Large-Scale Association Analysis Identifies 13 New Susceptibility Loci for Coronary Artery Disease

Nehal N. Mehta, MD, MSCE

Study Hypothesis
Recently, genome-wide association studies (GWAS) have identified several common variants that are associated with risk of coronary artery disease (CAD) and myocardial infarction (MI). The authors state that the current loci discovered in CAD and MI GWAS explain only a small fraction of the heritability of this complex disease. The authors hypothesized that a larger study would provide more power to discover common variants with modest effect sizes. Therefore, they formed the Coronary ARtery DIsease Genome-wide Replication And Meta-analysis (CARDioGRAM) consortium, which consisted of data from 14 GWAS of CAD and MI.

How Was the Hypothesis Tested?
The authors performed a meta-analysis of 14 GWAS of CAD comprising 22,233 cases and 64,762 control subjects, all of European ancestry. CAD was defined angiographically in a subset (n=7364) and by history in the entire sample. Presence of MI ranged from 48.1% to 100% of each cohort. After the discovery set (n=7364) and by history in the entire sample. Presence of MI ranged from 48.1% to 100% of each cohort. After the discovery set and a subset of previously reported CAD loci in up to 56,682 additional subjects (approximately half cases and half control subjects), Taking the number of loci into consideration, their replication study had 90% power to detect effect sizes observed in the GWAS meta-analysis. Finally, to understand potential mechanisms and intermediate pathways by which novel loci may mediate risk, the authors interrogated 3 genome-wide studies that also assessed gene expression in multiple tissues, using human cell lines, a genome-wide map of allelic expression imbalance, and other human disease traits.

Principal Findings
The analysis of approximately 135,000 individuals more than doubled the number of loci with CAD association, yielding 13 previously unidentified loci as well as 10 previously reported loci. CARDioGRAM also found that very few of the established and new loci appear to act through traditional risk factors. In fact, the majority of these loci reside in gene regions not previously thought to be involved in the pathogenesis of CAD. Finally, a substantial proportion of the CAD risk variants were also strongly associated with various other human disease traits in GWAS demonstrating a potential pleotropic effect of these gene regions.

Ten of the 12 loci (PCSK9 and SHB3 did not achieve genome-wide significance on initial meta-analysis) previously associated with CAD at a genome-wide significance level surpassed the same threshold of significance in CARDioGRAM in their initial meta-analysis (SORT1, MIA3, WDR12, MRAS, PHACTR1, LPA, 9p21, CXCL12, LDLR, MRPS6). Of note, at least 4 of these loci are related to low-density lipoprotein (LDL) cholesterol. Next, after meta-analysis, they selected 23 new loci with a significance level of P<5×10^-6 for follow-up genotyping. Of these 23 loci, replicated, using the a priori definition of a validated locus (showing independent replication after Bonferroni correction and achieving P<5×10^-8 in the combined discovery and replication data). These 13 loci (PPAP2B, ANKS1A, TCF21, ZC3HC1, ABO, CYP17A1/CNNM2/NT5C2, ZNF259/APOA5–A4–C3–A1, COL4A1/COL4A2, HHIPL1, ADAMTS7, SMG6/SRR, RASD1/SMC3/PEMT, and UBE2Z/GIP/ATP5G1/SNF8) had risk allele frequencies ranging from 0.13 to 0.91 and moderate effect sizes, with odds ratios ranging from 1.06 to 1.17, demonstrating the power of their study. Interestingly, given that the consortium had access to angiography, they performed various subgroup analyses starting with age of onset of CAD as well as ascertainment of CAD, which yielded the following findings: (1) 20 of 22 loci with P<5×10^-8 (known and new loci combined; for 1 locus, age subgroups were not available) had higher odds ratios for early-onset than for late-onset CAD (P=1.2×10^-4 for observed versus expected); (2) the odds ratios for most individual SNPs tended to be slightly greater for cases with angiographically proven CAD than for cases with unknown angiographic status (P=0.019 for observed versus expected); and (3) analyses in male and female subjects revealed no sex-specific effects for any risk alleles (P=0.4 for observed versus expected).
Next, to better understand the biology of these novel loci, the authors examined whether novel risk alleles were associated with traditional CAD risk factors, were coding variants, had functional significance, or were associated with other human disease traits. They found that 3 of the risk alleles on chromosome 14q23.3 (rs964184, ZNF259, APOA5-APOA4-APOC3-APOA1 gene region) were associated with increased LDL cholesterol and decreased high-density lipoprotein cholesterol, that the risk allele on chromosome 9q34.2 (rs579459, ABO) was associated with increased LDL and total cholesterol, and that variant rs12413409 on chromosome 10q24.32, representing the CYP17A1-CNNM2-NT5C2 gene region, was associated with hypertension.

