JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY
© 2016 BY THE AMERICAN COLLEGE OF CARDIOLOGY FOUNDATION
PUBLISHED BY ELSEVIER

VOL. 68, NO. 25, 2016
ISSN 0735-1097/$36.00
http://dx.doi.org/10.1016/j.jacc.2016.10.039

FOCUS SEMINAR: GENETICS

STATE-OF-THE-ART REVIEW

Genetics: Implications for Prevention and Management of Coronary Artery Disease

Themistocles L. Assimes, MD, PhD,[a] Robert Roberts, MD[b]

ABSTRACT

An exciting new era has dawned for the prevention and management of coronary artery disease (CAD) utilizing genetic risk variants. The recent identification of over 60 susceptibility loci for CAD confirms not only the importance of established risk factors, but also the existence of many novel causal pathways that are expected to improve our understanding of the genetic basis of CAD and facilitate the development of new therapeutic agents over time. Concurrently, Mendelian randomization studies have provided intriguing insights on the causal relationship between CAD-related traits, and highlight the potential benefits of long-term modifications of risk factors. Last, genetic risk scores of CAD may serve not only as prognostic, but also as predictive markers, and carry the potential to considerably improve the delivery of established prevention strategies. This review will summarize the evolution and discovery of genetic risk variants for CAD and their current and future clinical applications. (J Am Coll Cardiol 2016;68:2797–818)

Coronary artery disease (CAD), with its clinical sequelae of angina, myocardial infarction (MI), heart failure, and sudden death, has long been the number 1 cause of death in the developed world (1). Recently, it became the number 1 killer in the developing world, and so is now a global, sustained pandemic (1). This increased incidence is due to several factors, including increasing lifespans, changes in lifestyle, adoption of Western diets, and many others (1). Despite the global increased prevalence of CAD, its incidence in the developed world decreased substantially after the implementation of effective primary and secondary prevention strategies in the latter part of the 20th century (2,3). These strategies include the promotion of a healthy lifestyle anchored on optimal nutrition, smoking cessation, and increased physical activity, as well as the development of effective pharmacological therapies, such as platelet and 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase inhibitors. Collectively, these strategies are able to reduce the incidence of CAD by almost 50% in high-risk populations (2,3).

Improved application of established primary prevention strategies undoubtedly has the potential to further reduce the incidence of CAD (4,5). However, many individuals have substantial residual risk, remain reluctant to take medications on the basis of assessments of risk with current risk calculators, and/or are unable to institute and maintain the lifestyle changes that are necessary to substantially reduce risk (6–9). In this context, recent impressive progress in genomic medicine, guided by astonishing breakthroughs in laboratory technology and computing power, provides us with a golden opportunity to dissect the genetic basis of CAD. This knowledge, in turn, is being leveraged to improve our ability to identify subjects who are at high risk and to develop

Listen to this manuscript's audio summary by *JACC* Editor-in-Chief Dr. Valentin Fuster.

From the [a]Department of Medicine, Stanford University School of Medicine, Stanford, California; and the [b]International Society for Cardiovascular Translational Research, University of Arizona College of Medicine, Phoenix, Arizona. Dr. Assimes has received grants from Telomere Diagnostics and Regeneron Genetics Center. Dr. Roberts is a Scientific Advisor to Cumberland Pharmaceuticals.

Manuscript received August 15, 2016; revised manuscript received October 12, 2016, accepted October 24, 2016.

ABBREVIATIONS AND ACRONYMS

CAD = coronary artery disease

GRS = genetic risk score

GWAS = genome-wide association study

HDL = high-density lipoprotein

LDL = low-density lipoprotein

MI = myocardial infarction

MR = Mendelian randomization

SNP = single-nucleotide polymorphism

WES = whole-exome sequencing

WGS = whole-genome sequencing

novel therapies. The **Central Illustration** summarizes progress to date, as well as the topics covered in this review. The pharmacogenetics of drugs used to treat or prevent CAD, including oral antithrombotic agents and lipid-lowering drugs, is of substantial interest, but is reviewed in detail elsewhere (10,11).

ORIGIN OF HUMAN DEOXYRIBONUCLEIC ACID VARIATION RESPONSIBLE FOR PREDISPOSITION TO DISEASE

The human genome possesses 3.2 billion base pairs, of which approximately 99.5% are identical between any 2 individuals (12,13). The deoxyribonucleic acid (DNA) sequence defining the features unique to each human is contained in approximately 0.5% of the sequence (~15 million bases). Most of these 15 million base differences consist of large segments of DNA that are rearranged, or repetitive sequences that vary in the number of copies per genome (copy number variation) (13). However, the most common type of mutation contributing to the genome's uniqueness is a substitution of one base (adenine, guanine, thymine, or cytosine) for another, commonly referred to as a single nucleotide polymorphism (SNP) (12,13).

SNPs are due to DNA copying errors during replication of the 2 strands of DNA. DNA turns over every few days, and the base pairs are replaced appropriately and accurately, with only 1 error per billion base pairs created (mutation rate of 1.4×10^{-8}) (14,15). This error rate is enough to induce a significant number of mutations, which may be transmitted to the next generation if they occur in the germ line (egg or sperm) (16). SNPs are distributed throughout the genome, and their effect on the phenotype may be neutral, benign, or markedly influential. The number of SNPs is fairly constant, at about 3 million per human genome, and these variations are responsible for a sizable fraction of the variation in physiological traits, including predisposition to disease. However, SNPs are constantly evolving with each new generation of humans. Given a mutation rate of 1.0×10^{-8}, it is estimated that 40 to 60 new nonembryonic lethal mutations are introduced per generation. With the world's population of 7 billion, over 300 billion new mutations exist in the current generation alone (17).

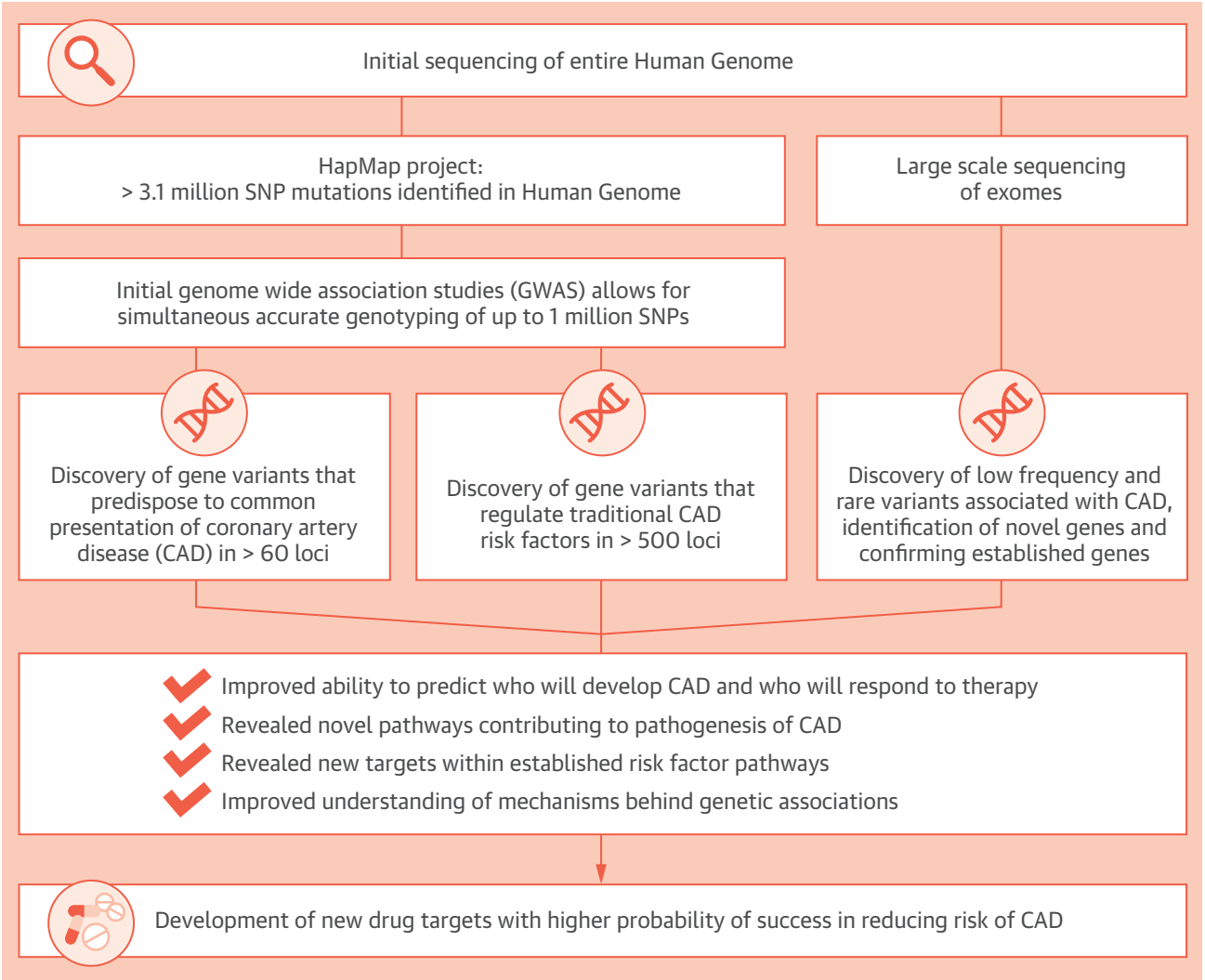
New mutations that are randomly assigned to each generation are initially rare, and their frequency in subsequent genomes will evolve according to Darwinian principles. SNPs that provide an advantage to or have no effect on survival and fitness may increase in frequency, whereas those that are detrimental will remain rare or disappear (18). In this context, genetic determinants of most common chronic diseases, including CAD, have not been under substantial selective pressure, given that these conditions have not been primary causes of morbidity and mortality until only recently (18). Even as CAD emerged as a leading cause of death in the last century, it generally still does not cause death before the end of the reproductive phase of life in humans.

FAMILY STUDIES SUGGEST THAT COMMON PRESENTATIONS OF CAD ARE MODERATELY HERITABLE

Several epidemiology studies involving unrelated individuals provided the first clues that non-Mendelian common presentations of CAD in middle to late adulthood were heritable (19). These familial aggregation studies estimated the contribution of genetic factors by comparing rates of CAD among subjects with a positive family history of CAD to rates among subjects with no such history. Overall, these studies have documented ~2.5- to 4-fold higher rates of CAD among individuals with a family history compared with individuals with no family history when adjusting only for age and sex (19–21). When further controlling for traditional risk factors, the excess risk of CAD is reduced to about 1.5- to 2.5-fold. This is consistent with the notion that at least part of this excess risk is a consequence of the familial aggregation of traditional risk factors (19–21). Familial aggregation studies also provide compelling evidence that a stronger family history or an earlier age of onset of disease in a family member further increases the risk for close relatives (19–21).

A weakness of familial aggregation studies is rooted in the possibility that a positive family history may serve as a proxy of exposure to common familial environmental risk factors (e.g., pollution and nongenetic influences, such as smoking, diet, and physical activity) in addition to genetic factors (22). Such confounding may result in an overestimate of the contribution of genetic factors. To address this concern, several twin studies have been conducted, which offer the advantage of nullifying the effect of a common familial environment by comparing concordance rates of disease among monozygotic twins to those in dizygotic twins. Difference in concordance rates are then transformed to unbiased estimates of heritability, assuming that the extent to which common familial environmental risk factors are shared

CENTRAL ILLUSTRATION **Genetics of CAD: Prevention and Management**



Assimes, T.L. et al. J Am Coll Cardiol. 2016;68(25):2797–818.

The figure summarizes recent progress and future directions in identifying the genetic determinants of CAD as well as the application of this knowledge to the management and prevention of this disease. CAD = coronary artery disease; GWAS = genome wide association studies; PCSK9 = proprotein convertase subtilisin/kexin type 9; SNP = single-nucleotide polymorphism.

between either set of twins is equal. Arguably, the most informative twin studies conducted to date for CAD have involved registries in Denmark and Sweden, which have estimated the heritability of CAD to be in the range of 40% to 60% (23–25). Consistent with the familial aggregation studies, the twin studies have also observed higher heritability when disease occurs at a younger age, and a partial attenuation of estimates when they are adjusted for traditional risk factors (23–25).

LINKAGE STUDIES FAIL TO IDENTIFY GENETIC DETERMINANTS OF COMMON PRESENTATIONS OF CAD

In the 1980s, the identification and mapping of several hundred highly polymorphic DNA markers that are equally spaced throughout the genome (referred to as *microsatellites*) enabled scientists to identify a region of the genome harboring highly penetrant mutations responsible for rare monogenic

disorders through linkage analysis. Such work was possible even before the sequencing of the entire human genome was complete. In such studies, microsatellites are typed in families with at least 1 affected member, and analyses are then performed to identify which markers within the family members segregate best with disease status. Subsequent DNA sequencing around the region of linkage is then used to refine the signal and to ultimately identify the causal variant.

Linkage analysis has proven to be a very effective means of identifying highly penetrant variants responsible for disease in large multigenerational families that include many affected individuals with a Mendelian or Mendelian-like mode of inheritance. For example, this approach was used to identify mutations in the beta-myosin heavy chain at the 14q1 chromosomal locus as a cause of familial hypertrophic cardiomyopathy (26–28) in the early 1990s. However, such families are not common and/or not easy to find for CAD. For many of these families identified to date, the culprit mutation has proved to be a highly damaging mutation within a gene that strongly influences circulating low-density lipoprotein (LDL) cholesterol levels (e.g., familial hypercholesterolemia [FH] genes, including *LDLR*, *APOB*, *PCSK9,* and, less commonly, *LDLRAP1*, *ABCG5*, and *ABCG8*) (29). In more recent studies, the culprit mutation has been found in genes predisposing to severe forms of the metabolic syndrome (e.g., *LRP6* and *DYRK1B*) (30,31). The initial observation that supported the highly multigenic nature of CAD for a majority of subjects presenting with CAD was the failure of linkage studies performed on families with a more typical low penetrance of disease to detect any susceptibly loci (29,32).

