Ruczinski<sup>2</sup>, Dhananjay Vaidya<sup>1</sup>, Lisa R. Yanek<sup>1</sup>, Arshed A. Quyyumi<sup>3</sup>, Riyaz S. Patel<sup>3</sup>, A Maziar Zafari<sup>3,4</sup>, Viola Vaccarino<sup>3,5</sup>, Elizabeth R. Hauser<sup>6</sup>, William E. Kraus<sup>7</sup>, Lewis C. Becker<sup>1</sup>, and Diane M. Becker<sup>1</sup>

<sup>1</sup>Department of Medicine GeneSTAR Research Program, The Johns Hopkins University School of Medicine, Baltimore, MD, USA

<sup>2</sup>The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

<sup>3</sup>Department of Medicine, Division of Cardiology, Emory University School of Medicine, Atlanta, GA, USA

<sup>4</sup>Atlanta Veterans Affairs Medical Center, Decatur, GA, USA

<sup>5</sup>Rollins School of Public Health, Emory University, Atlanta, GA, USA

<sup>6</sup>Center for Human Genetics, Duke University Medical Center, Durham, NC, USA.

<sup>7</sup>Department of Medicine, Division of Cardiology, Duke University Medical Center, Durham, NC, USA.

### Abstract

**Background**—A 58kb 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.

**Methods and Results**—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 SNPs across the 9p21.3 region in a GEE 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,

Users may view, print, copy, download and text and data- mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

Correspondence should be addressed to: Diane M. Becker, Sc.D., The Johns Hopkins GeneSTAR Research Program, 1830 East Monument Street, Room 8023, Baltimore, MD 21287 dbecker@jhmi.edu, Telephone: 410-955-7782, Fax: 410-955-0321.

\* Author contributions are equal

### Disclosures

None.

rs3217989, was common (minor allele [G] frequency 0.242), conveyed protection against CAD (OR=0.19, 95% 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).

**Conclusions**—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.

### 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<sup>1, 2</sup>, there is a paucity of data regarding genetic variants related to CAD in persons of African compared to European ancestry. Consistently implicated in several large genome-wide association studies (GWAS) in primarily white, North American or Northern European populations<sup>3-7</sup>, the 9p21.3 chromosomal region appears 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<sup>8, 9</sup>, Italian<sup>10</sup>, 11, South Korean<sup>12, 13</sup>, Japanese<sup>13, 14</sup>, and American Hispanic<sup>15</sup> populations but not in African Americans<sup>4, 15</sup>. 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 appears to consistently convey risk independent of conventional CAD risk factors in general population studies<sup>4, 11</sup>. The strength of association of the 9p21 locus with CAD was recently demonstrated in a systematic review of 47 distinct data sets, including 8 Asian studies, with 35,872 cases and 95,837 controls, but none with African Americans<sup>16</sup>.

Two cyclin-dependent kinase inhibitor genes, *CDKN2A* and *CDKN2B*, located ~115kb 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 play important roles in cell cycle regulation<sup>17</sup>. More recently, a large noncoding antisense transcribed RNA named ANRIL (*CDKN2BAS*) has been implicated as a possible regulatory element<sup>11, 18</sup>. The physical 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<sup>19</sup>. CAD at particularly young ages aggregates strongly in families, accounts for 60% of all CAD occurring prior to age 65<sup>20</sup>, and independently conveys an excess risk of CAD in first degree relatives 2-5 times that of the general population<sup>19, 21, 22</sup>. We have previously shown a high prevalence of CAD risk factors<sup>23</sup> and subclinical CAD<sup>24</sup> as well as a higher than expected incidence of

CAD events<sup>25</sup> in African Americans with a family history of premature CAD. In the current 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  $1.97 \pm 1.2$ . 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 (CABG) (n=81), percutaneous coronary intervention (PCI) (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 (i.e. AIDS, cancer), or were receiving chronic glucocorticosteroid therapy as previously described<sup>26</sup>. 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-hour overnight fast. A physical examination was performed, 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<sup>25</sup>. Participants were followed at five-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<sup>27</sup>. The first CAD event during the follow-up period was recorded and included sudden cardiac death, MI, unstable angina with CABG, unstable angina with PCI, stable angina with CABG or PCI, 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<sup>25</sup>.

