There are currently almost no genetic data associated with the acute coronary syndromes of unstable angina or myocardial infarction (MI). The only documented genetic variants associated with increased risk for MI, per se, are the A and B blood group variants. There are no genetic data differentially associated with ST segment elevation MI or non-ST segment elevation MI. Genetic variants predisposing specifically to plaque rupture are yet to be detected. The following review is primarily on the genetic risk variants predisposing to coronary atherosclerosis, the process underlying the acute coronary syndromes.

**Genetics of Coronary Artery Disease**

Robert Roberts

**Abstract:** There is almost no data on the genetics of acute coronary syndromes, so this review discusses primarily the 50 genetic risk variants associated with coronary artery disease that are of genome-wide significance in the discovery population and replicated in an independent population. All of these risk variants are extremely common with more than half occurring in >50% of the general population. They increased only minimally the relative risk for coronary artery disease. The most striking finding is that 35 of the 50 risk variants act independently of known risk factors, indicating there are several pathways yet to be appreciated, contributing to the pathogenesis of coronary atherosclerosis and myocardial infarction. All of the genetic variants seem to act through atherosclerosis, except for the ABO blood groups, which show that A and B are associated with increased risk for myocardial infarction, mediated by a prolonged von Willebrand plasma half life leading to thrombosis. The potential molecular mechanisms of 9p21 are discussed, including cell cycle kinase inhibitors. Discovery of risk variants associated with PCSK9 has led to the development of novel treatment for plasma low-density lipoprotein cholesterol. A monoclonal antibody inhibiting PCSK9 has already undergone phase I and II clinical trials, showing it is a potent inhibitor of low-density lipoprotein cholesterol and is mediated through more rapid removal of low-density lipoprotein cholesterol from the plasma. This therapy complements that of statin therapy, which inhibits the synthesis of cholesterol. The benefits of Mendelian randomization to assess safety and efficacy and their limitations are discussed along with future directions. (Circ Res. 2014;114:1890-1903.)

**Key Words:** base sequence ■ coronary artery disease ■ genomics
with an average frequency of one tenth of 1%. Responsible mutations usually occur in protein-coding regions, are highly penetrant, and disrupt the function of the protein, resulting in an easily recognized phenotype. Insertion of the mutant gene as a transgene in animal models induces the expected phenotype, indicating the mutation is both necessary and sufficient to induce the phenotype. Nevertheless, few genes were discovered because success required prior knowledge of the protein involved.

This changed rapidly in the 1980s with the development of the technique referred to as positional cloning. This technique made it possible to determine the approximate chromosomal location of the gene responsible for the disease without the need for knowledge of the protein. A pedigree of 2 to 3 generations having individuals affected with the disease could be genotyped with DNA markers of known chromosomal location. Analysis for genetic linkage could then differentially detect markers segregating individuals affected with the disease versus those unaffected. Markers co-inherited by individuals affected with the disease indicate the marker is in close physical proximity to the gene responsible for the disorder. This was followed by the development of highly informative markers, initially restriction fragment–linked polymorphisms, followed by variable repeats of di-, tri-, and tetranucleotide markers. These new markers could be analyzed by polymerase chain reaction within a few hours, whereas previous markers detected by Southern blot analysis required 10 to 12 days.

These developments markedly facilitated the discovery of genes responsible for rare single-gene disorders. The first cardiovascular disorder to undergo chromosomal mapping was familial hypertrophic cardiomyopathy,2,3 with the discovery of the responsible missense mutation in the cardiac β-myosin heavy chain gene.2 This was followed by the discovery of genes for multiple cardiovascular disorders, including long QT syndrome,3 Brugada syndrome,4 atrial fibrillation,5 dilated cardiomyopathy,6 Wolff–Parkinson–White syndrome,7 and arrhythmogenic right ventricular dysplasia.8 In the subsequent 2 decades of the estimated 7000 single-gene disorders, ≥1 mutant genes would be identified for >3000 of these rare disorders,11 including several mutations in the low-density lipoprotein receptor (LDL-R) causing familial hypercholesterolemia.

**Nonstandard Abbreviations and Acronyms**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAD</td>
<td>coronary artery disease</td>
</tr>
<tr>
<td>GWAS</td>
<td>genome-wide association study</td>
</tr>
<tr>
<td>HDL-C</td>
<td>high-density lipoprotein cholesterol</td>
</tr>
<tr>
<td>LDL-C</td>
<td>low-density lipoprotein cholesterol</td>
</tr>
<tr>
<td>LDL-R</td>
<td>low-density lipoprotein receptor</td>
</tr>
<tr>
<td>LIPG</td>
<td>lipase gene</td>
</tr>
<tr>
<td>MI</td>
<td>myocardial infarction</td>
</tr>
<tr>
<td>RCT</td>
<td>randomized controlled trial</td>
</tr>
<tr>
<td>SNP</td>
<td>single-nucleotide polymorphism</td>
</tr>
</tbody>
</table>

**Genome-Wide Association Study Era: Common Complex Diseases**

Despite 2 decades of great success in rare single-gene disorders, such success for common diseases would have to wait further developments. It was recognized even in the 1990s that the application of genetic linkage analysis to detect markers segregating the disease in affected families would not be appropriate for common diseases such as coronary artery disease (CAD) or cancer. These diseases are common and would be expected to be associated with genes that occur commonly.