The limited power of linkage studies to detect variants with low to modest effects for multigene disorders became glaringly clear by the mid-1990s (32,33). Researchers also recognized that association studies focused on directly comparing the frequency of variants among affected and unaffected subjects were much better powered than linkage studies to identify polygenic determinants of common disorders, such as CAD (32). All that was needed to institute such an approach was a complete catalog of human genetic variants.

THE 21ST CENTURY USHERS IN GENETICS OF MULTIGENE DISORDERS

COMPLETION OF THE HUMAN GENOME PROJECT AND FAILURE OF EARLY CANDIDATE GENE ASSOCIATION STUDIES. The initial sequencing of the human genome in the late 1990s and early 2000s resulted in the discovery of approximately 1.5 million SNPs (34–36). Scientists jumped on this resource to perform small-scale association studies limited to variants within or near genes previously suspected of being involved in the pathogenesis of disease (37). Unfortunately, this biased candidate gene approach was unsuccessful, as many positive initial reports of association could not be replicated (37). The reasons for the failure of candidate genes studies are many, but are rooted in the biased and premature approach of examining only a subset of SNPs within a subset of genes in search of what ultimately proved to be very modest effects, using relatively small sample sizes (37). As a result, many published studies were woefully underpowered and carried a high rate of false positive findings (38,39).

GENOME-WIDE ASSOCIATION STUDIES

Whole-genome sequencing of more humans by various groups led to the cataloging of several additional millions of SNPs in the first one-half of the 2000s. Between 2002 and 2007, the International HapMap Consortium subsequently carefully documented the allele frequencies of approximately 3.1 million of these SNPs in 270 subjects, representing 4 races/ethnic groups (40,41). This rich resource provided the DNA markers necessary to span the entire genome and set the stage for initial genome-wide association studies (GWAS) using newly developed, cost-efficient array technology, which allowed simultaneous accurate genotyping of up to 1 million SNPs (42,43).

A GWAS represents an unbiased approach to detect genetic variants associated with clinical outcomes (43,44). In essence, it extends the candidate gene association study approach not only to variants within all genes, but also to variants between genes that may be regulating the expression of genes nearby (**Figure 1**). Similar to candidate gene association studies, GWAS compares the frequency of each genetic variant among subjects with a disease of interest to the frequency among control subjects. Variants with frequencies that statistically differ between the 2 groups are assumed to either be pathogenic or to be highly correlated with the pathogenic variant(s) in the region (43,44).

The overwhelming failure of candidate gene association studies motivated the scientific community of genetic epidemiologists, statistical geneticists, and journal editors to adopt strict criteria for establishing novel findings from GWAS (43,45). Testing 1 million SNPs simultaneously can be problematic with respect

to traditional thresholds of significance, as about 50,000 of the tests would be expected to have a $p < 0.05$ by chance. Thus, to account for the multitude of tests, a Bonferroni correction to the p value threshold of significance was embraced at 5×10^{-8}, a threshold commonly referred to as the *genome-wide significance threshold* (33). Ideally, for a given SNP, both a discovery and a replication study would demonstrate similar magnitude of effect in the same direction, with a joint or combined p value leading to a smaller p value than that seen in the initial report (45). Additional criteria suggested for establishing a positive replication in GWAS include examining the same or a very similar population and phenotype in the replication study, using the same genetic model in both the discovery and the replication datasets, and providing a strong rationale for the SNPs selected for replication from the initial study (45).

DISCOVERY OF THE FIRST LOCUS PREDISPOSING TO COMMON PRESENTATIONS OF CAD

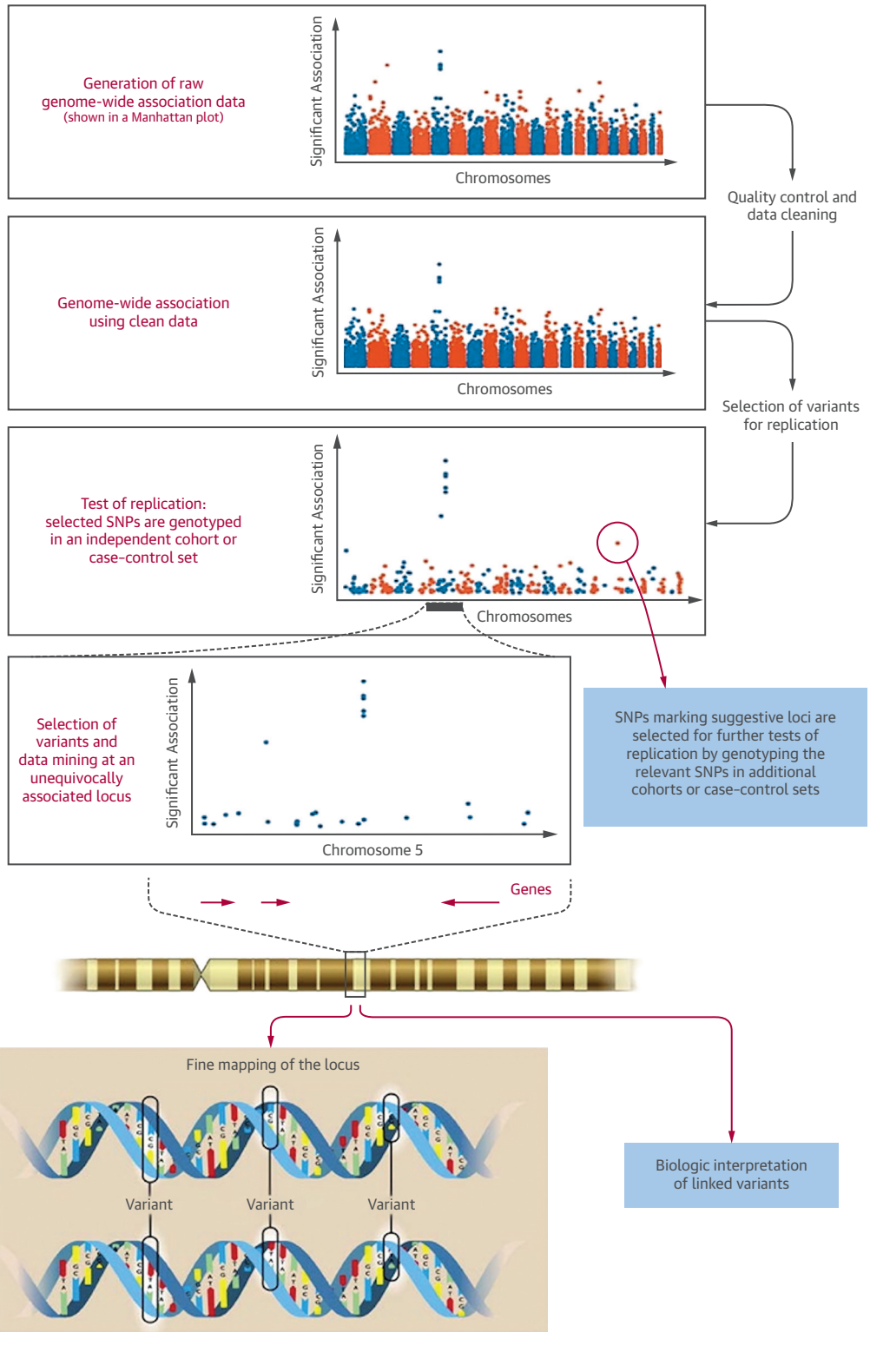
In 2007, 3 groups independently reported the first region of the genome to be linked to CAD utilizing the GWAS approach (46–48). All 3 groups made the discovery in European populations (**Table 1**). The susceptibility locus is located on the short arm (p) of chromosome 9 at band 2.1, and thus is commonly referred to as 9p21. The locus includes approximately 60 strongly correlated SNPs over ~53,000 base pairs, and is approximately 100,000 to 150,000 base pairs upstream of the genes encoding 2 cyclin-dependent kinase (CDK) inhibitors and known tumor suppressors, *CDKN2B* and *CDKN2A* (46–48). The locus also overlaps the last section of a long noncoding ribonucleic acid, which is transcribed antisense to CDKN2B (CDKN2BAS or ANRIL) (46–48) (**Figure 2**). The lead variants at 9p21 are common, with a minor allele frequency of about 50% in Europeans, resulting in about 75% of individuals of European ancestry carrying at least 1 allele that raises risk. The initial increase in risk for CAD observed was about 25% for 1 copy and about 50% for 2 copies, with somewhat higher risks per allele observed among those with early-onset CAD (46,47). The risk mediated by 9p21 appeared from the outset to be completely independent of all known risk factors, implying that other unknown factors are contributing substantially to the pathogenesis of CAD (46,47). The findings at 9p21 were quickly replicated in many other cohorts of Europeans and East Asians (49,50). Somewhat surprisingly, robust replication in African Americans remains elusive to this day, although some signals in the region have been detected (51–53). Replication studies have also demonstrated that the true excess risk per risk allele is likely closer to 20% in Europeans (49,54).

The exact mechanism of action linking genetic variation at 9p21 to the risk of CAD remains unclear to this day, although several sets of observations at the population level and in the laboratory have helped to narrow down the possibilities (55). First, the exact same variants have been linked to complications of atherosclerosis outside of the heart, including carotid plaque, ischemic stroke, and peripheral arterial disease, suggesting that the locus predisposes to atherosclerosis in all vascular beds (55). Second, the exact same variants have been linked to both abdominal aortic and intracranial arterial aneurysms, suggesting that the cells that are affected by these variants reside in the vessel wall of the artery and play some role in maintaining its integrity (55). Third, a growing body of evidence suggests that the SNPs in the high-risk region disrupt or create transcription factor binding sites that alter the expression levels, or the relative abundance of different transcripts of the noncoding ribonucleic acid, ANRIL, which in turn affects the expression levels of *CDKN2B* and/or *2A*. The protein products of these 2 genes, $p15^{INK4a}$ and $p16^{INK4a}$, then alter the function of resident macrophages and/or vascular smooth muscle cells, facilitating the formation of atherosclerotic plaque (55). Last, animal model studies to date suggest that these effects could involve increased proliferation and reduced apoptosis of resident macrophages and/or vascular smooth muscle cells (55).

FURTHER DISCOVERY OF SUSCEPTIBILITY LOCI FOR CAD THROUGH THE FORMATION OF MULTIPLE INTERNATIONAL CONSORTIA AND META-ANALYSIS OF GWAS

A total of 11 other loci for CAD were identified within a very short interval of the discovery of 9p21 by several groups, including the Welcome Trust Case-Control Consortium, the Cardiogenics Consortium, and the Myocardial Infarction Genetics Consortium (56–60). The result of these initial GWAS studies confirmed that common susceptibility variants for CAD carried minimal incremental risk, and would require very large sample sizes, in the tens of thousands to the hundreds of thousands, to be uncovered. Larger national and international consortia were formed to tackle this challenge, including the CARDIoGRAM (Coronary Artery Disease Genome Wide Replication and Meta-analysis) and the Coronary Artery Disease (C4D) Genetics consortia, which in 2011 confirmed

FIGURE 1 Stages of a GWAS



Continued on the next page

almost all of the previously reported loci and uncovered 17 new loci through meta-analysis (61,62). Through further collaboration between these 2 consortia, and the testing of a larger fraction of SNPs with a minor allele frequency >1% using advanced imputation algorithms, CARDIoGRAMplusC4D reported an additional 15 novel loci in 2013 and 8 loci in 2015, after the examination of >60,000 cases and >120,000 control subjects (54,63). These studies brought the total loci having reached genome-wide significance in largely European and, to a lesser extent, South Asian populations to 58 (54,63).

East Asians represent the next most studied race/ethnic group. A total of 12 loci have convincingly reached genome-wide significance for CAD in studies involving either Han Chinese, Korean, or Japanese subjects, with at least 6 of these loci overlapping with loci uncovered in Europeans (64–68). Thus, both ethnic-specific as well as cross-ethnic loci have been uncovered to date. The overlap in susceptibility loci across race/ethnic groups is expected to increase as larger studies are conducted in non-European populations, given that the genetic basis of most common diseases is believed to be identical across race/ethnic groups (69–71).

The decreasing cost of genotyping over time also allowed the IBC 50k CAD Consortium to revisit the candidate gene approach using a customized chip that comprehensively surveyed common genetic variation at ~2,100 candidate genes, representing a diverse spectrum of cardiovascular diseases (72). Well-powered studies using this array firmly established associations between CAD and genetic variants in *LPA*, the gene encoding lipoprotein(a), and confirmed several early GWAS discoveries for CAD (9p21, *COL4A1/COL4A2*, *ZC3HC1*, and *CYP17A1*) (73,74). However, these studies also confirmed a lack of association for a vast majority of variants within candidate genes related to CAD (73,74).

MANY SUSCEPTIBILITY LOCI FOR CONVENTIONAL RISK FACTORS FOR CAD HAVE ALSO BEEN IDENTIFIED THROUGH GWAS

The genetics of several known risk factors for CAD, including body mass index, blood lipid levels, blood pressure (BP), diabetes, insulin resistance, and smoking, are also being explored in parallel to the CAD outcome. In total, several hundred susceptibility loci have been identified for these traits, with many correlating with more than 1 trait (75–79). The genetic risk variants involved with these conventional risk factors may have indirect or direct effects on the risk of CAD. Indeed, as highlighted above for the 9p21 locus, the pattern of association between susceptibility SNPs and multiple traits may provide important clues on the mechanism of association (80).

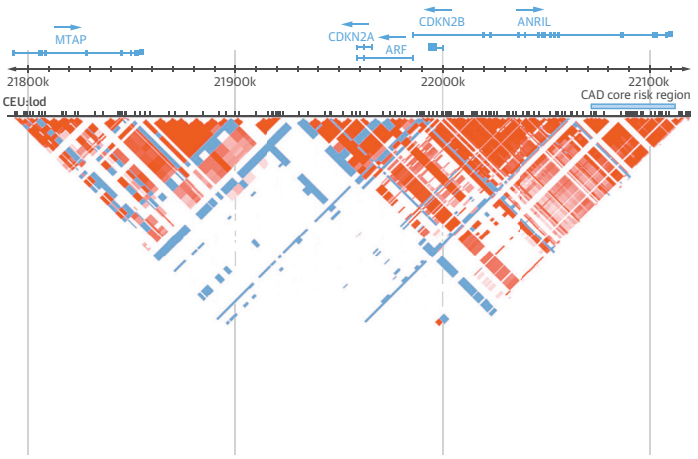
FEATURES OF GENETIC RISK VARIANTS FOR CAD

Table 1 lists regions of the genome where at least 1 genetic risk variant has reached genome-wide significance for association with CAD to date since 2007.

FIGURE 1 Continued

In the first stage, SNPs across the genome are genotyped, almost exclusively on chip-based products generated by 1 of 2 companies, Illumina (San Diego, California) or Affymetrix (Santa Clara, California). The genotyping content of these products differs, but can be combined through imputation algorithms that allow for the reliable calling of thousands of untyped SNPs on the basis of the pattern of variation observed in SNPs genotyped in the vicinity. Second, the SNP data are subjected to quality control and cleaning procedures to allow for the removal of samples from ethnically distant subjects and adjustment for any systematic differences between or within cohorts. Third, each SNP that survives quality control and cleaning is then tested for association with a disease or trait. Shown is a Manhattan plot, which is typically used in genome-wide association studies (GWAS) and plots the negative log of the p value against chromosomal position. A high false-positive rate is expected given the number of statistical tests. Therefore, genome-wide statistical significance is often set at $p < 5.0 \times 10^{-8}$ or less at this stage of the analysis. The additive model is the most frequently tested model, which assumes that the presence of 1 risk allele confers an intermediate risk between having no allele and having 2 alleles. Fourth, SNPs or loci are selected for replication in an independent sample set, ideally of the same or larger size than the sample analyzed in the genome-wide association. The selection of loci may be on the basis of statistical significance alone or a combination of statistical significance and biological plausibility. Fifth, replication experiments lead to any combination of 3 results: selected loci show clear and unequivocal association with disease, show no association signal, or show an association with disease that is not of sufficient magnitude to pass a pre-determined statistical threshold. Sixth, genotyping is performed in independent replication cohorts to determine whether an association with a disease is real. Seventh, data mining at unequivocally associated loci reveals transcripts in and around a locus, in addition to the mapping of all known genetic variation within the region. Further *fine mapping* of the locus is performed by a combination of deep-resequencing methods to discover new variants and the genotyping of untyped variants to determine which are most significantly associated with disease. The examination of association patterns in the region in other race/ethnic groups may help narrow the high-risk region or pinpoint a causal variant. Further analysis of the region is performed to determine the most critical variants, the pathologically relevant gene, and the likely biological effect. SNP = single-nucleotide polymorphism. Reproduced with permission from Hardy and Singleton (44). Copyright 2009 Massachusetts Medical Society.

FIGURE 2 Genes and Genetic Variants Around the CAD Core Risk Region on 9p21



The figure illustrates the blocks of correlated variants in Europeans (CEU). **Numbers on the scale bar** represent the chromosome 9 location (kb = kilobases). **Shading** indicates correlation (or linkage disequilibrium [LD]), ranging from D′ = 1 in **orange** (strong correlation) to D′ = 0 in **white** (no correlation). Genes are shown above the bar in **blue**, with **arrows** indicating the direction of transcription. The core risk region overlaps the last section of CDKN2BAS, a long noncoding ribonucleic acid, which is transcribed antisense to CDKN2B. CAD = coronary artery disease. Reprinted with permission from Cunnington and Keavney (174).

To minimize inflation of effects from "winner's curse," the table is restricted to the results of the most significant SNP in the region, as determined by the latest GWAS meta-analysis conducted by CARDIoGRAMplusC4D (54).

Table 1 highlights several features of these variants. First, the variants are common, have small effects, and are generally not located within protein-coding regions of the genome. These findings support the common variant-common disease hypothesis put forth many years ago by leading geneticists, which predicted that common disease-causing polymorphisms found in all major populations of the world are predominantly responsible for complex polygenic diseases through subtle effects on regulatory regions of the genome (81). The findings also highlight the lack of power and the accompanying high risk of false positive associations inherent in the candidate gene studies that preceded GWAS, which typically involved no more than a few hundred cases (37–39). Second, only about one-third of the variants appear to be influencing the risk of CAD through effects on traditional risk factors, namely, lipids and BP (63). This estimate was derived from cross-referencing the GWAS signals for CAD with those observed for risk factors (63). Thus, the mechanisms whereby most of these risk variants mediate risk remain unknown.

The small effect sizes observed with common variants should not be interpreted as pointing to genes or pathways with little clinical relevance. Although such effect sizes make risk prediction more challenging, they may still involve genes or pathways that have the potential to dramatically affect risk of disease through more extreme modulation of gene function or expression. For example, the initial GWAS signals for CAD include common variants with modest effects in or near all 3 of the FH genes (*LDLR, APOB,* and *PCSK9*). Collectively, these loci point to the cholesterol synthesis pathway within the hepatocyte, a pathway that has already been very effectively targeted by statins and proprotein convertase subtilisin/kexin type 9 (PCSK9) inhibitors (82,83). Thus, a better understanding of the causal genes and pathways responsible for GWAS findings to date for CAD should not only provide insight into the pathogenesis of coronary atherosclerosis, but also facilitate the development of effective novel therapeutic agents.

GENETIC RISK VARIANTS FOR CAD APPEAR TO PRIMARILY PROMOTE THE FORMATION OF CORONARY ATHEROSCLEROSIS

In the pursuit of genetic risk for coronary atherosclerosis, CAD and MI have been treated as the same phenotype. This approach is justified on the basis that MI, with rare exceptions, occurs secondary to significant coronary atherosclerosis (84,85). The pathophysiology of coronary atherosclerosis involves many factors, including endothelial dysfunction, monocyte adhesion, cholesterol deposition, and an inflammatory response with release of cytokines (84,85). The genetic predisposition leading to an MI would be through different targets and mechanisms than that of atherosclerosis, including mechanisms associated with plaque rupture and thrombosis (84,85).

A major limitation in determining whether a susceptibility locus for CAD predisposes specifically to coronary atherosclerosis versus MI is our inability to accurately quantify the burden of coronary atherosclerosis, as well as the number of prior symptomatic and silent plaque ruptures among individuals with coronary atherosclerosis (86–88). One way to partially circumvent this difficulty is to compare subjects with prior MIs to subjects with no prior MIs among subjects with documented clinically significant CAD. Such a design was leveraged to identify the *ABO* locus as an MI-predisposing locus, with the A and B risk variants shown to increase the risk of MI in the setting of CAD by ~20% (89). The protein coded by the *ABO* locus, α-1-3-N-acetylgalactoseaminyltransferase, is known to transfer a carbohydrate onto von Willebrand Factor

(vWF), prolonging its half-life and leading to levels in the plasma that are 25% higher in individuals with A, B, or AB blood groups compared with those with blood group O (90). The blood group O gene also codes for the same transferase protein, but was mutated a long time ago, resulting in loss of any biochemical activity. Individuals with blood group O do not have increased plasma levels of vWF and have a lower risk for MI. Thus, the biological plausibility of this association is reasonable, given that plasma vWF levels have been repeatedly shown to predict the occurrence of adverse cardiac events, including death, among patients with pre-existing vascular disease (91). However, the *ABO* locus has also been found to associate with LDL levels, as well as with CAD outcomes in conventional case-control studies, making it difficult to conclude with certainty whether the locus predisposes to MI independent of its ability to promote atherosclerosis (62,89).

Examination of correlation patterns between established variants for CAD and the burden of atherosclerosis on coronary angiography may provide some additional insight on mechanisms of association. Multiple studies have shown that genetic variants with CAD susceptibility loci correlate with burden of disease on angiography (92–94). However, such correlations could still be driven by a genetic predisposition to subclinical extraplaque or intraplaque ruptures and/or erosions of plaques, which have been previously shown to be responsible for the nonuniform growth of many moderate and high-grade lesions used to quantify an individual's burden of disease on angiography (88,95).

The study of the earliest lesions of coronary atherosclerosis not yet prone to plaque rupture or erosion represents another means to help distinguish between the 2 potential mechanisms of association. Salfati et al. (96) recently took advantage of the PDAY (Pathobiological Determinants of Atherosclerosis in Youth) resource to document associations between a genetic risk score (GRS) of high-risk alleles for clinical CAD and the degree of subclinical coronary atherosclerotic lesions documented through autopsy in young adults who died between 15 and 35 years of age from noncardiovascular causes. They found that the magnitude of association between the GRS and early lesions documented in PDAY was comparable to the association observed with coronary artery calcification in older subjects, and that these associations persisted, even when risk factor SNPs were removed from the GRS (96). Overall, these findings support the hypothesis that a majority of newly identified GWAS loci for CAD are atherosclerosis predisposing loci, rather than loci that predispose to plaque rupture or thrombosis (96). However, the PDAY resource is too small to allow for definitive conclusions regarding associations on an SNP by SNP basis. Thus, more studies using this approach are necessary before definitive conclusions can be made for each CAD susceptibility locus.

DETERMINING CAUSALITY OF RISK FACTORS FOR CAD THROUGH MENDELIAN RANDOMIZATION

A pervasive challenge in observational epidemiology is determining whether a link between a risk factor and an outcome is causal (97). Although a noncausal risk factor may still serve as a useful predictor, an intervention developed specifically to modify the degree of exposure to such a risk factor is not likely to succeed in improving outcomes (97).

The most reliable way to assess a causal association between a risk factor and an outcome has been through randomized controlled trials (RCTs) of interventions that modify risk factor levels (97). In some instances, such interventions have demonstrated the limitations of statistical adjustment of measured confounders in observational studies. Notable examples include a plethora of studies on vitamin levels and cardiovascular health, whose benefits in the observational setting could never be replicated through corresponding RCTs (98).

The advent of large-scale genotyping of common variants over the last 15 years has facilitated the development of Mendelian randomization (MR) studies (**Figure 3**) as an alternate method of assessing causality between a risk factor and an outcome in an observational setting (97). MR studies leverage the well-established method of instrumental variable (IV) analysis, used successfully in econometrics for many years to deal with issues of confounding, reverse causality, and regression dilution (97). In an analogous fashion, MR studies use 1 or more genetic variants associated with the risk factor of interest as the instrument to achieve the same goal in medicine (97) (**Figure 3**). A key prerequisite for successful MR studies is the identification of a subset of genetic variants that have been randomly allocated at conception, have been found to robustly influence the exposure levels of a risk factor, and are not linked to the outcome in any way other than through effects on the risk factor (97) (**Figure 3**).

MR studies have provided some intriguing insights on the causality of a variety of cardiometabolic risk factors and, in some cases, have shattered previously widely held beliefs (99). For example, MR studies to date suggest that homocysteine, fibrinogen, secretory

TABLE 1 Susceptibility Loci Reaching Genome-Wide Significance for CAD*

Lead SNP†	Chromosome	Nearest Genes‡	Frequency of Allele Raising Risk	OR (95% CI)	p Value§	Potential Mechanism of Action‖	Year Locus First Reported to Reach Genome-Wide Significance	Consortium and/or Author (Ref. #)
rs11206510	1	*PCSK9*	0.85	1.08 (1.05–1.11)	2.340×10^{-8}	LDL levels	2009	MIGen (59)
rs7528419	1	*SORT1*	0.79	1.12 (1.10–1.15)	1.970×10^{-23}	LDL levels	2007	WTCCC and Cardiogenics (56)
rs515135	2	*APOB*	0.79	1.07 (1.04–1.10)	3.090×10^{-8}	LDL levels	2013	CARDIoGRAM+C4D (63)
rs6544713	2	*ABCG5–ABCG8*	0.32	1.05 (1.03–1.07)	8.880×10^{-7}	LDL levels	2013	CARDIoGRAM+C4D (63)
rs56289821	19	*LDLR*	0.90	1.14 (1.11–1.18)	4.440×10^{-15}	LDL levels	2009	MIGen (59)
rs4420638	19	*APOE–APOC1*	0.17	1.10 (1.07–1.13)	7.070×10^{-11}	LDL levels	2013	CARDIoGRAM+C4D (63)
rs3184504	12	*SH2B3*	0.42	1.07 (1.04–1.09)	1.030×10^{-9}	LDL levels, BP	2009	deCODE (58)
rs55730499	6	*SLC22A3–LPAL2–LPA*	0.06	1.37 (1.31–1.44)	5.390×10^{-39}	Lp(a) levels	2009	WTCCC and Cardiogenics (60)
rs264	8	*LPL*	0.85	1.06 (1.03–1.09)	1.060×10^{-5}	TRIG levels	2013	CARDIoGRAM+C4D (63)
rs2954029	8	*TRIB1*	0.55	1.04 (1.03–1.06)	2.610×10^{-6}	TRIG levels	2013	CARDIoGRAM+C4D (63)
rs964184	11	*ZNF259–APOA5–APOA1*	0.18	1.05 (1.03–1.08)	5.600×10^{-5}	TRIG levels	2011	CARDIoGRAM (62)
rs17609940	6	*ANKS1A*	0.82	1.03 (1.00–1.05)	3.000×10^{-2}	HDL levels, height	2011	CARDIoGRAM (62)
rs3918226	7	*NOS3*	0.06	1.14 (1.09–1.19)	1.690×10^{-9}	BP	2015	1GP CARDIoGRAM+C4D (54)
rs2681472	12	*ATP2B1*	0.20	1.08 (1.05–1.10)	6.170×10^{-11}	BP	2012	Lu et al. (64)
rs17514846	15	*FURIN–FES*	0.44	1.05 (1.03–1.07)	3.100×10^{-7}	BP	2013	CARDIoGRAM+C4D (63)
rs72689147	4	*GUCY1A3*	0.82	1.07 (1.05–1.10)	6.070×10^{-9}	BP, cell growth/differentiation/ apoptosis	2013	CARDIoGRAM+C4D (63)
rs11830157	12	*KSR2*	0.36	1.12 (1.08–1.16)¶	2.120×10^{-9}	BMI	2015	1GP CARDIoGRAM+C4D (54)
rs663129	18	*PMAIP1–MC4R*	0.26	1.06 (1.04–1.08)	3.200×10^{-8}	BMI	2015	1GP CARDIoGRAM+C4D (54)
rs4252185	6	*PLG*	0.06	1.34 (1.28–1.41)	1.640×10^{-32}	coagulation	2013	CARDIoGRAM+C4D (63)
rs2519093	9	*ABO*	0.19	1.08 (1.06–1.11)	1.190×10^{-11}	coagulation, LDL levels	2011	CARDIoGRAM (62)
rs9349379	6	*PHACTR1*	0.43	1.14 (1.12–1.16)	1.810×10^{-42}	Arterial vessel wall endothelial cell	2009	MIGen (59)
rs9319428	13	*FLT1*	0.31	1.04 (1.02–1.06)	7.130×10^{-5}	Arterial vessel wall endothelial cell	2013	CARDIoGRAM+C4D (63)
rs8042271	15	*MFGE8–ABHD2*	0.90	1.10 (1.06–1.14)	3.680×10^{-8}	Arterial vessel wall endothelial cell	2015	1GP CARDIoGRAM+C4D (54)
rs7212798	17	*BCAS3*	0.15	1.08 (1.05–1.11)	1.880×10^{-8}	Arterial vessel wall endothelial cell	2015	1GP CARDIoGRAM+C4D (54)
rs4593108	4	*EDNRA*	0.80	1.07 (1.05–1.10)	8.820×10^{-10}	Arterial vessel wall smooth muscle cell	2013	CARDIoGRAM+C4D (63)
rs17087335	4	*REST–NOA1*	0.21	1.06 (1.04–1.09)	4.590×10^{-8}	Arterial vessel wall smooth muscle cell	2015	1GP CARDIoGRAM+C4D (54)
rs12202017	6	*TCF21*	0.70	1.07 (1.05–1.09)	1.980×10^{-11}	Arterial vessel wall smooth muscle cell	2011	CARDIoGRAM (62)
rs2107595	7	*HDAC9*	0.20	1.08 (1.05–1.10)	8.050×10^{-11}	Arterial vessel wall smooth muscle cell	2013	CARDIoGRAM+C4D (63)
rs2891168	9	*CDKN2BAS (9p21)*	0.49	1.21 (1.19–1.24)	2.290×10^{-98}	Arterial vessel wall smooth muscle cell	2007	McPherson et al., deCODE, WTCCC (46–48)
rs11191416	10	*CYP17A1–CNNM2–NT5C2*	0.87	1.08 (1.05–1.11)	4.650×10^{-9}	Arterial vessel wall smooth muscle cell	2011	CARDIoGRAM (62)
rs4468572	15	*ADAMTS7*	0.59	1.08 (1.06–1.10)	4.440×10^{-16}	Arterial vessel wall smooth muscle cell	2011	C4D, Reilly, CARDIoGRAM (61,62,89)
rs10840293	11	*SWAP70*	0.55	1.06 (1.04–1.08)	1.280×10^{-8}	Arterial vessel wall smooth muscle cell, inflammation/immune system/ cell migration-adhesion	2015	1GP CARDIoGRAM+C4D (54)
rs17678683	2	*ZEB2–ACO74093.1*	0.09	1.10 (1.07–1.14)	3.000×10^{-9}	Cell growth/differentiation/apoptosis	2013	CARDIoGRAM+C4D (63)
rs2128739	11	*PDGFD*	0.32	1.07 (1.05–1.09)	7.050×10^{-11}	Cell growth/differentiation/apoptosis	2011	C4D (61)
rs56062135	15	*SMAD3*	0.79	1.07 (1.05–1.10)	4.520×10^{-9}	Cell growth/differentiation/apoptosis	2015	1GP CARDIoGRAM+C4D (54)
rs46522	17	*UBE2Z*	0.51	1.04 (1.02–1.06)	1.840×10^{-5}	Cell growth/differentiation/apoptosis	2011	CARDIoGRAM (62)

Continued on the next page

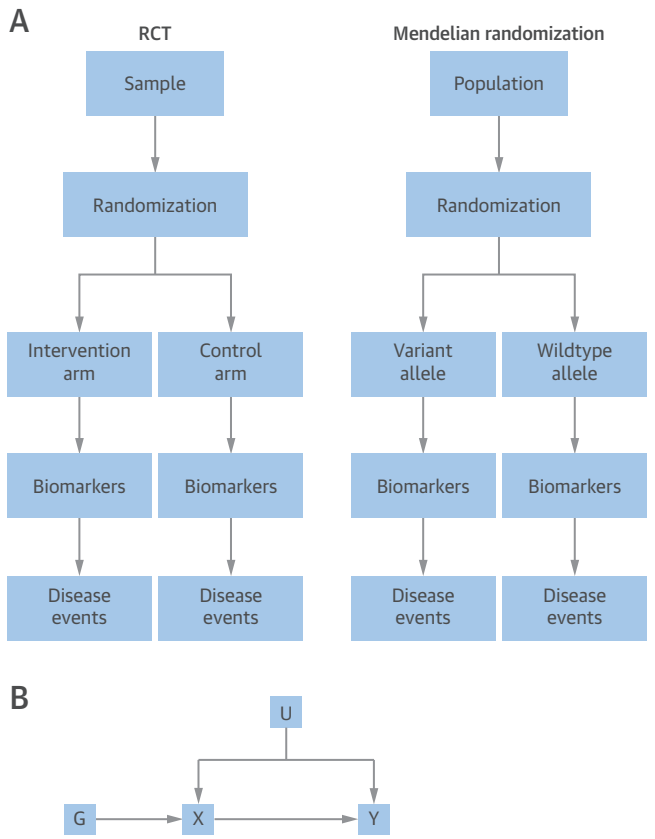
TABLE 1 Continued

Lead SNP†	Chromosome	Nearest Genes‡	Frequency of Allele Raising Risk	OR (95% CI)	p Value§	Potential Mechanism of Action‖	Year Locus First Reported to Reach Genome-Wide Significance	Consortium and/or Author (Ref. #)
rs9970807	1	*PPAP2B*	0.92	1.13 (1.10–1.17)	5.000×10^{-14}	Inflammation/Immune system/cell migration-adhesion	2011	CARDIoGRAM (62)
rs2487928	10	*KIAA1462*	0.42	1.06 (1.04–1.08)	4.410×10^{-11}	Inflammation/immune system/cell migration-adhesion	2011	C4D (61)
rs1870634	10	*CXCL12*	0.64	1.08 (1.06–1.10)	5.550×10^{-15}	Inflammation/immune system/cell migration-adhesion	2007	WTCCC and Cardiogenics (56)
rs1412444	10	*LIPA*	0.37	1.07 (1.05–1.09)	5.150×10^{-12}	Inflammation/immune system/cell migration-adhesion	2011	C4D (61)
rs6689306	1	*IL6R*	0.45	1.06 (1.04–1.08)	2.600×10^{-9}	Inflammation/immune system/cell migration-adhesion, cell growth/differentiation/apoptosis	2013	CARDIoGRAM+C4D (63)
rs67180937	1	*MIA3*	0.66	1.08 (1.06–1.11)	1.010×10^{-12}	Extracellular matrix	2007	WTCCC & Cardiogenics (56)
rs11838776	13	*COL4A1/A2*	0.26	1.07 (1.05–1.09)	1.830×10^{-10}	Extracellular matrix	2011	CARDIoGRAM (62)
rs16986953	2	*AK097927*	0.10	1.09 (1.06–1.12)	1.450×10^{-8}	Other/unknown	2013	CARDIoGRAM+C4D (63)
rs7568458	2	*VAMP5-VAMP8-GGCX*	0.45	1.06 (1.04–1.08)	3.620×10^{-10}	Other/unknown	2013	CARDIoGRAM+C4D (63)
rs6725887	2	*WDR12*	0.11	1.14 (1.11–1.18)	9.510×10^{-18}	Other/unknown	2009	MIGen (59)
rs9818870	3	*MRAS*	0.14	1.07 (1.04–1.10)	2.210×10^{-6}	Other/unknown	2009	Cardiogenics (57)
rs273909	5	*SLC22A4-SLC22A5*	0.12	1.06 (1.03–1.09)	1.240E−04	Other/unknown	2013	CARDIoGRAM+C4D (63)
rs6903956	6	*ADTRP-C6orf105*	0.35	1.00 (0.98–1.02)	9.600×10^{-1}	Other/unknown	2011	Wang et al. (176)
rs56336142	6	*KCNK5*	0.81	1.07 (1.04–1.09)	1.850×10^{-8}	Other/unknown	2013	CARDIoGRAM+C4D (63)
rs10953541	7	*7q22*	0.78	1.05 (1.03–1.08)	1.020×10^{-5}	Other/unknown	2011	C4D (61)
rs11556924	7	*ZC3HC1*	0.69	1.08 (1.05–1.10)	5.340×10^{-11}	Other/unknown	2011	CARDIoGRAM (62)
rs10139550	14	*HHIPL1*	0.42	1.06 (1.04–1.08)	1.380×10^{-8}	Other/unknown	2011	CARDIoGRAM (62)
rs216172	17	*SMG6*	0.35	1.05 (1.03–1.07)	5.070×10^{-7}	Other/unknown	2011	CARDIoGRAM (62)
rs12936587	17	*RAI1-PEMT-RASD1*	0.61	1.03 (1.01–1.05)	8.240×10^{-4}	Other/unknown	2011	CARDIoGRAM (62)
rs12976411	19	*ZNF507-LOC400684*	0.91	1.49 (1.38–1.67)¶	1.180×10^{-14}	Other/unknown	2015	1GP CARDIoGRAM+C4D (54)
rs28451064	21	*KCNE2* (gene desert)	0.12	1.14 (1.10–1.17)	1.330×10^{-15}	Other/unknown	2009	MIGen (59)
rs180803	22	*POM121L9P-ADORA2A*	0.97	1.20 (1.13–1.27)	1.640×10^{-10}	Other/unknown	2015	1GP CARDIoGRAM+C4D (54)

*Table shows susceptibility loci having reached genome-wide significance for CAD at any time, ORs, and p values for the lead SNP within each locus from the most recent CARDIoGRAMplusC4D meta-analysis, involving largely subjects of European (77%) and South Asian (13%) descent, sorted by potential mechanism of action and chromosomal location. Adapted from Supplemental Tables 2 and 4 in Nikpay et al. (54). †The most significant SNP in the locus, as determined by the CARDIOGRAMplusC4D 1,000 genomes meta-analysis. ‡The nearest 1 to 3 genes in the locus where the signal has been found. Although these genes are the most likely to be affected, it is possible that another gene that is slightly further away, but still nearby, may be responsible for the association, which would change the potential mechanism of action. For many loci, the causal gene in the region remains unclear. §The p value for several of these loci that have reached genome-wide significance ($p < 5 \times 10^{-8}$) in prior studies is greater than the genome-wide significance in the latest meta-analysis by CARDIoGRAMplusC4D, most likely due to statistical fluctuation. All of the lead SNPs, except for rs17609940 (ANKS1A) and rs6903956 (ADTRP-C6orf105), show a strong trend toward genome-wide significance, and likely represent true CAD susceptibility loci. The rs17609940 (ANKS1A) region requires further validation. rs6903956 (ADTRP-C6orf105) is a locus that was first identified in East Asians in 2011, but never replicated in both additional East Asian and non-East Asian populations, despite adequate power to detect the original effect size. ‖Potential mechanisms of action are on the basis of a combination of cross-referencing these regions with findings in the region for GWAS of risk factors for CAD and other related phenotypes (e.g., other complications of atherosclerosis, arterial aneurysms, migraines, and so on) and/or what is already known about the function of the nearby genes. ¶These loci were identified using a recessive model of inheritance; all other loci were identified using an additive model of inheritance.

BMI = body mass index; BP = blood pressure; CAD = coronary artery disease; CI = confidence interval; HDL = high-density lipoprotein; LDL = low-density lipoprotein; Lp(a) = lipoprotein(a); OR = odds ratio; SNP = single-nucleotide polymorphism; TRIG = triglycerides.