## 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 prior history of MI, CABG or percutaneous transluminal coronary angioplasty (PTCA) were defined as cases ( $n=321$ ). Subjects with no evidence of coronary artery disease on cardiac catheterization and no prior 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, North Carolina) through the CATHGEN biorepository, consisting of subjects greater than 18 years of age, recruited sequentially through the cardiac catheterization laboratories from 2001-2005. 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 prior to 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 index  $\geq 28-23$  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, Inc. using the Human 1Mv1\_C array from Illumina, Inc. 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 CAD5, and this SNP was genotyped separately using a Taqman assay on an ABI Prims 7900HT Sequence Detection System at deCODE Genetics, Inc., 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 (HWE) 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 HWE for any of the SNPs at our threshold of 0.0005 (i.e.  $p=0.05/\text{number of SNPs tested}$ ). 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

([www.hapmap.org](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, Inc. using a Taqman assay on an ABI Primis 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 (i.e. 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 (GEE) framework with an exchangeable covariance matrix to correct for familial correlation<sup>30</sup>. Data were analyzed adjusting for traditional CAD covariates including age, gender, LDL cholesterol, HDL 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 Inc., Cary, NC] and SUDAAN v 10.0 [Research Triangle Institute, Research Triangle Park, NC]. 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<sup>31</sup>. 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 4.232.

### 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 ORs, 95% 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  $46.9 \pm 7.0$  years (range 26 to 60 years of age). Baseline population characteristics are shown in Table 1. Traditional CAD risk factors are prevalent, as we have previously reported<sup>23</sup>. During a mean follow-up time of  $8 \pm 3$  years (range 5 to 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, LDL cholesterol, HDL 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 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'UTR 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).

Due to its lack of representation on the GWAS array used, additional genotyping was done on rs10757278, one of the most published SNPs in the 9p21 region associated with MI and CAD in populations other than African ancestry<sup>5</sup>. We found no significant association between this SNP and CAD (p=0.709, MAF=21%). Furthermore, rs10757278 appears 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 which is 6.4kb away from rs10757278. However this SNP did not pass our permutation-based criteria for statistical significance. Additionally, 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'UTR of *CDKN2B* and is independent from the 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 five-fold 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.

We postulate that the heterogeneity in the magnitude of the odds ratio 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 to 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 odds ratio 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 CAD4. However, in studies published to date, the statistical power has been limited4. 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<sup>15</sup>. 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*<sup>4</sup>. These genes encode protein inhibitors of cyclin-dependent kinases, p16<sup>INK4a</sup> and p15<sup>INK4b</sup>, respectively, expressed at high levels in endothelial and inflammatory cells, and are also thought to help regulate cell proliferation, cell aging, and apoptosis<sup>33-35</sup>.

*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*<sup>36</sup>. *ANRIL* expression has been documented in atheromatous vessels, vascular endothelial cells, monocyte-derived macrophages, and coronary smooth muscle cells<sup>11</sup>. In a subset of individuals in the Ottawa Heart Study<sup>18</sup> 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.<sup>18</sup> Alternatively, the variant we report in the 3'-UTR of *CDKN2B* may contribute to an alteration in *CDKN2B* expression and/or function independently of *ANRIL*. For example, the 3'UTR 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 prior 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.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## 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.