Epidemiological studies had strongly indicated that genetic predisposition resulted from a combination of multiple genes, each having only minimal to moderate effect on the phenotype. Thus, in contrast to rare single-gene disorders, no 1 gene was necessary or sufficient to induce the phenotype. These disorders do not exhibit a phenotype following a Mendelian pattern; therefore, an approach different than linkage analysis of pedigrees would be required to identify the genes responsible for such genetic predisposition. Furthermore, computer modeling by Kruglyak and Nickerson12 postulated that it would require hundreds of thousands of markers, spaced at intervals of not more than 6000 bases,13 in contrast to single-gene disorders that require only a few hundred markers spaced at intervals of ≥10 million bases. Second, instead of family studies, it would require unrelated individuals in what is known as a case–control association study.14,15 This approach compares the frequency of the marker in cases versus controls, and if significantly more common in cases, it would indicate that the marker is in a DNA region associated with increased predisposition for the disease.7 This would require sample sizes involving thousands of individuals with (cases) and without (controls) the disease. In the 1990s, because such an array of markers was not available, the first approach to the case–control study was to select candidate genes based on their function and compare their frequency in cases versus controls. Studies reported >100 candidates associated with genetic predisposition to CAD,16,17 of which almost none would be confirmed in subsequent studies.

The first major breakthrough came in 2005, with the development of a microarray of 500 000 DNA markers18 followed by arrays containing 1 million DNA markers.19 These markers consisted of single-nucleotide polymorphisms (SNPs), which are relatively uniformly distributed throughout the genome. The HapMap project20 had annotated more than a million of these SNPs to their chromosomal location, and it indicated that the average number of SNPs per human genome is ≥3 million. Furthermore, although the human genome sequences are 99.5% identical, of the 15 million base pairs that differ, studies would indicate that 3 million SNPs are in large part responsible for human variation, including susceptibility to disease.21 This was followed by platforms enabling the analysis of such massive data.19 It was now possible to assess the frequency of markers distributed throughout the human genome and determine those markers occurring more commonly with the disease than with controls. This approach is unbiased and requires no presupposition of which marker or gene is more likely to be a risk candidate. This genome-wide case–control association approach is now referred to as a genome-wide association study (GWAS). If 1 genotype with 1 million SNPs and accepting a P value of 0.05 to be statistically significant, it would be associated with ≥50 000 false-positives. A stringent statistical correction to account for false-positives was
proposed by Risch et al., namely the Bonferroni correction derived by dividing 1 million into 0.05 giving a $P$ value of $5 \times 10^{-8}$. This $P$ value of $5 \times 10^{-8}$ is now referred to as genome-wide significant $P$ value. Recognizing this stringent statistical requirement, sample sizes would have to include thousands of individuals with the disease (cases) and thousands without the disease (controls). Additional stringency was added, namely those SNPs showing increased frequency of genome-wide significance in cases (discovery population) would be genotyped and analyzed in an independent population for replication (replication population).

**Heritability and CAD**

For decades, epidemiological and familial studies have claimed that 40% to 60% of susceptibility for CAD is inherited. Results of follow-up studies from the Framingham Study showed that a family history of CAD is associated with a 2.4-fold increased risk of CAD in men and a 2.2-fold increased risk in women. The frequency of CAD in monozygotic twins compared with dizygotic twins averages 44% and 14%, respectively, based on the Danish Twin Registry Study.

More recently, the INTERHEART study, after correcting for other risk factors, showed that a family history of CAD increases the risk for CAD by 1.5-fold. Several studies including Prospective Cardiovascular Münster Study (PROCAM) showed that a family history of MI was an independent risk factor of CAD. These findings were confirmed in a more recent study. In most of these studies, the phenotype for CAD is a cardiovascular event, such as MI or death from cardiovascular causes.

**Evolution of Human DNA Variation**

Life on earth is currently documented to have begun ≈3.8 billion years ago. The evolution of the human genus (homo) began 6 million years ago when it separated from the ancestral tree it shared with the upper primates, particularly the chimpanzee. Several species of the genus homo evolved with homo erectus, developing ≈2 million years ago and quickly spreading throughout the planet. Between 200000 to 300000 years ago, the homo sapiens emerged, referred to as the modern human species, and migrated out of Africa between 120000 and 150000 years ago.

Throughout this long evolutionary interval, despite most DNA sequences being similar across genii and species, sequences unique to each species were maintained. Throughout the homo sapiens species, the DNA sequence is 99.5% identical with 0.5% or 15 million base pairs being different. Most of these differences are because of large chunks of DNA from chromosomal rearrangement, duplication (copy number variation), and translocations. Based on the current information, most of the human DNA variation, whether it is responsible for hair color or predisposition to disease, is because of 3 million SNPs contained in each human genome. These SNPs are distributed throughout the genome, and their effect on the phenotype may be neutral, relatively benign, or markedly influential. Most of the DNA variation is created by DNA copying errors during replication of the 2 strands of DNA. DNA turns over every few days with <1 error per billion bases created; nevertheless, this error over thousands of generations is enough to induce a significant number of mutations, which when occur in the egg or sperm can be transmitted to the next generation. Ninety-four percent of the copy errors are substitutions of a single nucleotide and 4.5% are because of deletion of 1 to 4 bp and the remainder because of insertion of 1 to 4 bp. Kruglyak and Nickerson estimated a mutation rate of $2 \times 10^{-8}$ per base pair per generation, whereas more recent estimates by Sun et al. claim a mutation rate of $1.4 \times 10^{-8}$. The latter mutation rate would amount to 40 new mutations per generation, and with the world’s population of 7 billion, it results in ≈400 billion new mutations in the current generation. This biodiversity provides the fodder for an intense research effort to elucidate the function of these mutations and their association with disease. The biodiversity is greater than it may seem. Although each human genome has only 3 million SNPs to account for its variation, they are selected from well >17 million SNPs circulating in the general population. Mutations including SNPs evolve according to Darwinian principles. Those that provide an advantage to the species are retained and with each subsequent generation become increasingly more frequent, whereas those having deleterious consequences are eliminated or their frequency is kept to a minimum. Included in the latter would be these rare deleterious mutations that induce single-gene disorders.

**Discovery of Multiple Novel Genetic Risk Variants for CAD Acting Primarily Through Unknown Pathways**

The discovery of the 9p21 risk variant in 2007 was soon followed by several international groups discovering other genetic risk variants for CAD, resulting in an additional 12 genetic risk variants for CAD, all of genome-wide significance and replicated in independent populations. It was recognized by the GWAS groups that to detect risk variants of low frequency and minimal effect would require even larger sample sizes than initially anticipated. This led to the formation of Coronary Artery Disease Genome-Wide Replication and Meta-Analysis (CARDIoGRAM), which united investigators from the United States, United Kingdom, Canada, Germany, and Iceland—the largest collaborative study ever in cardiology. The sample size was initially 88,000, and a meta-analysis led to the confirmation of 10 of the previously discovered genetic risk variants plus a discovery of 13 novel risk variants for CAD. The subsequent addition to CARDIoGRAM of several other GWAS groups pursuing CAD risk variants, such as CAD, led to a name change to CARDIoGRAMplusC4D with a sample size >240,000 individuals. Meta-analysis of this larger sample size led to a discovery of 15 novel genetic risk variants for CAD, with confirmation of 31 previously discovered risk variants for a total of 46 genetic risk variants associated with CAD. Wang et al. identified a genetic variant at 6p21, which increases the risk for CAD in the Chinese Han population; however, it has no risk effect in the white population. The international consortium, IBC 50K (ITMAT-Broad Care array), identified 3 additional genetic risk variants for CAD. All of these risk variants exhibited genome-wide significant association with CAD and were confirmed.
A total of 50 genetic risk variants for CAD are indicated in the Table.

The genetic risk variants for CAD as for other polygenic disorders have several features in common. All of the genetic risk variants for CAD are common, with 50% occurring in more than half of the population and 25% occurring in >75% of the population. Their relative risk is usually minimal to moderate, averaging only 18% increased relative risk with odds ratio varying from 2% to 90%. The majority of the genetic risk variants are located in noncoding DNA sequences, implying the risk is mediated through regulation of downstream or upstream protein-coding sequences, with possibility of affecting protein-coding genes on other chromosomes. The DNA genetic risk variants need only be assessed once because they do not change in an individual’s lifetime and they do not vary with time, nutrition, drugs, or sex.

The most striking feature of genetic variants for CAD, illustrated in the Table, is the observation that 35 of the 50 genetic risk variants act through mechanisms independent of known risk factors, such as cholesterol, blood pressure, diabetes mellitus, or C-reactive protein. This observation has many implications, notably several pathways contributing to the pathogenesis of CAD are yet to be appreciated. The discovery of a genetic risk variant in the major histocompatibility locus at 6p21.3 would strongly indicate that it acts through the innate immune system to enhance the vascular inflammatory response.43 There are other potential candidates that have been shown to be part of the inflammatory network, including the interleukin 6 receptor, CXCL12, mitochondrial ribosomal protein S6, and muscle RAS oncogene homolog protein. To prove that their risk is mediated through inflammation will require confirmation by more direct in vivo and in vitro studies.44 Comprehensive prevention of CAD will require development of drugs targeted to these pathways.