FIGURE 3 MR Study Provides Evidence of Causality Between a Risk Factor and an Outcome



(A) Similarities between RCTs and MR studies. Both designs eliminate confounding through randomization. **(B)** The MR model. To infer a causal role of a risk factor, X, in a disease, Y, 3 assumptions must be held. First, the genetic instrument, G, must associate with X. Second, G must be independent of confounders, U, of the X–Y relationship. Third, G must be associated with Y only through its effect on X. If these 3 relationships can be demonstrated, the risk factor, X, can be said to be causally related to the disease. MR = Mendelian randomization; RCT = randomized controlled trial. Reprinted with permission from Swerdlow et al. (175).

phospholipase IIA, lipoprotein-associated PLA2, high-sensitivity C-reactive protein, and high-density lipoprotein (HDL) are not causally associated with CAD (100). These observations are supported by the failure of RCTs that investigated folic acid supplementation, secretory phospholipase IIA and lipoprotein-associated PLA2 inhibitors, and, most recently, cholesterol ester transfer protein (CETP) inhibitors (101–105). The failure of the CETP inhibitors has been particularly enigmatic, given the profound elevations of HDL achieved with these drugs.

Other MR studies have supported a causal role for elevated body mass index, elevated BP, diabetes, LDL, lipoprotein(a), and triglycerides (100,106). These findings are supported by prior RCTs of weight loss and of a variety of antihypertensive drugs, statins, and ezetimibe (100,106). However, not all interventions influencing risk factors deemed causal through IV analysis have proven successful, indicating that MR studies cannot be trusted entirely to predict the success of all interventions influencing one's degree of exposure to causal risk factors. Notable failures in this respect include recent trials of niacin performed on a background of full-dose statin therapy, which did not result in a reduction of CAD events, despite marked lowering of triglyceride levels (101). Similarly, fibrates and omega-3 fatty acid supplements have not been convincingly shown to reduce the risk of CAD, despite impressive reductions in triglyceride levels (101,107). Last, the CETP inhibitor evacetrapib was recently reported to have no effect on outcomes, despite lowering LDL by ~37%, in addition to raising HDL (108). These failures are likely a consequence of an incomplete understanding of the various physiological pathways that control biomarker levels, and highlight a need to identify and target the subset of pathways that also influence the risk of CAD. Discrepancies between MR studies and results of drug trials could be resolved in the future, as we gain a better understanding of the mechanisms behind each genetic association related to risk factors of CAD. For example, MR studies restricted to the subset of variants associated with triglyceride levels, as well as effects on the level of expression of genes influenced by niacin or fibrates, would be expected to be negative.

Another insight derived through MR studies on risk factors for CAD has been a better appreciation of the potential benefit of long-term pharmacological therapy. Causal estimates derived from MR studies for quantitative traits, such as lipids, reflect the lifetime effects of a risk factor, as compared to the typical 3- to 5-year intervention achieved through a clinical trial (109). For example, MR studies suggest that the beneficial effects of LDL and BP lowering over a lifetime are substantially greater than has been observed in short-term observational studies, and RCTs of statin and antihypertensive drug therapies (99,110–112). In a comprehensive MR study involving 9 LDL-related SNPs in 6 genomic regions, as well as a meta-analysis of multiple studies involving >300,000 individuals and thousands of events, Ference et al. (112) documented that individuals with inherited LDL lowering experience a marked reduction in cardiac morbidity and mortality over their lifetime (112). They estimated that a decrease in 1 unit of LDL (38.7 mg/dl) since birth is associated with 54.5% reduction in cardiac events, which is 3-fold

greater than that observed in clinical trials involving statin therapies given over shorter periods of time later in life (112). In a more recent study, individuals with LDL cholesterol >190 mg/dl on a single measure, who were carrying damaging mutations within FH genes (*LDLR*, *APOB*, and *PCSK9*), were found to be at 2 to 3× higher risk of clinical CAD than individuals who were not carrying such mutations within every stratum of LDL (113). Furthermore, FH mutation carriers were found to have a higher cumulative exposure to LDL cholesterol than noncarriers in an analysis of participants with repeated measures over many years (113). These results reinforce the potential of early primary prevention of CAD, especially among those at very high lifetime risk at a young age, such as individuals experiencing FH or other polygenic causes of very elevated LDL levels.

An increasing appreciation of the strengths and weaknesses of MR studies has evolved over the last few years. First, the consequences of an underpowered MR study, either because of the use of a weak instrument or due to an inadequate number of events, is now very well understood (97). Over the next few years, mega-cohort studies involving biobanks and genetic data, combined with stronger instruments as more genetic discoveries are made, are expected to optimize the power of IV analyses and minimize the probability of either missing a causal association or falsely declaring its presence. Second, important new methods have been recently developed to identify potentially problematic instruments, which may result in bias and false conclusions about the causality of a risk factor (114–116). Thus, MR studies are expected to continue to serve as an important means of assessing the causal nature of a risk factor and estimating the probability of success of a novel therapeutic agent in improving outcomes of CAD.

CAD GENETIC RISK SCORES MAY IMPROVE OUR ABILITY TO IDENTIFY INDIVIDUALS AT HIGH RISK WHO ARE MOST LIKELY TO RESPOND TO THERAPY

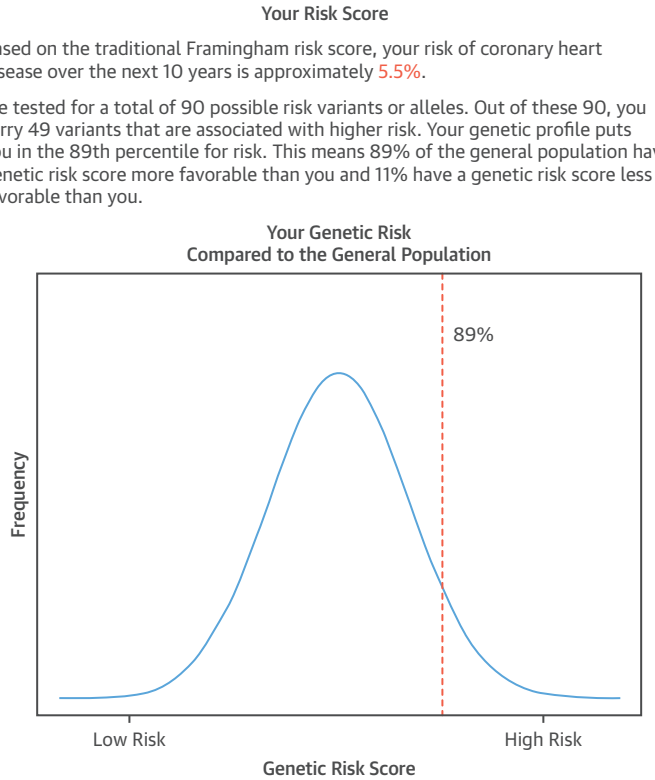
A major anticipated clinical application for genetic risk variants predisposing to CAD is their utilization to improve our ability to risk-stratify individuals (117,118). Although current clinical risk prediction scores for CAD perform relatively well compared with scores for other chronic diseases, substantial room for improvement exists, as a large pool of individuals with incident events carry either only 1 modifiable risk factor or only borderline risk factors (119–121).

The dilemma in selecting individuals for primary prevention can be illustrated in the case of a 47-year-old, white, pre-menopausal woman with total plasma cholesterol of 320 mg/dl, an HDL of 50 mg/dl, an LDL of 175 mg/dl, untreated BP of 120/80 mm Hg, and no other risk factors. According to the American College of Cardiology (ACC)/American Heart Association (AHA) pooled cohort risk calculator, her 10-year risk is only 2%, and treatment with statins to lower her LDL is not indicated unless she develops other risk factors or has a cardiac event (122). This individual is likely to have appreciable subclinical CAD, which may not be identified through a coronary artery calcification scan. Decreasing her plasma LDL would be expected to delay or prevent development of clinical CAD. Improving our ability to better predict which subset of women with similar risk factor profiles will develop an event through the addition of a novel biomarker, such as one's genetic susceptibility to CAD, would be expected to improve outcomes through more efficient application of established primary prevention therapies.

The most practical way to currently integrate genetics into risk prediction models, such as the Framingham Risk Score or the ACC/AHA pooled cohorts calculator, is through the calculation of a genetic risk score (GRS) for individuals (123). A GRS is a single variable that summarizes one's exposure to variants that increase risk for CAD (124). A GRS is typically calculated by summing the product of the number of high-risk variants inherited by each individual for each susceptibility variant and the log of the odds ratio previously determined in a GWAS for the same variant (123). **Figure 4** provides an example of how a GRS can be integrated into the Framingham Risk Score to update an individual's 10-year risk.

The use of GRS in clinical practice has been slow to materialize for several reasons, including the high cost of genotyping, the more modest effects of genetic variants on the risk of CAD than originally anticipated, and the challenge of improving a clinical risk score, such as the Framingham or ACC/AHA risk score, that already performs quite well (118,123,125). Consequently, it has been difficult to demonstrate substantial improvements to standard model performance metrics, such as the C-statistic, with the addition of a GRS, even though ample evidence now exists that a GRS of known CAD loci predicts incident CAD events independent of all other traditional risk factors (123,126–131). The challenge of demonstrating the clinical utility of a novel biomarker through observational studies alone has been further amplified by concerns of the power and reliability of some of the most commonly used model performance

FIGURE 4 Incorporating a GRS Into Standard Clinical Risk Assessment for CAD



The figure shows how a GRS of CAD variants can be integrated with a clinical risk score to update an individual's 10-year risk. The increase in the estimated 10-year risk illustrated here could result in a recommendation to initiate statins. CAD = coronary artery disease; GRS = genetic risk score. Reprinted with permission from Goldstein et al. (123).

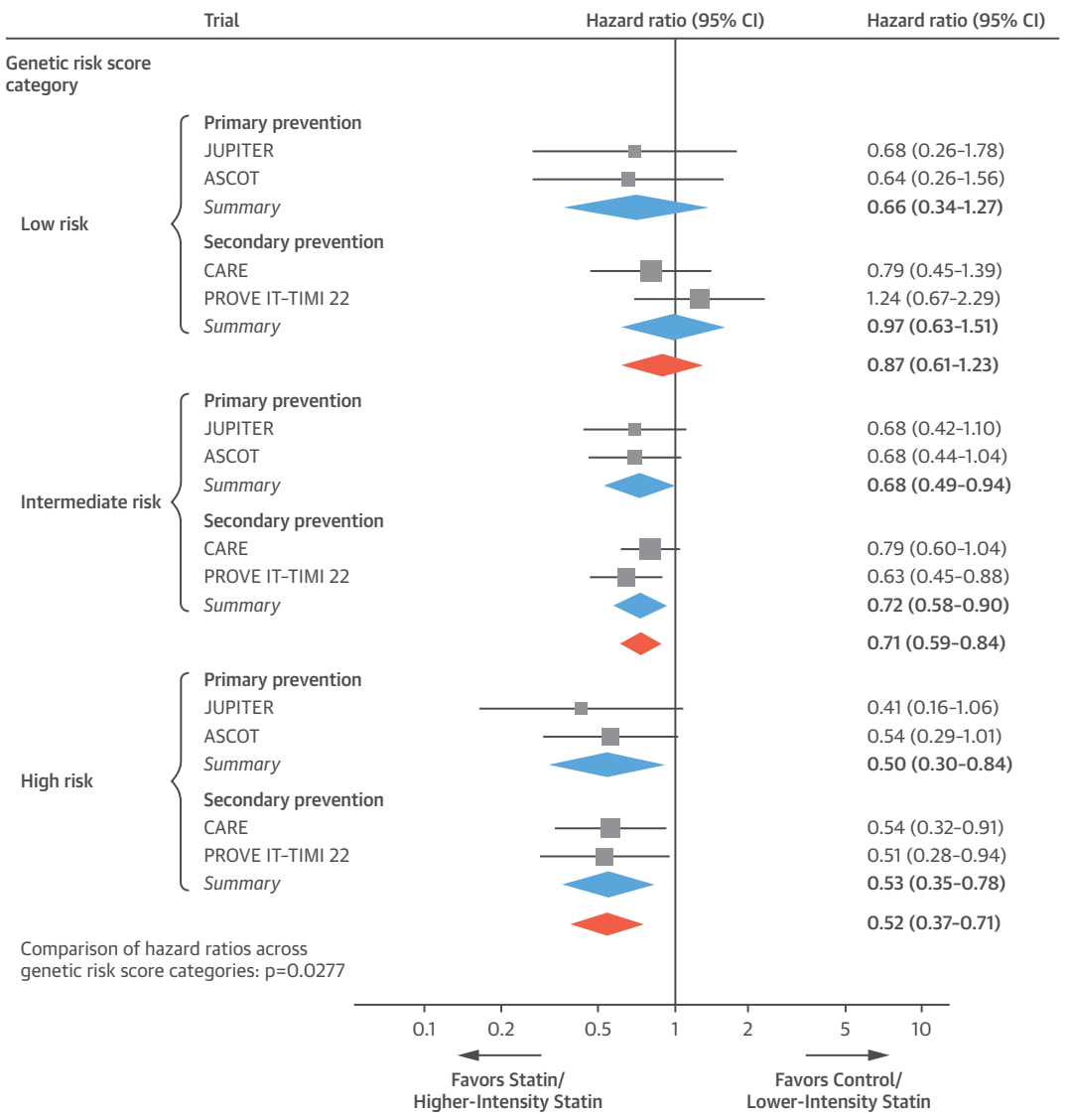
metrics for novel biomarkers, including the C-statistic, as well as the net reclassification index (132–134).

Family history would intuitively serve as a substitute for genetic risk. However, individual variants, as well as GRS of CAD, have been shown to predict clinical complications of CAD independent of family history (130). Furthermore, the magnitude of the association between family history and the risk of incident CAD is not obviously reduced when a GRS is introduced into a multivariate prediction model that includes family history (130). These observations support the notion that family history reflects exposure to common familial environmental risk factors of CAD to a larger degree than previously suspected, and the proportion of the heritability of CAD reflected by current GRS is not high enough to noticeably erode the predictive power of family history.