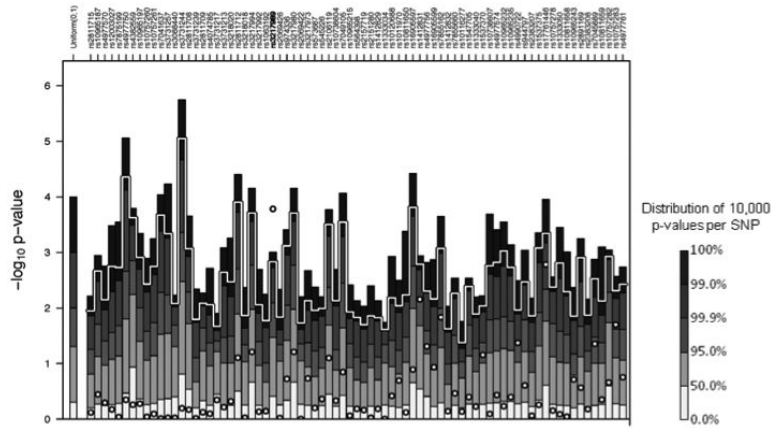
## REFERENCES

1. Ferdinand KC. Coronary artery disease in minority racial and ethnic groups in the United States. *Am. J. Cardiol.* 2006; 97:12A–19A.
2. Carnethon MR, Lynch EB, Dyer AR, Lloyd-Jones DM, Wang R, Garside DB, et al. Comparison of risk factors for cardiovascular mortality in black and white adults. *Arch Intern. Med.* 2006; 166:1196–1202. [PubMed: 16772247]
3. Samani NJ, Erdmann J, Hall AS, 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.* 2007; 357:443–453. [PubMed: 17634449]
4. McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, et al. A common allele on chromosome 9 associated with coronary heart disease. *Science.* 2007; 316:1488–1491. [PubMed: 17478681]
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.* 2007; 316:1491–1493. [PubMed: 17478679]
6. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007; 447:661–678. [PubMed: 17554300]
7. Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, Mannucci PM, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat. Genet.* 2009; 41:334–341. [PubMed: 19198609]
8. Zhang Q, Wang XF, Cheng SS, Wan XH, Cao FF, Li L, et al. Three SNPs on chromosome 9p21 confer increased risk of myocardial infarction in Chinese subjects. *Atherosclerosis.* 2009; 207:26–28. [PubMed: 19467655]
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.* 2009; 2:338–346. [PubMed: 20031605]
10. Shen GQ, 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.* 2008; 53:144–150. [PubMed: 18066490]
11. Broadbent HM, Peden JF, 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.* 2008; 17:806–814. [PubMed: 18048406]
12. Shen GQ, Li L, Rao S, Abdullah KG, Ban JM, Lee BS, et al. Four SNPs on chromosome 9p21 in a South Korean population implicate a genetic locus that confers high cross-race risk for

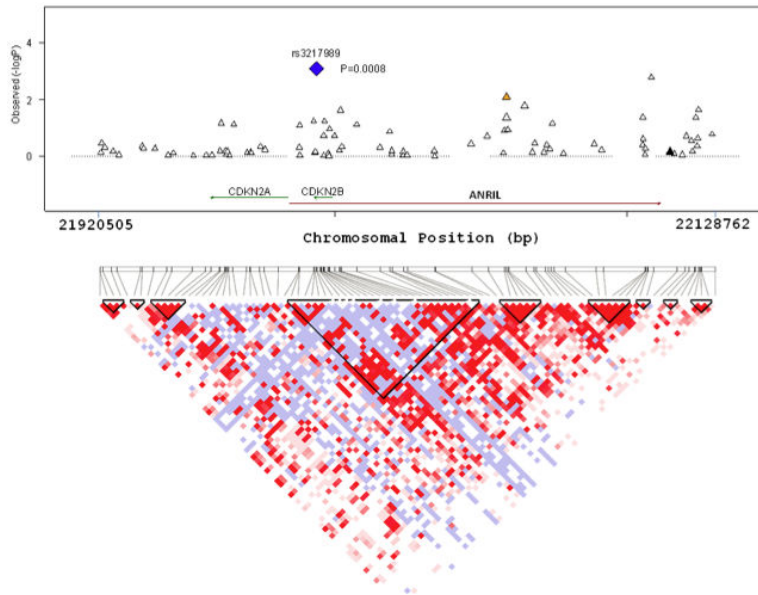
development of coronary artery disease. *Arterioscler. Thromb. Vasc. Biol.* 2008; 28:360–365. [PubMed: 18048766]