Indicated in the Table are 7 genetic variants mediating their risk through LDL cholesterol (LDL-C). Most of these are in genes well recognized to be associated with LDL-C, including lipoprotein-a, apolipoprotein-B, LDL-R, apolipoprotein-E, and ABCG5. However, the Sortilin gene (SORT1) on 1p13.3 represents a novel pathway for regulating LDL-C. The non-coding polymorphism creates a C/EBP (CAAAT enhancer-binding protein) transcription factor–binding site, which alters the hepatic expression of SORT1.45 With knockdown studies in the mouse, they demonstrated that SORT1 alters plasma LDL-C by modulating very low-density lipoprotein secretion. This represents a novel regulatory pathway for lipoprotein metabolism and also provides targets for the development of novel therapy.46 Discovery of the PCSK9 variant located on 1p32.3 has already led to the development of a new therapeutic agent complementary to statins, which will be discussed below. The discovery of 2 risk variants associated with triglycerides adds support to the emerging data indicating that triglyceride is a risk factor for CAD (Table).

9p21: The First Genetic Risk Variant for CAD Whose Function Is Unknown,Profiles Most Genetic Risk Variants

The first genetic risk variant for CAD was discovered on the small arm p of chromosome 9 at band 21.3 and is now referred to as 9p21, and the features typify those of other risk variants to come, not just for CAD but for most diseases. The features observed by our group47 were identical to those found by the Icelandic group,48 despite being simultaneously and independently discovered. The frequency of 9p21 was as expected common, with 25% of the population having 2 copies and 50% having 1 copy. The increased relative risk for CAD was ≈25% for 1 copy and 50% for 2 copies. In individuals with premature CAD, the relative risk for CAD associated with 9p21 is increased 2-fold.49 9p21, because of its frequency, accounts for 21% of the general population attributable risk and 31% of population attributable risk in individuals with premature onset of CAD.46 Both studies showed that 9p21 is independent of known risk factors (eg, cholesterol) and acting through some novel unknown pathway. The 9p21 risk variant is not associated with C-reactive protein.45–48 All of these studies were performed in whites. The 9p21 risk variant was confirmed in the Wellcome Trust group49 and around the world in diverse ethnic groups, including Korean,50 Japanese,50 Italian,51 and East Indians.52 However, 9p21 does not seem to be associated with risk for CAD in blacks.53 In these studies, as in a more recent study,54 the 9p21 risk variant is consistently shown to act independently of known risk factors.

Fine mapping of the 9p21 risk region shows that it is confined to a 58,000-bp region with several SNPs showing significant association with CAD, all in linkage disequilibrium with rs4973754.53 The 58,000-bp region does not contain any protein-coding genes, but does overlap with a long noncoding RNA of 126,000 bp referred to as antisense noncoding in the INK4 locus (ANRIL) (Figure 1).56,57 ANRIL has no known function, but its expression indicates the expression of 3 nearby genes, CDKN2a, CDKN2b, and ARF, that encode for Cyclin-dependent kinase inhibitors. CDKN2a and CDKN2b encode for proteins that inhibit cell cycle through inhibition of the retinoblastoma pathway.57 This represents an arrest of the cell cycle in the G1 phase, resulting in decreased cell proliferation and growth. A third kinase inhibitor, ARF, is an alternate reading frame of the CDKN2a gene, which inhibits the cell cycle by either the retinoblastoma pathway or the p53 pathway arresting the cell cycle in G1 and G2 phases, and also inhibits apoptosis.58 ARF does this by degrading MDM2, a protein common to both pathways. All of these cell cycle kinase inhibitors are found in many tissues, including coronary vasculature, adipose tissue, and circulating leukocytes.59–62 The 9p21 risk variant may decrease the expression of ANRIL, which in turn results in reduced expression of the kinase inhibitors enabling cell proliferation and growth along with increased apoptosis.59 The 9p21 variant is a risk factor not only for CAD but also for abdominal aortic aneurysms, intracranial aneurysms,63 and large-vessel stroke.64 Although there may be some atheroma associated with abdominal aortic aneurysms, athromia is not associated with intracranial aneurysms.65 The inhibition of apoptosis by ARF is particularly interesting in view of a recent observation that in a CDKN2a-knockout mouse, aortic aneurysms developed with thinning of the wall and increased smooth muscle proliferation.66 This would imply that 9p21 leads to a defect in the vessel wall, making it prone to atheroma formation or dilatation.
### Table: Chronological List of 50 Genetic Variants (Genome-Wide Significant) Associated With Coronary Artery Disease or MI

<table>
<thead>
<tr>
<th>Chromosomal Location</th>
<th>SNP</th>
<th>Nearby Genes (Allele)</th>
<th>Risk Allele Frequency</th>
<th>Odds Ratio (95% CI)</th>
<th>Year Discovered</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Risk variants associated with LDL cholesterol</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6q25.3</td>
<td>rs3798220</td>
<td>LPA</td>
<td>0.02 (C)</td>
<td>1.92 (1.48–2.49)</td>
<td>2009</td>
</tr>
<tr>
<td>2p24.1</td>
<td>rs515135</td>
<td>APOB</td>
<td>0.83 (G)</td>
<td>1.03</td>
<td>2012</td>
</tr>
<tr>
<td>1p13.3</td>
<td>rs599839</td>
<td>SORT1</td>
<td>0.78 (A)</td>
<td>1.29 (1.18–1.40)</td>
<td>2007</td>
</tr>
<tr>
<td>19p13.2</td>
<td>rs1122608</td>
<td>LDLR</td>
<td>0.77 (G)</td>
<td>1.14 (1.09–1.19)</td>
<td>2009</td>
</tr>
<tr>
<td>19q13.32</td>
<td>rs2075650</td>
<td>APOE</td>
<td>0.14 (G)</td>
<td>1.14 (1.09–1.19)</td>
<td>2011</td>
</tr>
<tr>
<td>2p21</td>
<td>rs6544713</td>
<td>ABCG5-ABC8</td>
<td>0.29 (G)</td>
<td>1.07 (1.04–1.11)</td>
<td>2011</td>
</tr>
<tr>
<td>1p32.3</td>
<td>rs11206510</td>
<td>PCSK9</td>
<td>0.82 (T)</td>
<td>1.15 (1.10–1.21)</td>
<td>2009</td>
</tr>
<tr>
<td><strong>Risk variant associated with HDL cholesterol</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6p21.31</td>
<td>rs12205331</td>
<td>ANKS1A</td>
<td>0.81 (C)</td>
<td>1.04</td>
<td>2012</td>
</tr>
<tr>
<td><strong>Risk variants associated with triglycerides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8q24.13</td>
<td>rs10808546</td>
<td>TRIB1</td>
<td>0.65 (A)</td>
<td>1.08 (1.04–1.12)</td>
<td>2011</td>
</tr>
<tr>
<td>11q23.3</td>
<td>rs964184</td>
<td>ZNF259, APOA5-A4-C3-A1</td>
<td>0.13 (G)</td>
<td>1.13 (1.10–1.16)</td>
<td>2011</td>
</tr>
<tr>
<td><strong>Risk variants associated with hypertension</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12q24.12</td>
<td>rs3184504</td>
<td>SH2B3</td>
<td>0.44 (T)</td>
<td>1.13 (1.08–1.18)</td>
<td>2009</td>
</tr>
<tr>
<td>10q32.32</td>
<td>rs12413409</td>
<td>CYP17A1, CNNM2, NT5C2</td>
<td>0.89 (G)</td>
<td>1.12 (1.08–1.16)</td>
<td>2011</td>
</tr>
<tr>
<td>4q31.1</td>
<td>rs7692387</td>
<td>GUCY4A</td>
<td>0.81 (G)</td>
<td>1.13</td>
<td>2012</td>
</tr>
<tr>
<td>15q26.1</td>
<td>rs17514846</td>
<td>FURIN-FES</td>
<td>0.44 (A)</td>
<td>1.04</td>
<td>2012</td>
</tr>
<tr>
<td><strong>Risk variant associated with myocardial infarction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9q34.2‡</td>
<td>rs579459</td>
<td>ABO</td>
<td>0.21 (C)</td>
<td>1.10 (1.07–1.13)</td>
<td>2011</td>
</tr>
<tr>
<td><strong>Risk variants (mechanism of risk unknown)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9p21.3</td>
<td>rs4977574</td>
<td>CDKN2A, CDKN2B</td>
<td>0.46 (G)</td>
<td>1.25 (1.18–1.31) to 1.37 (1.26–1.48)</td>
<td>2007</td>
</tr>
<tr>
<td>1q41</td>
<td>rs17465637</td>
<td>MIA3</td>
<td>0.74 (C)</td>
<td>1.20 (1.12–1.30)</td>
<td>2007</td>
</tr>
<tr>
<td>10q11.21</td>
<td>rs1746048</td>
<td>CXCL12</td>
<td>0.87 (C)</td>
<td>1.33 (1.20–1.48)</td>
<td>2007</td>
</tr>
<tr>
<td>2p33.1</td>
<td>rs6725887</td>
<td>WDR12</td>
<td>0.15 (C)</td>
<td>1.16 (1.10–1.22)</td>
<td>2009</td>
</tr>
<tr>
<td>6p24.1</td>
<td>rs12526453</td>
<td>PHACTR1</td>
<td>0.67 (C)</td>
<td>1.13 (1.09–1.17)</td>
<td>2009</td>
</tr>
<tr>
<td>21q22.11</td>
<td>rs9982601</td>
<td>MRPS6</td>
<td>0.15 (T)</td>
<td>1.19 (1.13–1.27)</td>
<td>2009</td>
</tr>
<tr>
<td>3p22.3</td>
<td>rs2306374</td>
<td>MRA5</td>
<td>0.18 (C)</td>
<td>1.15 (1.11–1.19)</td>
<td>2009</td>
</tr>
<tr>
<td>10p11.23</td>
<td>rs2505083</td>
<td>KAA1462</td>
<td>0.42 (C)</td>
<td>1.07 (1.04–1.09)</td>
<td>2010</td>
</tr>
<tr>
<td>1p32.2</td>
<td>rs17114036</td>
<td>PPAP2B</td>
<td>0.91 (A)</td>
<td>1.17 (1.13–1.22)</td>
<td>2011</td>
</tr>
<tr>
<td>5q31.1</td>
<td>rs2706399</td>
<td>IL5</td>
<td>0.48 (A)</td>
<td>1.02 (1.01–1.03)</td>
<td>2011</td>
</tr>
<tr>
<td>6q23.2</td>
<td>rs12190287</td>
<td>TCF21</td>
<td>0.62 (C)</td>
<td>1.08 (1.06–1.10)</td>
<td>2011</td>
</tr>
<tr>
<td>7q22.3</td>
<td>rs10953541</td>
<td>BCAP29</td>
<td>0.75 (C)</td>
<td>1.08 (1.05–1.11)</td>
<td>2011</td>
</tr>
<tr>
<td>7q32.2</td>
<td>rs11556924</td>
<td>ZC3HC1</td>
<td>0.62 (C)</td>
<td>1.09 (1.07–1.12)</td>
<td>2011</td>
</tr>
<tr>
<td>10q32.31</td>
<td>rs1412444</td>
<td>LIPA</td>
<td>0.34 (T)</td>
<td>1.09 (1.07–1.12)</td>
<td>2011</td>
</tr>
<tr>
<td>11q22.3</td>
<td>rs974819</td>
<td>PDGF</td>
<td>0.29 (T)</td>
<td>1.07 (1.04–1.09)</td>
<td>2011</td>
</tr>
<tr>
<td>13q34</td>
<td>rs4773144</td>
<td>COL4A1, COL4A2</td>
<td>0.44 (G)</td>
<td>1.07 (1.05–1.09)</td>
<td>2011</td>
</tr>
<tr>
<td>14q32.2</td>
<td>rs2895811</td>
<td>HHIP1</td>
<td>0.43 (C)</td>
<td>1.07 (1.05–1.10)</td>
<td>2011</td>
</tr>
<tr>
<td>15q25.1</td>
<td>rs3825807</td>
<td>ADAMTS7</td>
<td>0.57 (A)</td>
<td>1.08 (1.06–1.10)</td>
<td>2011</td>
</tr>
<tr>
<td>17p13.3</td>
<td>rs216172</td>
<td>SMG6, SRR</td>
<td>0.37 (C)</td>
<td>1.07 (1.05–1.09)</td>
<td>2011</td>
</tr>
<tr>
<td>17p11.2</td>
<td>rs12936587</td>
<td>RASD1, SMCR3, PEMT</td>
<td>0.56 (G)</td>
<td>1.07 (1.05–1.09)</td>
<td>2011</td>
</tr>
<tr>
<td>17q21.32</td>
<td>rs46522</td>
<td>UBE2Z, GIP, ATP5G1, SNF8</td>
<td>0.53 (T)</td>
<td>1.06 (1.04–1.08)</td>
<td>2011</td>
</tr>
<tr>
<td>5p13.3*</td>
<td>rs11748327</td>
<td>IRX1, ADAMTS16</td>
<td>0.76 (C)</td>
<td>1.25 (1.18–1.33)</td>
<td>2011</td>
</tr>
<tr>
<td>6p22.1†</td>
<td>rs6929846</td>
<td>BTN2A1</td>
<td>0.06 (T)</td>
<td>1.51 (1.28–1.77)</td>
<td>2011</td>
</tr>
<tr>
<td>6p24.1†</td>
<td>rs6903956</td>
<td>G6b1f105</td>
<td>0.07 (A)</td>
<td>1.65 (1.44–1.90)</td>
<td>2011</td>
</tr>
<tr>
<td>6p21.3</td>
<td>rs3869109</td>
<td>HCG27 and HLA-C</td>
<td>0.60 (C)</td>
<td>1.15</td>
<td>2012</td>
</tr>
<tr>
<td>1q21</td>
<td>rs4845625</td>
<td>IL6R</td>
<td>0.47 (T)</td>
<td>1.09</td>
<td>2012</td>
</tr>
<tr>
<td>Chr4</td>
<td>rs1878406</td>
<td>EDNRA</td>
<td>0.15 (T)</td>
<td>1.09</td>
<td>2012</td>
</tr>
</tbody>
</table>

(Continued)
ANRIL is necessary to recruit the polycomb repression complex 2, which trimethylates the histone H3 at lysine 27 and suppresses the expression of CDKN2a and CDKN2b. These genes occupy a 42,000-bp region on human chromosome 9p21 adjacent to and overlapping with the ANRIL sequence. This region is extensively studied because of the possible link between cancer and the antiproliferative properties of CDKN2a and ARF. These genes are well recognized as tumor suppressor genes, and their expression is known to be inhibited in 40% of human cancers.

Thus, an attractive hypothesis is that the 9p21 risk variant reduces the expression of ANRIL and the kinase inhibitors, which enables the growth of vascular atheroma. Expression studies in cell culture showed the expression of the 9p21 risk variant to be associated with decreased expression of CDKN2a and ARF. Liu et al. showed that the 9p21 risk variant was associated with decreased expression of CDKN2a and ARF in purified peripheral blood T cells obtained from 170 individuals. In the same studies, they found no association between the expression of 9p21 risk variants and that of methylthioadenosine phosphorylase. They concluded that the risk of 9p21 is through decreased expression of the cyclin-dependent kinase inhibitors CDKN2a or CDKN2ab. Other studies also support the hypothesis that 9p21 mediates its effect through suppression of the expression of CDKN2a and ARF, but unfortunately, other in vitro and in vivo studies do not confirm this hypothesis.

Holdt et al. showed a strong correlation between the expression of ANRIL and atherosclerosis in peripheral monocytes, but showed no correlation with the expression of CDKN2a and ARF, or methylthioadenosine phosphorylase. Other studies also showed a lack of correlation between the 9p21 risk variant and the expression of the Cyclin-dependent kinase inhibitors. The hypothesis remains viable because of the large number of supporting studies, despite the contradictory results of other studies.

Visel et al. pursued in vivo studies of 9p21 in a mouse model. The 58,000-bp region in humans corresponds to a 70,000-bp region in the mouse. They confirmed that in the mouse, as in humans, there are several promoter sequences in the corresponding region. Through a complicated sequential analysis, the whole 70 kbp region was knocked out and the resulting phenotype carefully analyzed. They confirmed that the region had no effect on plasma lipids, indicating that the 9p21 region is independent of plasma lipids as observed in human studies. They also observed excessive proliferation of smooth muscle cells and several neoplasms. These studies confirm that the 9p21 region is a regulator of the Cyclin-dependent kinase inhibitors in the mouse as in humans. However, despite the smooth muscle proliferation, the mice did not develop fatty streaks or any evidence of coronary atherosclerosis.

The absence of atherosclerosis may, in part, be because of the mouse’s resistance to atherosclerosis or the lack of the 9p21 variant, given there is only 50% sequence homology. Comparative DNA analysis suggests that the 9p21 risk variant did not appear in the genome until that of higher primates and thus would not be expected to be present in the mouse.

Harismendy et al. performed elegant studies suggesting an alternative mechanism whereby 9p21 mediates its risk for CAD through interferon-γ. The interferons are a family of cytokines that include type 1 (interferon-α and -β), type 2 (interferon-γ), and type 3 encoded by a cluster of genes located on chromosome 9, <1 million base pairs from the 9p21 risk region. This cluster of genes encodes for 13 isoforms of interferon-1, including interferon-γ and interferon-α-21. Type 1 interferon is part of the innate immune response associated with autoimmune and inflammatory diseases, such as systemic lupus erythematosus. The DNA sequence of the 9p21 risk region forms an active domain through the formation of a chromatin loop, which approximates the interferon region to that of the 9p21 risk region, providing the potential for 9p21 to regulate the expression of interferon genes. They observed a SNP in a STAT1-binding site located in the 58,000-bp DNA region of 9p21. Because interferon mediates its effect on inflammation, in part through the STAT1 sequence, the investigators postulated that interferon-γ may mediate the risk of 9p21 for CAD. However, we subsequently showed that the effects of interferon-γ on CDKN2a and ARF are independent of the 9p21 risk variant, and more recently, Erridge et al. showed that the effects of type 1 interferon, including interferon-α-21, are not associated with the 9p21 risk variant. The molecular mechanism whereby 9p21 mediates its risk for CAD remains unknown.
Although the molecular mechanisms of 9p21 remain unknown, there is strong evidence to indicate that it is proatherogenic and its site of action is at the vessel wall. The deCode group observed that the 9p21 risk variant is associated with increased risk for abdominal aortic aneurysms and intracranial aneurysms. The 9p21 risk variant, although strongly associates with coronary atherosclerosis, shows no genetic association with MI. Because MI is usually due to plaque rupture and thrombosis, superimposed on coronary atherosclerosis, it would suggest that 9p21 is not associated with plaque rupture or thrombosis. Multiple studies consistently show that the 9p21 risk variant is associated with coronary atherosclerosis and not MI. Several studies also indicate that 9p21 risk variant is associated with the progression of coronary atherosclerosis based on the number of vessels involved (Figure 2) as well as the Gensini and Duke scores for severity of coronary atherosclerosis. The most recent of these studies is a large meta-analysis by Chan et al., involving 33,673 subjects (cases and controls). Nevertheless, there are a few other studies, albeit with small sample sizes, showing that 9p21 is not associated with progression, but simply the presence of coronary atheroma. All of these studies have been cross-sectional, and it may require a longitudinal study to provide a definitive solution to whether 9p21 is associated with the progression of atherosclerosis. Studies to further delineate the molecular mechanism of 9p21 should focus on the various factors contributing to the pathogenesis of atherosclerosis, such as macrophages, adhesion molecules, and inflammation. Unfortunately, the mouse may not be the appropriate animal model, because the 9p21 risk variant seems to have entered the genome until the evolution of higher primates.