A recent report highlights the great potential for a GRS of CAD to improve outcomes, not only through primary prevention, but also possibly secondary prevention (127). In this report, a GRS involving 27 variants previously proven to be associated with risk for CAD was constructed after genotyping DNA biobanked at baseline from participants in 1 community-based cohort study (the Malmo Diet and Cancer Study), 2 primary prevention trials of statins (JUPITER and ASCOT), and 2 secondary prevention trials assessing the efficacy of statin therapy (CARE and PROVE IT-TIMI 22) (127). Among the 48,421 individuals and 3,477 events included in this study, investigators showed that the GRS not only predicted incident CAD events, but also predicted recurrent CAD events, independent of all traditional risk factors, as well as family history (127). Furthermore, a significant gradient of increasing relative risk reductions ($p = 0.0277$), summarized in **Figure 5**, as well as absolute risk reductions ($p = 0.0101$), was documented across the low, intermediate, and high genetic risk categories among the 4 randomized trials (127). The absolute risk reductions estimated a roughly 3-fold decrease in the number needed to treat to prevent 1 CAD event in the primary prevention trials (127). Specifically, in the primary prevention trials, the number needed to treat to prevent 1 such event in 10 years was 66 in people at low genetic risk, 42 in those at intermediate genetic risk, and 25 in those at high genetic risk in JUPITER, and 57, 47, and 20, respectively, in ASCOT. Thus, a GRS has the potential to serve not only as a prognostic marker, but also as a marker that can predict response to the single most important primary and secondary therapy already available (135). Such a combination of prognostic and predictive power for a single biomarker is uncommon, and has generally been observed only for certain markers of cancer (136). This combination offers the potential to substantially improve the efficiency of both primary and secondary established primary prevention strategies.

These findings, if successfully replicated, have important implications for the prevention and clinical management of CAD. Utilizing GRS, one could target large populations that are at high risk, such as premenopausal women with high genetic risk and optimal response to statin therapy, for primary prevention. Furthermore, the incremental predictive ability of the GRS is expected to improve over time, although the pace and magnitude of improvement remains unclear (118). However, a substantial improvement in the predictive ability of a GRS may eventually justify the early application of primary prevention

FIGURE 5 **A GRS of CAD Variants Identifies Optimal Responders to Established Prevention Therapy**



The figure shows hazard ratios for CAD with statin therapy across GRS categories. The **boxes** indicate the point estimates, and the **size of each box** represents the weight of a trial's data within that subgroup. The **horizontal lines** are the 95% CIs. The **diamonds** provide summary data. In PROVE IT-TIMI 22, the control group is moderate-intensity statin therapy (pravastatin 40 mg) and the statin group is high-intensity statin therapy (atorvastatin 80 mg). Subjects in the high genetic risk score category show the largest reductions in relative risk, whereas subjects at low risk have a nonsignificant reduction in relative risk. CI = confidence interval; other abbreviations as in Figure 4. Reprinted with permission from Mega et al. (127).

strategies, such as statins, to asymptomatic individuals who are at very high genetic risk, despite not having any other apparent risk factor. Such an approach is already indirectly supported by recent MR studies for LDL and CAD risk (112). Of course, such approaches would have to be supported by carefully conducted RCTs demonstrating improvement in outcomes or acceptable intermediate traits with the use of a GRS. Thus, the testing of GRS in prospective RCTs holds great promise for reducing the incidence of CAD.

WHOLE-EXOME AND -GENOME SEQUENCING STUDIES OF CAD

GWAS have been exceptionally successful in identifying susceptibility loci harboring common variants

with modest effects on the risk of CAD. These common variants are likely responsible for the majority of the heritability of CAD (137–141). However, less common variants with larger effect sizes exist both within loci already uncovered through GWAS, as well as in loci not yet linked to CAD. As many of these rare causal variants have not been previously observed, the only way to reliably identify them is through the direct sequencing of the DNA of subjects with disease.

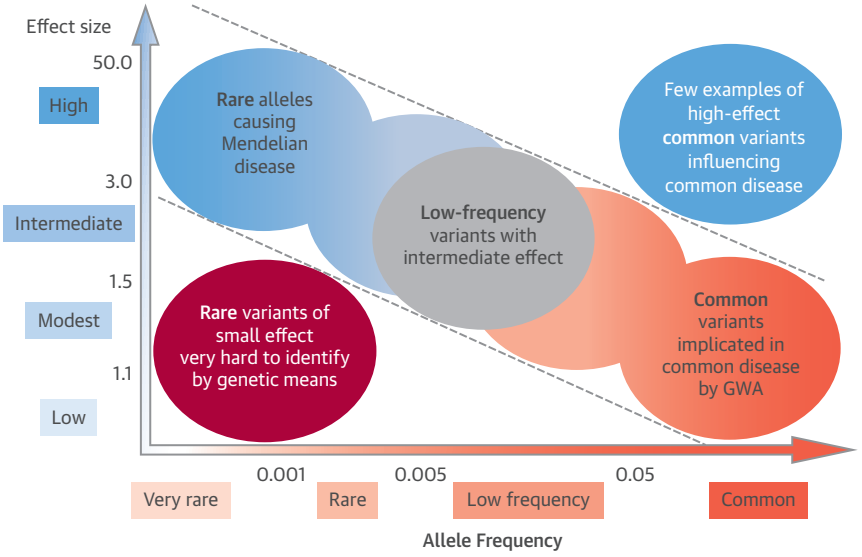
Next-generation sequencing platforms have been developed and continuously enhanced over the last decade (142). They are now mature enough to allow for large-scale genome sequencing projects. Two types of genome sequencing are possible: genome sequencing limited to the exons of all protein coding genes, referred to as whole-exome sequencing (WES), and genome sequencing of the entire genome, referred to as whole-genome sequencing (WGS). WES became economically feasible to conduct in large numbers approximately 6 years ago (143). In 2010, the National Heart, Lung, and Blood Institute (NHLBI) funded several groups to conduct WES association studies through the Exome Sequencing Project, a collaborative endeavor that included several sample sets with individuals experiencing clinical CAD (144). A subset of the rare variants uncovered through the NHLBI's WES project and other similar projects across the world were then incorporated into custom exome genotyping chips to facilitate the large-scale and cost-efficient replication of the findings in the discovery WES studies (145). This subset was restricted to variants predicted to most likely have severe consequences on protein function, such as damaging missense variants and variants affecting intron-exon splice sites. Overall, the discoveries to date using this strategy have been limited, primarily because effect sizes of less common and rare variants have been observed to be not much higher than the effect sizes observed for more common variants (146). Consequently, the study of very large sample sizes involving tens of thousands of individuals has been necessary to reliably identify novel associations involving rare variants uncovered through WES (145). Furthermore, it has been extremely challenging to examine rare variants individually (145). Instead, inference in WES studies have generally relied on comparing the aggregate number of rare variants within a gene observed among cases to that observed among control subjects (146).

Discoveries to date using WES and exome chip genotyping related to CAD have almost exclusively pointed to previously established genes involved in lipoprotein metabolism. These genes include genes influencing plasma LDL levels (*PSCK9, LDLR, and NPC1L1*) as well as genes influencing plasma triglyceride levels (*APOA5*, *APOC3*, *LPL*, and *ANGPTL4*) (147–150). A low-frequency variant in *SVEP1* was also found to be associated with CAD whose mechanism of action most likely involves effects on BP (150). Despite the limited novelty of these findings with respect to mechanisms of disease, they include *PCSK9* and *NPC1L1,* 2 genes with Food and Drug Association–approved drugs on the market that inhibit the activity of the gene's protein and lead to a reduction of both plasma LDL levels as well as the risk of CAD (106,151,152). These findings, combined with the results of MR studies of triglyceride levels and CAD, have reinvigorated the search for and the development of novel therapeutic agents that reduce the risk of CAD by interfering with the metabolism of triglyceride-rich lipoproteins (153–156). In other smaller studies, WES has recently been used to suggest that rare nonsynonymous variants in *SPTBN5*, *NID2*, and *ADAMTSL4* may protect against CAD and to implicate a 2-base insertion in *RECQL5* as a cause of early onset MI in a Chinese Han family (157,158).

Discoveries from WES are expected to expand as larger sample sizes become available. However, WES will undoubtedly be supplanted by WGS as the differential in price between the 2 approaches dwindles over time. WGS offers the additional benefit of documenting the presence of less common and rare variants in noncoding regulatory regions of an individual's genome whose degree of existence until recently was substantially underappreciated (159). Empirical evidence already exists that the effects sizes for such variants are not likely to be substantially larger than the effect sizes observed for causal variants within exons or common noncoding causal variants discovered through GWAS (140) (**Figure 6**). Thus, the ability to link such variants will also be heavily dependent on the profiling of very large sample sizes. The discovery of loci not previously implicated through GWAS or WES may be particularly challenging (**Figure 6**).

The first large-scale National Institutes of Health–funded projects to explore the WGS approach in the context of CAD and its risk factors have only recently been initiated, and include the NHLBI's WGS project within the Trans-Omics for Precision Medicine (160) program, as well as the National Human Genome Research Institute's new Genome Sequencing Program, which includes 4 Centers for Common Disease Genomics (161). The Trans-Omics for Precision Medicine program is currently completing WGS in >60,000 samples from 28 studies, of which at least 14 relate to coronary

FIGURE 6 Genetic Variants by Risk-Allele Frequency and Strength of Genetic Effect



Genome-wide association studies to date for CAD and other complex traits have explored common variants and established many associations with modest to low effect size on risk **(orange)**. Whole-exome and -genome sequencing studies seek to further identify all low-frequency and rare variants within genes or across the entire genome, respectively, that cannot be genotyped or imputed. The effect sizes of less common and rare variants are on average greater than more common variants, although the initial exome sequencing studies for CAD suggest that these differences in effect sizes are small, requiring very large sample sizes for discovery. CAD = coronary artery disease; GWA = genome-wide association. Reprinted with permission from Manolio et al. (141).

atherosclerosis or its risk factors. The initial disease focus of the Centers for Common Disease Genomics is early onset myocardial infarction, hemorrhagic stroke, and autism.

FUTURE DIRECTIONS

We have undoubtedly entered a golden era in the genetics of CAD early in the 21st century, highlighted by a rapid pace of discovery that has been facilitated by technological advances, which have not only made it feasible to perform high-throughput genetic profiling in a cost-efficient manner, but also allowed research groups around the world to share data and collaborate. These collaborations have resulted in the identification of over 60 susceptibility loci, which not only highlight the importance of well-established mechanisms of disease, such as cholesterol metabolism, but also reveal the presence of many novel pathways related to coronary atherosclerosis. The same large-scale population genetic studies are also providing us with opportunities to better understand the relationship between established or emerging risk factors and CAD through MR and to improve our ability to identify individuals who are at high risk of CAD through the incorporation of genetic risk into clinical risk scores.

Many challenging tasks remain to build on this remarkable progress and maximize the clinical utility of the knowledge gained. These tasks include the identification of the remaining susceptibility loci for CAD, proving the clinical utility of genetic data in the prevention of CAD, and acquiring a solid appreciation of the cellular and/or extracellular mechanisms responsible for genetic associations observed at the population level. Such knowledge will serve as the foundation for the development of novel therapeutic agents. Extremely large sample sizes are needed for additional discoveries, given the distribution of effect sizes observed to date for both common and rare variants, as well as the estimated proportion of the heritability of CAD explained by these variants to date. In the coming years, this need should be fulfilled by mega-biobanks involving at least one-half million participants, including, but not limited to, the UK Biobank, the China Kadoorie Biobank, the Million Veteran Program, and the soon-to-be-established National Institutes of Health Precision Medicine Initiative cohort (162–165). Such biobanks will likely also be leveraged to gain a better understanding of the

clinical utility of genetic risk scores, and to conduct additional, well-powered MR studies to complement studies published to date. Perhaps the most challenging of these tasks will be to understand the cellular mechanisms behind these associations. This task requires the careful interpretation of genetic association signals in combination with ongoing initiatives, such as the ENCODE (ENCyclopedia Of DNA Elements), GTEx (Genotype-Tissue Expression), and Roadmap Epigenomics projects, which are carefully documenting the complex and abundant regulatory regions of the genome across all human cell types (159). This task also requires compelling evidence to support the causality of variants within susceptibility loci through in vitro and in vivo targeted genome editing experiments (159). These new resources and techniques have already provided important mechanistic insights for several novel susceptibility loci for CAD, including those regions harboring the genes *CDKN2B, SORT1, TCF21, ADAMTS7, SMAD3,* and other loci (166–173). Although this understanding has not yet been translated to therapeutic agents targeting these loci, the rapid path of translation for other susceptibility loci, such as *PCSCK9,* provides some reassurance that such developments are forthcoming and are likely to provide us with groundbreaking opportunities to further reduce, and possibility eliminate, the threat of CAD in the 21st century.

REPRINT REQUESTS AND CORRESPONDENCE: Dr. Themistocles L. Assimes, Stanford University School of Medicine, Population Health Sciences Building, 1070 Arastradero Road, Palo Alto, California 94304. E-mail: tassimes@stanford.edu.

REFERENCES

1. Gaziano TA, Prabhakaran D, Gaziano JM. Global burden of cardiovascular disease. In: Mann DL, Zipes DP, Libby P, Bonow RO, Braunwald E, editors. Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine. 10th edition. Philadelphia, PA: Elsevier/Saunders, 2015:1–20.

2. Ridker PM, Libby P, Buring JE. Risk markers and the primary prevention of cardiovascular disease. In: Mann DL, Zipes DP, Libby P, Bonow RO, Braunwald E, editors. Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine. 10th edition. Philadelphia, PA: Elsevier/Saunders, 2015:891–933.

3. Morrow DA, Boden WE. Stable ischemic heart disease. In: Mann DL, Zipes DP, Libby P, Bonow RO, Braunwald E, editors. Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine. 10th edition. Philadelphia, PA: Elsevier/Saunders, 2015:1182–244.