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.* 2008; 53:357–359. [PubMed: 18264662]
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.* 2008; 72:1213–1217. [PubMed: 18654002]
15. Assimes TL, Knowles JW, 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.* 2008; 17:2320–2328. [PubMed: 18443000]
16. Palomaki GE, Melillo S, Bradley LA. Association between 9p21 genomic markers and heart disease: a meta-analysis. *JAMA.* 2010; 303:648–656. [PubMed: 20159873]
17. Kim WY, Sharpless NE. The regulation of INK4/ARF in cancer and aging. *Cell.* 2006; 127:265–275. [PubMed: 17055429]
18. Jarinova O, Stewart AF, 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.* 2009; 29:1671–1677. [PubMed: 19592466]
19. Nora JJ, Lortscher RH, Spangler RD, Nora AH, Kimberling WJ. Genetic--epidemiologic study of early-onset ischemic heart disease. *Circulation.* 1980; 61:503–508. [PubMed: 7353240]
20. Williams RR. Understanding genetic and environmental risk factors in susceptible persons. *West. J. Med.* 1984; 141:799–806. [PubMed: 6596795]
21. Deutscher S, Ostrander LD, Epstein FH. Familial factors in premature coronary heart disease--a preliminary report from the Tecumseh Community Health Study. *Am. J. Epidemiol.* 1970; 91:233–237. [PubMed: 5414824]
22. Rissanen AM. Familial occurrence of coronary heart disease: effect of age at diagnosis. *Am. J. Cardiol.* 1979; 44:60–66. [PubMed: 453047]
23. Becker DM, Yook RM, Moy TF, Blumenthal RS, Becker LC. 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.* 1998; 82:1046–1051. [PubMed: 9817479]
24. Blumenthal RS, Becker DM, Yanek LR, Moy TF, Michos ED, Fishman EK, 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.* 2006; 97:328–333. [PubMed: 16442390]
25. Vaidya D, Yanek LR, Moy TF, Pearson TA, Becker LC, Becker DM. Incidence of coronary artery disease in siblings of patients with premature coronary artery disease: 10 years of follow-up. *Am. J. Cardiol.* 2007; 100:1410–1415. [PubMed: 17950799]
26. Blumenthal RS, Becker DM, Moy TF, Coresh J, Wilder LB, Becker LC. Exercise thallium tomography predicts future clinically manifest coronary heart disease in a high-risk asymptomatic population. *Circulation.* 1996; 93:915–923. [PubMed: 8598082]
27. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation.* 1998; 97:1837–1847. [PubMed: 9603539]
28. Smith LR, Harrell FE Jr, Rankin JS, Califf RM, Pryor DB, Muhlbaier LH, et al. Determinants of early versus late cardiac death in patients undergoing coronary artery bypass graft surgery. *Circulation.* 1991; 84(5 Suppl):III245–253. [PubMed: 1934415]
29. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 2007; 81:559–575. [PubMed: 17701901]
30. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics.* 1986; 42:121–130. [PubMed: 3719049]
31. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 2006; 38:904–909. [PubMed: 16862161]

32. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005; 21:263–265. [PubMed: 15297300]
33. Kamb A, Gruis NA, Weaver-Feldhaus J, Liu Q, Harshman K, Tavtigian SV, et al. A cell cycle regulator potentially involved in genesis of many tumor types. *Science*. 1994; 264:436–440. [PubMed: 8153634]
34. Pomerantz J, Schreiber-Agus N, Liegeois NJ, 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*. 1998; 92:713–723. [PubMed: 9529248]
35. Hannon GJ, Beach D. p15INK4B is a potential effector of TGF-beta-induced cell cycle arrest. *Nature*. 1994; 371:257–261. [PubMed: 8078588]
36. Pasmant E, Laurendeau I, Heron D, Vidaud M, Vidaud D, Bieche I. Characterization of a germline 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*. 2007; 67:3963–3969. [PubMed: 17440112]



**Figure 1. Permutation-based p-values per SNP for 77 SNPs having MAF > 4%**  
 Shaded regions represent the 100<sup>th</sup>, 99.9<sup>th</sup>, 99<sup>th</sup>, 95<sup>th</sup>, and 50<sup>th</sup> 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 cut-off for each SNP necessary to achieve significance after a Bonferroni 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 bolded SNP name. From Table 2, rs13298881 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 Helgadóttir et al 2007<sup>5</sup>; 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 noncoding 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).

**Table 1**

Baseline population characteristics of the GeneSTAR African American siblings (n=548)\*

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	3.52 ± 1.12
HDL cholesterol, mmol/L	1.40 ± 0.42
Triglycerides, mmol/L	1.37 ± 0.86
Systolic blood pressure, mmHg	135.3 ± 16.2
Fasting glucose, mmol/L	5.72 ± 2.35
Current smoking, (%)	32.1
BMI, kg/m <sup>2</sup>	31.3 ± 6.6

\* Continuous variables presented as mean ± 1 standard deviation

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

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

**Table 2**

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 *	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

\* rs13298881 (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)5 in GeneSTAR African American siblings; one pair (rs3217989 and rs1412831) has an  $r^2 > 0.7$

GeneSTAR SNPs	rs3217989	rs17761446	rs107572785
rs3217989	---	0.017	0.054
rs13298881	0.011	0.000	0.004
rs1412831	<b>0.743</b>	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

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript