JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY
VOL. 69, NO. 7, 2017

ISSN 0735-1097/$36.00
PUBLISHED BY ELSEVIER
http://dx.doi.org/10.1016/j.jacc.2016.12.007

EDITORIAL COMMENT

# Unraveling the Complex Genetics of Coronary Artery Disease*

Jason C. Kovacic, MD, PhD

Unraveling the full genetic basis of cardiometabolic diseases, such as impaired glucose metabolism, diabetes, obesity, dyslipidemias, and hypertension—which are among the key causes of stroke and coronary artery disease (CAD)—is proving to be one of the most profoundly complicated tasks facing contemporary biomedical research. In this issue of the *Journal*, the Myocardial Infarction Genetics and CARDIoGRAM Exome Consortia Investigators present the latest in a series of landmark studies that have added piecemeal to our understanding of the genetic basis of CAD (1).

SEE PAGE 823

The backdrop to this study was another recently published undertaking, where this well-known consortium assembled a large deoxyribonucleic acid sequence dataset from 72,868 CAD patients and 120,770 controls to explore and identify rare coding sequence variants associated with CAD (2). Notably, the consortium was successful in identifying low-frequency variants and several novel loss-of-function mutations in the genes *LPL, SVEP1,* and *ANGPTL4*, the latter being associated with altered triglyceride levels (2). In the current study in the *Journal*, the consortium leveraged the same dataset to look for novel common variants associated with CAD. As a point of distinction, in the former study, the consortium used analytic techniques to specifically study and identify variants with a minor allele frequency of <5% ("rare variants"); for the current study, the focus was variants with minor allele frequency >5% ("common variants"). Subdividing their dataset into discovery and replication cohorts, as well as replicating previously known CAD risk loci, the consortium identified and validated 6 new variants associated with CAD at genome-wide significance, being respectively associated with the genes *KCNJ13-GIGYF2*, *C2*, *MRVI1-CTR9*, *LRP1*, *SCARB1,* and *CETP*. In a consistent trend with prior studies, some of these new loci are associated with genes that might plausibly be related to CAD on the basis of prior knowledge: *LRP1* (low-density lipoprotein [LDL] receptor related protein-1), *SCARB1* (which encodes SR-B1, a receptor for high-density lipoprotein cholesterol), and *CETP* (cholesterol ester transfer protein). Conversely, and again consistent with prior studies, the genes associated with other new loci have no immediately apparent links to known major biological aspects of CAD, although the variant associated with *C2* (which encodes complement C2 protein) introduces the possibility that the complement system may have a more prominent role in atherosclerosis and CAD than currently appreciated. Previously, there were 56 validated CAD risk loci (3,4); adding the 6 new novel variants brings the total to 62 validated CAD risk loci. Of note, there are at least another 100 identified loci potentially associated with CAD that are yet to be validated (3–5). Furthermore, in another recent study, 17 additional novel CAD risk loci were identified (6).

The consortium then evaluated potential associations between the 62 validated CAD risk loci and both traditional CAD risk factors (lipid traits, blood pressure, body mass index, diabetes, and smoking) and a wide range of other diseases and traits, including diseases/traits as diverse as coronary artery calcification, stroke, lupus, and autism. These analyses

*Editorials published in the *Journal of the American College of Cardiology* reflect the views of the authors and do not necessarily represent the views of *JACC* or the American College of Cardiology.

From the Zena and Michael A. Wiener Cardiovascular Institute, Icahn School of Medicine at Mount Sinai, New York, New York. Dr. Kovacic has received research support from the National Institutes of Health (K08HL111330, R01HL130423), the American Heart Association (14SFRN20490315, 14SFRN20840000), and The Leducq Foundation (Transatlantic Network of Excellence Award).

JACC VOL. 69, NO. 7, 2017
FEBRUARY 21, 2017:837–40

were motivated by mounting evidence suggesting that single nucleotide polymorphisms (SNPs), such as the authors identified, exhibit substantial overlap across common complex disorders. Adding support to this paradigm, they found that 24 of 62 loci (38.7%) showed statistical association with traditional cardiovascular risk factors (most commonly with lipid traits), with some loci showing multiple associations. Additionally, almost one-half of the 62 SNPs (29 of 62; 47%) were associated with other diseases/traits.

This paper is an important undertaking and a notable advance in our knowledge of CAD-risk SNPs and the heritability of common complex disorders. Before considering the advances this study brings, let us first address its limitations. The study was performed by compiling 20 individual studies for the discovery cohort and a further 8 studies for validation. Although every effort was made to harmonize these datasets, it is inevitable that biases and inconsistencies, both known and unknown, were present across these studies. For example, among the 28 studies, the definition of a CAD "case" included such diverse designations as physician-assigned International Classification of Diseases codes, coronary stenosis ≥50%, abnormal stress test (that may include false positives), and fatal myocardial infarction (MI). This inhomogeneity in defining a CAD "case" adds uncertainty to this study, particularly because the biology of acute MI (plaque rupture and thrombosis) differs from that of stable CAD. Moreover, the inconsistent definition of CAD cases also affects the definition of the control group, where a significant but unknown portion may have had subclinical CAD (7).

As further limitations, the populations studied were overwhelmingly Caucasian Europeans, and the generalizability of these SNPs to other ethnicities is yet to be proven. In addition, by nature of the retrospective compilation of existing datasets, this study used a now superseded Illumina (San Diego, California) platform from circa 2011 that looked for variance at ~30,000 loci. Contemporary bead chips cover dramatically more loci, whereas deoxyribonucleic acid sequencing (that can disclose novel SNPs) is rapidly becoming the gold standard for gene discovery studies. In time, as these more powerful methods are used to interrogate large cohorts, we can expect the list of CAD risk loci to grow further.

Scientifically, what does this study add? Foremost, this study discloses the profound pleiotropy that exists not just among CAD-risk SNPs, but across also common complex disorders and traits in general. To be clear, what is meant by pleiotropy is that a single risk locus is associated with multiple different diseases and traits. Perhaps it might have been anticipated that a proportion of CAD-risk SNPs would also show association with cardiovascular risk factors (because certain SNPs may promote CAD by affecting the risk factor itself; e.g., by raising LDL cholesterol). Indeed, this was the case for 24 of 62 loci. However, what is particularly revealing about the biology of CAD and other complex disorders is that almost one-half (29 of 62) of the CAD-risk SNPs were also associated with other diseases or traits. What this tells us about the biology of CAD, common complex disorders, and the human genome is that an extremely complex balance exists among the heritability of different disease states.

Indeed, not only can a single SNP be associated with several different diseases, but adding even greater complexity—and by entirely unknown mechanisms—the risk association can be in a different direction for different diseases. For example, at the SNP rs9349379, which is associated with the gene *PHACTR1*, either an adenine (A) or guanine (G) may be present. Remarkably, regardless of which nucleotides are present (AA, AG, or GG), a disease association exists. Thus, an (A) at rs9349379 is associated with increased risk of cervical artery dissection (8), migraine headache (9), and fibromuscular dysplasia (10). Conversely, the (G) allele at rs9349379 is associated with CAD (3,11,12), coronary artery calcification (13), and MI (3,12,13). Therefore, the risk of having CAD or specific diseases is related, in certain cases, both directly and reciprocally with the risk of having other diseases.

Seminal studies from the 1990s suggested that heritability accounts for ~50% of the likelihood of developing CAD (14), with the remaining risk likely attributable to environmental- and lifestyle-related factors such as smoking, sedentary lifestyle, obesity, salt intake, diet, and other factors (**Figure 1**). At present, one of the great mysteries of common complex diseases is that of "missing heritability." That is, if we take all 62 validated CAD risk loci, even if we include the >100 nonvalidated loci, these directly account for only ~15% of the heritable likelihood of having clinically manifest CAD. On the one hand, the addition of 6 new SNPs from the current study to the already known or suspected CAD risk loci does little to increase our understanding of "missing heritability." On the other hand, the truly fascinating aspect of the current paper is perhaps not the 6 new loci, but the profound pleiotropy that was shown among complex diseases. Along with other recent studies like STARNET (Stockholm-Tartu Atherosclerosis Reverse Network Engineering Task) (11), these papers have opened up the possibility that heritable risk for CAD and common complex diseases may be partially attributable to interactions among diseases and traits. Some of these interactions