Published in final edited form as:
Circ Cardiovasc Genet. 2011 June 1; 4(3): 327–329. doi:10.1161/CIRCGENETICS.111.960443.

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

Nehal N. Mehta, MD, MSCE
Cardiovascular Institute, Center for Clinical Epidemiology and Biostatistics University of Pennsylvania, Philadelphia, PA

Keywords

genomics; genetics; coronary artery disease; genome wide association; metaanalysis

Study Hypothesis

Recently, genome-wide association studies (GWAS) have identified several common variants which 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 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 fourteen 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 controls, all of European ancestry. CAD was defined angiographically in a subset (n=7,364) and by history in the entire sample. Presence of MI ranged from 48.1% to 100% of each cohort. Following the meta-analysis, they then genotyped the lead SNPs within the most promising (defined *a priori* as $P < 5 \times 10^{-8}$) previously unidentified loci as well as a subset of previously reported CAD loci in up to 56,682 additional subjects (approximately half cases and half controls). 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 three 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.

Correspondence: Nehal N. Mehta, MD, MSCE, 6 Penn Tower, One Convention Center Blvd., Philadelphia, PA 19104, Phone: 215.662.7988, Fax: 215.349.5927, nehal.mehta@uphs.upenn.edu.
1 The CARDIoGRAM Consortium. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nature Genetics; 43:333–338.

Conflict of Interest Disclosures: None

Principal Findings

The analysis of approximately 135,000 individuals more than doubled the number of loci with CAD association, yielding 13 previously unidentified loci and confirming at least 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 four of these loci are related to low density lipoprotein (LDL) cholesterol. Next, following meta-analysis, they selected 23 new loci with a significance level of $P < 5 \times 10-6$ for follow up genotyping. Of these 23 loci, 13 replicated using the *a priori* definition of a validated locus (showing independent replication after Bonferroni correction and also achieving $P < 5 \times 10-8$ 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/SMCR3/PEMT, UBE2Z/GIP/ATP5G1/SNF8*) had had risk allele frequencies ranging from 0.13 to 0.91 and moderate effect sizes with odds ratios ranging from 1.06–1.17 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 out of 22 loci with $P < 5 \times 10^{-8}$ (known and new loci combined; for one locus, age subgroups were not available) had higher odds ratios for early onset than for late onset CAD ($P = 1.2 \times 10^{-4}$ 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); 3) analyses in males and females 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 associated with other human disease traits. They found that 3 of the risk allele on chromosome 11q23.3 (rs964184, *ZNF259, APOA5-APOA4-APOC3-APOA1* gene region) was associated with increased low-density lipoprotein (LDL) cholesterol and decreased high-density lipoprotein (HDL) 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 non-synonymous coding variants or were in high linkage disequilibrium (LD) with such SNPs. Variants at 7q32.2 (rs11556924) and 15q25.1 (rs3825807) encoded changes in *ZC3HC1* (p.Arg363His) and *ADAMTS7* (p.Ser214Pro), respectively, whereas the lead SNP at 14q32.2 (rs2895811) was in strong LD (r2 = 0.82) with the p.Val691Ala variant in *HHIPL1*. Lastly, the lead SNP at 17q21.32 (rs46522) is in strong LD (r2 = 0.94) with two 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 ($r2 \geq 0.85$) with the strongest expressed SNP in the region: rs12190287 at 6q23.2 (*TCF21*), rs12936587 at 17p11.2 (*RASD1, SMCR3* and *PEMT*) and rs46522 at 17q21.32 (*UBE2Z*). They also interrogated new loci in a genome-wide map of allelic expression imbalance which 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 ($r2 > 0.8$) with variants that have previously been associated with many other traits or diseases ranging from expected traits such as LDL cholesterol, HDL cholesterol, triglycerides to aneurysms, soluble levels of adhesion molecules, coagulation factor VIII and von Willebrand factor (at $P < 5 \times 10-8$) 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 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 utilized 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 indeed demonstrate the importance of disease ascertainment to avoid potential misclassification of a measured outcome (in this case, angiographic CAD compared to CAD by history). Their analysis restricted to angiographic proven CAD showed stronger findings compared to 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 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 2 where the authors found differential associations in those with angiographic CAD versus controls (*ADAMTS7*; $p=4.98\times10^{-13}$) compared to those with angiographic CAD who had myocardial infarction versus those with angiographic CAD but no myocardial infarction (*ABO* locus; $p=7.62\times10^{-9}$). 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 myocardial infarction in the presence of coronary atherosclerosis. Furthermore, more carefully designed studies such as this will hopefully 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 utilize re-sequencing data

to facilitate discovery of novel loci which will continue to redefine the disease process of atherosclerosis and eventually inform novel therapeutic targeting.

References

1. Schunkert H, Konig IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss M, Stewart AF, Barbalic M, Gieger C, Absher D, Aherrahrou Z, Allayee H, Altshuler D, Anand SS, Andersen K, Anderson JL, Ardissino D, Ball SG, Balmforth AJ, Barnes TA, Becker DM, Becker LC, Berger K, Bis JC, Boekholdt SM, Boerwinkle E, Braund PS, Brown MJ, Burnett MS, Buysschaert I, Carlquist JF, Chen L, Cichon S, Codd V, Davies RW, Dedoussis G, Dehghan A, Demissie S, Devaney JM, Diemert P, Do R, Doering A, Eifert S, Mokhtari NE, Ellis SG, Elosua R, Engert JC, Epstein SE, de Faire U, Fischer M, Folsom AR, Freyer J, Gigante B, Girelli D, Gretarsdottir S, Gudnason V, Gulcher JR, Halperin E, Hammond N, Hazen SL, Hofman A, Horne BD, Illig T, Iribarren C, Jones GT, Jukema JW, Kaiser MA, Kaplan LM, Kastelein JJ, Khaw KT, Knowles JW, Kolovou G, Kong A, Laaksonen R, Lambrechts D, Leander K, Lettre G, Li M, Lieb W, Loley C, Lotery AJ, Mannucci PM, Maouche S, Martinelli N, McKeown PP, Meisinger C, Meitinger T, Melander O, Merlini PA, Mooser V, Morgan T, Muhleisen TW, Muhlestein JB, Munzel T, Musunuru K, Nahrstaedt J, Nelson CP, Nothen MM, Olivieri O, Patel RS, Patterson CC, Peters A, Peyvandi F, Qu L, Quyyumi AA, Rader DJ, Rallidis LS, Rice C, Rosendaal FR, Rubin D, Salomaa V, Sampietro ML, Sandhu MS, Schadt E, Schafer A, Schillert A, Schreiber S, Schrezenmeir J, Schwartz Schwartz, Siscovick DS, Sivananthan M, Sivapalaratnam S, Smith A, Smith TB, Snoep JD, Soranzo N, Spertus JA, Stark K, Stirrups K, Stoll M, Tang WH, Tennstedt S, Thorgeirsson G, Thorleifsson G, Tomaszewski M, Uitterlinden AG, van Rij AM, Voight BF, Wareham NJ, Wells GA, Wichmann HE, Wild PS, Willenborg C, Witteman JC, Wright BJ, Ye S, Zeller T, Ziegler A, Cambien F, Goodall AH, Cupples LA, Quertermous T, Marz W, Hengstenberg C, Blankenberg S, Ouwehand WH, Hall AS, Deloukas P, Thompson JR, Stefansson K, Roberts R, Thorsteinsdottir U, O'Donnell CJ, McPherson R, Erdmann J, Samani NJ. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet. 2011; 43:333–338. [PubMed: 21378990]

2. Reilly MP, Li M, He J, Ferguson JF, Stylianou IM, Mehta NN, Burnett MS, Devaney JM, Knouff CW, Thompson JR, Horne BD, Stewart AF, Assimes TL, Wild PS, Allayee H, Nitschke PL, Patel RS, Martinelli N, Girelli D, Quyyumi AA, Anderson JL, Erdmann J, Hall AS, Schunkert H, Quertermous T, Blankenberg S, Hazen SL, Roberts R, Kathiresan S, Samani NJ, Epstein SE, Rader DJ. Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. Lancet. 2011; 377:383–392. [PubMed: 21239051]