4. Wald NJ, Law MR. A strategy to reduce cardiovascular disease by more than 80%. BMJ 2003;326:1419.

5. Yusuf S, Hawken S, Ounpuu S, et al., for the INTERHEART Study Investigators. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet 2004;364:937–52.

6. Stampfer MJ, Hu FB, Manson JE, et al. Primary prevention of coronary heart disease in women through diet and lifestyle. N Engl J Med 2000;343:16–22.

7. Tang L, Patao C, Chuang J, et al. Cardiovascular risk factor control and adherence to recommended lifestyle and medical therapies in persons with coronary heart disease (from the National Health and Nutrition Examination Survey 2007–2010). Am J Cardiol 2013;112:1126–32.

8. Colantonio LD, Bittner V. Managing residual risk after myocardial infarction among individuals with low cholesterol levels. Endocrinol Metab Clin North Am 2016;45:171–84.

9. Shapiro MD, Fazio S. From lipids to inflammation: new approaches to reducing atherosclerotic risk. Circ Res 2016;118:732–49.

10. Alfonsi JE, Hegele RA, Gryn SE. Pharmacogenetics of lipid-lowering agents: precision or indecision medicine? Curr Atheroscler Rep 2016;18:24.

11. Ross S, Nejat S, Paré G. Use of genetic data to guide therapy in arterial disease. J Thromb Haemost 2015;13 Suppl 1:S281–9.

12. Stranger BE, Forrest MS, Dunning M, et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 2007;315:848–53.

13. Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. Nature 2006;444:444–54.

14. Bhangale TR, Rieder MJ, Livingston RJ, et al. Comprehensive identification and characterization of diallelic insertion-deletion polymorphisms in 330 human candidate genes. Hum Mol Genet 2005;14:59–69.

15. Weiner MP, Gabriel SB, Stephens JC. Genetic Variation: A Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2007.

16. Sun JX, Helgason A, Masson G, et al. A direct characterization of human mutation based on microsatellites. Nat Genet 2012;44:1161–5.

17. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing [Erratum in Nature 2011;473:544]. Nature 2010;467:1061–73.

18. Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? Am J Hum Genet 2001;69:124–37.

19. Friedlander Y. Familial clustering of coronary heart disease: a review of its significance and role as a risk factor for the disease. In: Goldbourt U, De Faire U, Berg K, editors. Genetic Factors in Coronary Heart Disease. Boston, MA: Kluwer Academic, 1994:37–53.

20. Chow CK, Islam S, Bautista L, et al. Parental history and myocardial infarction risk across the world: the INTERHEART Study. J Am Coll Cardiol 2011;57:619–27.

21. Nielsen M, Andersson C, Gerds TA, et al. Familial clustering of myocardial infarction in first-degree relatives: a nationwide study. Eur Heart J 2013;34:1198–203.

22. de Faire U, Pedersen N. Studies of twins and adoptees in coronary heart disease. In: Goldbourt U, De Faire U, Berg K, editors. Genetic Factors in Coronary Heart Disease. Boston, MA: Kluwer Academic, 1994:55–68.

23. Zdravkovic S, Wienke A, Pedersen NL, et al. Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. J Intern Med 2002;252:247–54.

24. Wienke A, Holm NV, Skytthe A, et al. The heritability of mortality due to heart diseases: a correlated frailty model applied to Danish twins. Twin Res 2001;4:266–74.

25. Marenberg ME, Risch N, Berkman LF, et al. Genetic susceptibility to death from coronary heart disease in a study of twins. N Engl J Med 1994;330:1041–6.

26. Hejtmancik JF, Brink PA, Towbin J, et al. Localization of gene for familial hypertrophic cardiomyopathy to chromosome 14q1 in a diverse US population. Circulation 1991;83:1592–7.

27. Jarcho JA, McKenna W, Pare JA, et al. Mapping a gene for familial hypertrophic cardiomyopathy to chromosome 14q1. N Engl J Med 1989;321:1372–8.

28. Geisterfer-Lowrance AAT, Kass S, Tanigawa G, et al. A molecular basis for familial hypertrophic cardiomyopathy: a β cardiac myosin heavy chain gene missense mutation. Cell 1990;62:999–1006.

29. Musunuru K, Kathiresan S. Genetics of coronary artery disease. Annu Rev Genomics Hum Genet 2010;11:91–108.

30. Keramati AR, Fathzadeh M, Go GW, et al. A form of the metabolic syndrome associated with mutations in DYRK1B. N Engl J Med 2014;370:1909–19.

31. Mani A, Radhakrishnan J, Wang H, et al. LRP6 mutation in a family with early coronary disease and metabolic risk factors. Science 2007;315:1278–82.

32. Risch NJ. Searching for genetic determinants in the new millennium. Nature 2000;405:847–56.

33. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996;273:1516–7.

34. Lander ES, Linton LM, Birren B, et al., for the International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome [Errata in Nature 2001;411:720 and Nature 2001;412:565]. Nature 2001;409:860–921.

35. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. Science 2001;291:1304–51.

36. Sachidanandam R, Weissman D, Schmidt SC, et al., for the International SNP Map Working Group. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 2001;409:928–33.

37. Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. Nat Rev Genet 2002;3:391–7.

38. Wacholder S, Chanock S, Garcia-Closas M, et al. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst 2004;96:434–42.

39. Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 2003;19:149–50.

40. International HapMap Consortium. The International HapMap Project. Nature 2003;426:789–96.

41. International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851–61.

42. LaFramboise T. Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. Nucleic Acids Res 2009;37:4181–93.

43. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 2005;6:95–108.

44. Hardy J, Singleton A. Genomewide association studies and human disease. N Engl J Med 2009;360:1759–68.

45. Chanock SJ, Manolio T, Boehnke M, et al., for the NCI-NHGRI Working Group on Replication in Association Studies. Replicating genotype-phenotype associations. Nature 2007;447:655–60.

46. McPherson R, Pertsemlidis A, Kavaslar N, et al. A common allele on chromosome 9 associated with coronary heart disease. Science 2007;316:1488–91.

47. Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 2007;316:1491–3.

48. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661–78.

49. Schunkert H, Götz A, Braund P, et al. Repeated replication and a prospective meta-analysis of the association between chromosome 9p21.3 and coronary artery disease. Circulation 2008;117:1675–84.

50. Shen GQ, Li L, Rao 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 2008;28:360–5.

51. Barbalic M, Reiner AP, Wu C, et al. Genome-wide association analysis of incident coronary heart disease (CHD) in African Americans: a short report. PLoS Genet 2011;7:e1002199.

52. Franceschini N, Hu Y, Reiner AP, et al. Prospective associations of coronary heart disease loci in African Americans using the MetaboChip: the PAGE study. PLoS One 2014;9:e113203.

53. Lettre G, Palmer CD, Young T, et al. Genome-wide association study of coronary heart disease and its risk factors in 8,090 African Americans: the NHLBI CARe Project. PLoS Genet 2011;7:e1001300.

54. Nikpay M, Goel A, Won HH, et al., for the CARDIoGRAMplusC4D Consortium. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet 2015;47:1121–30.

55. Hannou SA, Wouters K, Paumelle R, et al. Functional genomics of the CDKN2A/B locus in cardiovascular and metabolic disease: what have we learned from GWASs? Trends Endocrinol Metab 2015;26:176–84.

56. Samani NJ, Erdmann J, Hall AS, et al., for the WTCCC and the Cardiogenics Consortium. Genomewide association analysis of coronary artery disease. N Engl J Med 2007;357:443–53.

57. Erdmann J, Grosshennig A, Braund PS, et al. New susceptibility locus for coronary artery disease on chromosome 3q22.3. Nat Genet 2009;41:280–2.

58. Gudbjartsson DF, Bjornsdottir US, Halapi E, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. Nat Genet 2009;41:342–7.

59. Myocardial Infarction Genetics Consortium. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants [Erratum in Nat Genet 2009;41:762]. Nat Genet 2009;41:334–41.

60. Trégouët DA, König IR, Erdmann J, et al. Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. Nat Genet 2009;41:283–5.

61. Coronary Artery Disease (C4D) Genetics Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. Nat Genet 2011;43:339–44.

62. Schunkert H, König IR, Kathiresan S, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet 2011;43:333–8.

63. Deloukas P, Kanoni S, Willenborg C, et al., for the CARDIoGRAMplusC4D Consortium. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet 2013;45:25–33.

64. Lu X, Wang L, Chen S, et al. Genome-wide association study in Han Chinese identifies four new susceptibility loci for coronary artery disease. Nat Genet 2012;44:890–4.

65. Takeuchi F, Yokota M, Yamamoto K, et al. Genome-wide association study of coronary artery disease in the Japanese. Eur J Hum Genet 2012;20:333–40.

66. Lee JY, Lee BS, Shin DJ, et al. A genome-wide association study of a coronary artery disease risk variant. J Hum Genet 2013;58:120–6.

67. Aoki A, Ozaki K, Sato H, et al. SNPs on chromosome 5p15.3 associated with myocardial infarction in Japanese population. J Hum Genet 2011;56:47–51.

68. Hirokawa M, Morita H, Tajima T, et al. A genome-wide association study identifies PLCL2 and AP3D1-DOT1L-SF3A2 as new susceptibility loci for myocardial infarction in Japanese. Eur J Hum Genet 2015;23:374–80.

69. Ioannidis JP, Ntzani EE, Trikalinos TA. 'Racial' differences in genetic effects for complex diseases. Nat Genet 2004;36:1312–8.

70. Coram MA, Duan Q, Hoffmann TJ, et al. Genome-wide characterization of shared and distinct genetic components that influence blood lipid levels in ethnically diverse human populations. Am J Hum Genet 2013;92:904–16.

71. Carlson CS, Matise TC, North KE, et al., for the PAGE Consortium. Generalization and dilution of association results from European GWAS in populations of non-European ancestry: the PAGE study. PLoS Biol 2013;11:e1001661.

72. Keating BJ, Tischfield S, Murray SS, et al. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. PLoS One 2008;3:e3583.

73. Clarke R, Peden JF, Hopewell JC, et al., for the PROCARDIS Consortium. Genetic variants associated with Lp(a) lipoprotein level and coronary disease. N Engl J Med 2009;361:2518–28.

74. IBC 50K CAD Consortium. Large-scale gene-centric analysis identifies novel variants for coronary artery disease. PLoS Genet 2011;7:e1002260.

75. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015;518:197–206.

76. Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. Nature Genet 2013;45:1274–83.

77. International Consortium for Blood Pressure Genome-Wide Association Studies. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature 2011;478:103–9.

78. Morris AP, Voight BF, Teslovich TM, et al., for the DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet 2012;44:981–90.

79. Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nat Genet 2010;42:441–7.

80. Solovieff N, Cotsapas C, Lee PH, et al. Pleiotropy in complex traits: challenges and strategies. Nat Rev Genet 2013;14:483–95.

81. Reich DE, Lander ES. On the allelic spectrum of human disease. Trends Genet 2001;17:502–10.

82. Taylor F, Huffman MD, Macedo AF, et al. Statins for the primary prevention of cardiovascular disease. Cochrane Database Syst Rev 2013;1:CD004816.

83. Rodriguez F, Knowles JW. PCSK9 inhibition: current concepts and lessons from human genetics. Curr Atherosclerosis Rep 2015;17:487.

84. Fuster V, Walsh RA, Harrington RA, editors. Hurst's The Heart. 13th edition. New York, NY: McGraw-Hill, 2011.

85. Mann DL, Zipes DP, Libby P, Bonow RO, Braunwald E, editors. Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine. 10th edition. Philadelphia, PA: Elsevier/Saunders, 2015.

86. Hong MK, Mintz GS, Lee CW, et al. Plaque ruptures in stable angina pectoris compared with acute coronary syndrome. Int J Cardiol 2007;114:78–82.

87. Hong MK, Mintz GS, Lee CW, et al. Comparison of coronary plaque rupture between stable angina and acute myocardial infarction: a three-vessel intravascular ultrasound study in 235 patients. Circulation 2004;110:928–33.

88. Burke AP, Kolodgie FD, Farb A, et al. Healed plaque ruptures and sudden coronary death: evidence that subclinical rupture has a role in plaque progression. Circulation 2001;103:934–40.

89. Reilly MP, Li M, He J, et al. 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;377:383–92.

90. Gill JC, Endres-Brooks J, Bauer PJ, et al. The effect of ABO blood group on the diagnosis of von Willebrand disease. Blood 1987;69:1691–5.

91. Spiel AO, Gilbert JC, Jilma B. von Willebrand factor in cardiovascular disease: focus on acute coronary syndromes. Circulation 2008;117:1449–59.

92. Dandona S, Stewart AF, Chen L, et al. Gene dosage of the common variant 9p21 predicts severity of coronary artery disease. J Am Coll Cardiol 2010;56:479–86.

93. Chan K, Patel RS, Newcombe P, et al. Association between the chromosome 9p21 locus and angiographic coronary artery disease burden: a collaborative meta-analysis. J Am Coll Cardiol 2013;61:957–70.

94. Bjornsson E, Gudbjartsson DF, Helgadottir A, et al. Common sequence variants associated with coronary artery disease correlate with the extent of coronary atherosclerosis. Arterioscler Thromb Vasc Biol 2015;35:1526–31.

95. Bentzon JF, Otsuka F, Virmani R, et al. Mechanisms of plaque formation and rupture. Circ Res 2014;114:1852–66.

96. Salfati E, Nandkeolyar S, Fortmann SP, et al. Susceptibility loci for clinical coronary artery disease and subclinical coronary atherosclerosis throughout the life-course. Circ Cardiovasc Genet 2015;8:803–11.