Furthermore, the lead SNP at 4 of the 13 new loci were either nonsynonymous coding variants or were in high linkage disequilibrium (LD) with such SNPs. Variants at 7q32.2 (rs11556924) and 15q25.1 (rs3825807) encoded changes in ZC3H1 (p.Arg363His) and ADAMTS7 (p.Ser214Pro), respectively, whereas the lead SNP at 14q23.2 (rs2895811) was in strong LD ($r^2=0.82$) with the p.Val691Ala variant in HHIPL1. Last, the lead SNP at 17q21.32 (rs46522) is in strong LD ($r^2=0.94$) with 2 potential functional variants in GIP: p.Ser103Gly (rs2291725) and a variant influencing the splice site of intron 3 (rs2291726), leading to a truncated transcript.

In terms of expression, 3 of the new CAD risk variants showed convincing association with regional gene expression (cis effect) by either representing the most significant expressed SNP in the region or by being in high LD ($r^2=0.85$) with the strongest expressed SNP in the region: rs12190287 at 6q23.2 (TCF21), rs12936587 at 17p11.2 (RASD1, SMCR3, and PERT), and rs46522 at 17q21.32 (UBE2Z). They also interrogated new loci in a genome-wide map of allelic expression imbalance that provided further support for the expression quantitative trait locus findings at the 17q21.32 locus, which yielded strong evidence for cis effects for the 17p13.3 locus lead SNP (rs216172) on the expression of SMG6.

Finally, the authors identified 5 new loci [9q34 (ABO), 10q24 (CYP17A1), 11q23 (ZNF259/APO A5–A4–C3–A1), 15q25 (ADAMTS7), and 17p13 (SMG6, SRR)] where the CAD risk variant was fully or strongly correlated ($r^2>0.8$) with variants that have previously been associated with many other traits or diseases ranging from expected traits such as LDL cholesterol, high-density lipoprotein cholesterol, triglycerides to atherosperms, soluble levels of adhesion molecules, coagulation factor VIII, and von Willebrand factor (at $P<5\times10^{-8}$), demonstrating that a subset of the new CAD risk loci appear to have pleiotropic effects.

The authors state that their study focused on common risk variants, and, by assuming a heritability of 40% for CAD, the lead SNPs of previously established loci combined with the loci discovered in their study explain approximately 10% of the additive genetic variance of CAD, suggesting that many other common susceptibility variants of similar or lower effects and/or rare variants contribute to the risk of CAD.

**Implications**

This study represents how larger sample sizes can potentially reveal common susceptibility variants with small effect sizes in complex traits. Furthermore, this study is noteworthy in that it also demonstrates the importance of multiple parallel approaches to dissect the potential biological relevance of newly discovered loci in CAD. The authors used availability of risk factor endophenotypes, gene expression, and human disease traits to understand potential pathways and cell lines involved in the relationship between these loci and CAD to place the discovery of these loci in context. Finally, the authors demonstrate the importance of disease ascertainment to avoid potential misclassification of a measured outcome (in this case, angiographic CAD compared with CAD by history). Their analysis, restricted to angiographically proven CAD, showed stronger findings compared with analyzing CAD in those with a reported history. This also suggests that loci may differentially behave, depending on the definition and background phenotype of CAD.

Although the post hoc subgroup analysis in CARDIoGRAM suggested this nuance of a differential effect of novel loci on CAD, depending on the definition of phenotype, CARDIoGRAM was not designed to answer this question. This distinction of the phenotypic heterogeneity of CAD was recently demonstrated in a carefully designed study of CAD and MI in approximately 12 000 individuals, in which the authors found differential associations in those with angiographic CAD versus control subjects (ADAMTS7; $P=4.98\times10^{-13}$) compared with those with angiographic CAD who had MI versus those with angiographic CAD but no MI (ABO locus; $P=7.62\times10^{-8}$). These results, which complement findings from the CARDIoGRAM study, indicate that specific genetic predispositions may promote the development of coronary atherosclerosis, whereas others lead to MI in the presence of coronary atherosclerosis. We hope that more carefully designed studies such as this will improve understanding as to where novel loci fit into the spectrum of atherosclerosis and MI. These future studies should strive to improve phenotyping of CAD in large-scale studies as well as to use resequencing data to facilitate discovery of novel loci, which will continue to redefine the disease process of atherosclerosis and eventually inform novel therapeutic targeting.

**Sources of Funding**

Dr Mehta is supported by National Institutes of Health (NIH) research grant NHLBI K23 HL97151-2.

**Disclosures**

None.

**References**


Large-Scale Association Analysis Identifies 13 New Susceptibility Loci for Coronary Artery Disease
Nehal N. Mehta

_Circ Cardiovasc Genet._ 2011;4:327-329
doi: 10.1161/CIRCGENETICS.111.960443
_Circulation: Cardiovascular Genetics_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2011 American Heart Association, Inc. All rights reserved.
Print ISSN: 1942-325X. Online ISSN: 1942-3268

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/4/3/327

**Permissions:** Requests for permissions to reproduce figures, tables, or portions of articles originally published in _Circulation: Cardiovascular Genetics_ can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

**Reprints:** Information about reprints can be found online at:
http://www.lww.com/reprints

**Subscriptions:** Information about subscribing to _Circulation: Cardiovascular Genetics_ is online at:
http://circgenetics.ahajournals.org/subscriptions/