97. Burgess S, Thompson SG. Mendelian Randomization: Methods for Using Genetic Variants in Causal Estimation. Boca Raton, FL: Chapman & Hall/CRC, 2015.

98. Lawlor DA, Davey Smith G, Kundu D, et al. Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence? Lancet 2004;363:1724–7.

99. Voight BF, Peloso GM, Orho-Melander M, et al. Plasma HDL cholesterol and risk of myocardial infarction: a Mendelian randomisation study. Lancet 2012;380:572–80.

100. Jansen H, Samani NJ, Schunkert H. Mendelian randomization studies in coronary artery disease. Eur Heart J 2014;35:1917–24.

101. Keene D, Price C, Shun-Shin MJ, et al. Effect on cardiovascular risk of high density lipoprotein targeted drug treatments niacin, fibrates, and CETP inhibitors: meta-analysis of randomised controlled trials including 117,411 patients. BMJ 2014;349:g4379.

102. Stability Investigators. Darapladib for preventing ischemic events in stable coronary heart disease. N Engl J Med 2014;370:1702–11.

103. Nicholls SJ, Kastelein JJ, Schwartz GG, et al., for the VISTA-16 Investigators. Varespladib and cardiovascular events in patients with an acute coronary syndrome: the VISTA-16 randomized clinical trial. JAMA 2014;311:252–62.

104. O'Donoghue ML, Braunwald E, White HD, et al., for the SOLID-TIMI 52 Investigators. Effect of darapladib on major coronary events after an acute coronary syndrome: the SOLID-TIMI 52 randomized clinical trial. JAMA 2014;312:1006–15.

105. Yang HT, Lee M, Hong KS, et al. Efficacy of folic acid supplementation in cardiovascular disease prevention: an updated meta-analysis of randomized controlled trials. Eur J Intern Med 2012;23:745–54.

106. Cannon CP, Blazing MA, Giugliano RP, et al., for the IMPROVE-IT Investigators. Ezetimibe added to statin therapy after acute coronary syndromes. N Engl J Med 2015;372:2387–97.

107. Walz CP, Barry AR, Koshman SL. Omega-3 polyunsaturated fatty acid supplementation in the prevention of cardiovascular disease. Can Pharm J (Ott) 2016;149:166–73.

108. Nicholls SJ. Assessment of clinical effects of cholesteryl ester transfer protein inhibition with evacetrapib in patients at a high risk for vascular outcomes—ACCELERATE. Paper presented at: American College of Cardiology Annual Scientific Session; April 3, 2016; Chicago, IL.

109. Pierce BL, Van der Weele TJ. The effect of non-differential measurement error on bias, precision and power in Mendelian randomization studies. Int J Epidemiol 2012;41:1383–93.

110. Kathiresan S, Melander O, Anevski D, et al. Polymorphisms associated with cholesterol and risk of cardiovascular events. N Engl J Med 2008;358:1240–9.

111. Lieb W, Jansen H, Loley C, et al. Genetic predisposition to higher blood pressure increases coronary artery disease risk. Hypertension 2013;61:995–1001.

112. Ference BA, Yoo W, Alesh I, et al. Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. J Am Coll Cardiol 2012;60:2631–9.

113. Khera AV, Won HH, Peloso GM, et al. Diagnostic yield and clinical utility of sequencing familial hypercholesterolemia genes in patients with severe hypercholesterolemia. J Am Coll Cardiol 2016;67:2578–89.

114. Burgess S, Bowden J, Fall T, Ingellson E, Thompson SG. Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. Epidemiology 2017;28:30–42.

115. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44:512–25.

116. Burgess S, Harshfield E. Mendelian randomization to assess causal effects of blood lipids on coronary heart disease: lessons from the past and applications to the future. Curr Opin Endocrinol Diabetes Obes 2016;23:124–30.

117. Manolio TA. Genomewide association studies and assessment of the risk of disease. N Engl J Med 2010;363:166–76.

118. Thanassoulis G, Vasan RS. Genetic cardiovascular risk prediction: will we get there? Circulation 2010;122:2323–34.

119. Berger JS, Jordan CO, Lloyd-Jones D, et al. Screening for cardiovascular risk in asymptomatic patients. J Am Coll Cardiol 2010;55:1169–77.

120. Barlow WE, White E, Ballard-Barbash R, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. J Natl Cancer Inst 2006;98:1204–14.

121. Vasan RS, Sullivan LM, Wilson PW, et al. Relative importance of borderline and elevated levels of coronary heart disease risk factors

[Erratum in Ann Intern Med 2005;142:681]. Ann Intern Med 2005;142:393–402.

122. Goff DC Jr., Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines [Erratum in J Am Coll Cardiol 2014;63:3026]. J Am Coll Cardiol 2014;63:2935–59.

123. Goldstein BA, Knowles JW, Salfati E, et al. Simple, standardized incorporation of genetic risk into non-genetic risk prediction tools for complex traits: coronary heart disease as an example. Front Genet 2014;5:254.

124. Morrison AC, Bare LA, Chambless LE, et al. Prediction of coronary heart disease risk using a genetic risk score: the Atherosclerosis Risk in Communities Study. Am J Epidemiol 2007;166:28–35.

125. Wilson PW. Challenges to improve coronary heart disease risk assessment. JAMA 2009;302:2369–70.

126. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. Clin Chem 2008;54:17–23.

127. Mega JL, Stitziel NO, Smith JG, et al. Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. Lancet 2015;385:2264–71.

128. Krarup NT, Borglykke A, Allin KH, et al. A genetic risk score of 45 coronary artery disease risk variants associates with increased risk of myocardial infarction in 6041 Danish individuals. Atherosclerosis 2015;240:305–10.

129. Ganna A, Magnusson PK, Pedersen NL, et al. Multilocus genetic risk scores for coronary heart disease prediction. Arterioscler Thromb Vasc Biol 2013;33:2267–72.

130. Tada H, Melander O, Louie JZ, et al. Risk prediction by genetic risk scores for coronary heart disease is independent of self-reported family history. Eur Heart J 2016;37:561–7.

131. Abraham G, Havulinna AS, Bhalala OG, et al. Genomic prediction of coronary heart disease. Eur Heart J 2016 Sep 21 [E-pub ahead of print].

132. Kerr KF, Wang Z, Janes H, et al. Net reclassification indices for evaluating risk prediction instruments: a critical review. Epidemiology 2014;25:114–21.

133. Pepe MS, Janes H, Li CI. Net risk reclassification P values: valid or misleading? J Natl Cancer Inst 2014;106:dju041.

134. Pepe MS, Kerr KF, Longton G, et al. Testing for improvement in prediction model performance. Stat Med 2013;32:1467–82.

135. Težak Z, Kondratovich MV, Mansfield E. US FDA and personalized medicine: in vitro diagnostic regulatory perspective. Per Med 2010;7:517–30.

136. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet 2011;378:771–84.

137. Agarwala V, Flannick J, Sunyaev S, Altshuler D. Evaluating empirical bounds on complex disease genetic architecture. Nat Genet 2013;45:1418–27.

138. Park JH, Gail MH, Weinberg CR, et al. Distribution of allele frequencies and effect sizes and their interrelationships for common genetic susceptibility variants. Proc Natl Acad Sci U S A 2011;108:18026–31.

139. Lee SH, Yang J, Chen GB, et al. Estimation of SNP heritability from dense genotype data. Am J Hum Genet 2013;93:1151–5.

140. Fuchsberger C, Flannick J, Teslovich TM, et al. The genetic architecture of type 2 diabetes. Nature 2016;536:41–7.

141. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747–53.

142. van Dijk EL, Auger H, Jaszczyszyn Y, et al. Ten years of next-generation sequencing technology. Trends Genet 2014;30:418–26.

143. Kim DW, Nam SH, Kim RN, et al. Whole human exome capture for high-throughput sequencing. Genome 2010;53:568–74.

144. Fu W, O'Connor TD, Jun G, et al. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. Nature 2013;493:216–20.

145. Guo Y, He J, Zhao S, et al. Illumina human exome genotyping array clustering and quality control. Nat Protoc 2014;9:2643–62.

146. Auer PL, Lettre G. Rare variant association studies: considerations, challenges and opportunities. Genome Med 2015;7:16.

147. Do R, Stitziel NO, Won HH, et al. Exome sequencing identifies rare LDLR and APOA5 alleles conferring risk for myocardial infarction. Nature 2015;518:102–6.

148. TG and HDL Working Group of the Exome Sequencing Project, National Heart Lung Blood Institute. Loss-of-function mutations in APOC3, triglycerides, and coronary disease. N Engl J Med 2014;371:22–31.

149. Myocardial Infarction Genetics Consortium Investigators. Inactivating mutations in NPC1L1 and protection from coronary heart disease. N Engl J Med 2014;371:2072–82.

150. Myocardial Infarction Genetics and CARDIoGRAM Exome Consortia Investigators. Coding variation in ANGPTL4, LPL, and SVEP1 and the risk of coronary disease. N Engl J Med 2016;374:1134–44.

151. Robinson JG, Farnier M, Krempf M, et al., for the ODYSSEY LONG TERM Investigators. Efficacy and safety of alirocumab in reducing lipids and cardiovascular events. N Engl J Med 2015;372:1489–99.

152. Sabatine MS, Giugliano RP, Wiviott SD, et al., for the Open-Label Study of Long-Term Evaluation against LDL Cholesterol (OSLER) Investigators. Efficacy and safety of evolocumab in reducing lipids and cardiovascular events. N Engl J Med 2015;372:1500–9.

153. Do R, Willer CJ, Schmidt EM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. Nat Genet 2013;45:1345–52.

154. Burgess S, Freitag DF, Khan H, et al. Using multivariable Mendelian randomization to disentangle the causal effects of lipid fractions. PLoS One 2014;9:e108891.

155. C Reactive Protein Coronary Heart Disease Genetics Collaboration (CCGC). Association between C reactive protein and coronary heart disease: Mendelian randomisation analysis based on individual participant data. BMJ 2011;342:d548.

156. Kathiresan S. Developing medicines that mimic the natural successes of the human genome: lessons from NPC1L1, HMGCR, PCSK9, APOC3, and CETP. J Am Coll Cardiol 2015;65:1562–6.

157. Xie X, Zheng YY, Adi D, et al. Exome sequencing in a family identifies RECQL5 mutation resulting in early myocardial infarction. Medicine (Baltimore) 2016;95:e2737.

158. Abramowitz Y, Roth A, Keren G, et al. Whole-exome sequencing in individuals with multiple cardiovascular risk factors and normal coronary arteries. Coronary Artery Dis 2016;27:257–66.

159. Tak YG, Farnham PJ. Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. Epigenetics Chromatin 2015;8:57.

160. National Heart, Lung, and Blood Institute. NHLBI Trans-Omics for Precision Medicine Whole Genome Sequencing Program. 2016. Available at: https://www.nhlbiwgs.org/. Accessed October 24, 2016.

161. National Human Genome Research Institute. NHGRI Genome Sequencing Program (GSP): Centers for Common Disease Genomics. 2016. Available at: https://www.genome.gov/27563570. Accessed October 24, 2016.

162. Allen NE, Sudlow C, Peakman T, et al., for the Biobank UK. UK biobank data: come and get it. Sci Transl Med 2014;6:224ed4.

163. Chen Z, Chen J, Collins R, et al., for the China Kadoorie Biobank (CKB) Collaborative Group. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. Int J Epidemiol 2011;40:1652–66.

164. Gaziano JM, Concato J, Brophy M, et al. Million Veteran Program: a mega-biobank to study genetic influences on health and disease. J Clin Epidemiol 2016;70:214–23.

165. Shah SH, Arnett D, Houser SR, et al. Opportunities for the cardiovascular community in the precision medicine initiative. Circulation 2016;133:226–31.

166. Musunuru K, Strong A, Frank-Kamenetsky M, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature 2010;466:714–9.

167. Kjolby M, Andersen OM, Breiderhoff T, et al. Sort1, encoded by the cardiovascular risk locus 1p13.3, is a regulator of hepatic lipoprotein export. Cell Metab 2010;12:213–23.

168. Kessler T, Zhang L, Liu Z, et al. ADAMTS-7 inhibits re-endothelialization of injured arteries

and promotes vascular remodeling through cleavage of thrombospondin-1. Circulation 2015; 131:1191–201.

169. Bauer RC, Tohyama J, Cui J, et al. Knockout of Adamts7, a novel coronary artery disease locus in humans, reduces atherosclerosis in mice. Circulation 2015;131:1202–13.

170. Nurnberg ST, Zhang H, Hand NJ, et al. From loci to biology: functional genomics of genome-wide association for coronary disease. Circ Res 2016;118:586–606.

171. Zhao Y, Chen J, Freudenberg JM, et al. Network-based identification and prioritization of key regulators of coronary artery disease loci. Arterioscler Thromb Vasc Biol 2016;36:928–41.

172. Miller CL, Pjanic M, Wang T, et al. Integrative functional genomics identifies regulatory mechanisms at coronary artery disease loci. Nat Commun 2016;7:12092.

173. Pjanic M, Miller CL, Wirka R, et al. Genetics and genomics of coronary artery disease. Curr Cardiol Rep 2016;18:102.

174. Cunnington MS, Keavney B. Genetic mechanisms mediating atherosclerosis susceptibility at the chromosome 9p21 locus. Curr Atheroscler Rep 2011;13:193–201.

175. Swerdlow DI, Hingorani AD, Humphries SE. Genetic risk factors and Mendelian randomization in cardiovascular disease. Curr Cardiol Rep 2015; 17:33.

176. Wang F, Xu CQ, He Q, et al. Genome-wide association identifies a susceptibility locus for coronary artery disease in the Chinese Han population. Nat Genet 2011;43: 345–9.

KEY WORDS genetic variation, Mendelian randomization, risk scores, sequencing