**Genetic Risk Variants and MI**

In all of the GWAS, the phenotypes of MI and CAD are treated as synonymous. This is justified on the basis that MI is precipitated by a thrombus, superimposed on significant coronary atherosclerosis. The thrombus is precipitated by the rupture of an atheromatous plaque. Nevertheless, the genetic variants contributing to thrombosis, such as platelet adhesion, or release of tissue coagulating factors, or the triggering plaque rupture, are most certainly distinct from those contributing to coronary atherosclerosis by mechanisms such as monocyte adhesion, cholesterol deposition, or release of inflammatory cytokines. In the Ottawa Heart Genomics Study, most of the cases selected for the GWAS also underwent coronary angiography. Analysis of the data showed that 9p21 is associated with both coronary atherosclerosis and MI, but the association is mainly accounted for atherosclerosis rather than MI. These findings were confirmed in similar studies by Reilly et al.

It is somewhat surprising that the only genetic risk variant currently shown to be related to acute coronary syndrome, namely MI, is at the ABO blood group locus. It is of note that epidemiological studies have suggested a strong association between the ABO blood groups at 9q34.2 and MI. This association was confirmed in CARDIoGRAM, showing that the A and B risk variants increase the risk for MI by ~20%. Blood groups A and B encode for a protein (α-1-3-N-acetylgalactosaminyltransferase), which transfers a carbohydrate onto von Willebrand Factor, prolonging the half life of von Willebrand Factor with predisposition to coronary thrombosis and MI. Plasma levels of von Willebrand Factor are ~25% higher in individuals with A, B, or AB blood groups compared with blood group O. The blood group O gene also code for the same transferase protein, but because of a mutation lacks any biochemical activity. In individuals with blood group O, the half life of von Willebrand Factor is not prolonged, and neither there is any increased risk for thrombosis or MI. In the recent Nurse’s Health Study of >90,000 individuals, 4070 developed heart disease. In this 20-year follow-up study, blood group A or B was associated with a 10% increased frequency of MI; however, the combination of A and B blood groups increased the risk to 20%.

**Conventional Risk Factors for CAD Such as Plasma Lipids Are Also Under Intense Genetic Regulation**

In assessing the genetic risk for CAD, it must also be realized that the conventional risk factors such as plasma lipids are highly regulated by genetic variants. It is well accepted from different sources that plasma levels of high-density lipoprotein cholesterol (HDL-C) and LDL-C are 60% to 80% determined by genetic variants. It is known that 70% to
80% of cholesterol is synthesized endogenously under the control of many genetic variants. CARDioGRAM formed a collaboration with the Lipids Genetics Consortium using the GWAS technique to pursue the genetic variants regulating plasma lipid levels. This provided a total sample size of >100,000 individuals who were genotyped for genes related to plasma levels of total cholesterol, LDL-C, HDL-C, and triglycerides and identified a total of 95 associated genetic variants. In a follow-up study, a similar GWAS was performed by the Global Lipids Genetics Consortium involving a sample size of 188,577 individuals. This study identified 62 novel genetic variants associated with the regulation of plasma lipid levels. The results of these 2 studies provided a total of 157 genetic variants associated with plasma lipid levels: 55 associated with plasma HDL-C, 37 associated with plasma LDL-C, 54 associated with total plasma cholesterol levels, and 24 associated with plasma triglyceride levels.

These studies emphasize the genetic endogenous control of plasma lipid levels and their potential importance in the pathogenesis of CAD. In considering the broad picture of genetic risk for CAD, ultimately, it must also include those genetic risk factors involved with known conventional risk factors as well as novel ones identified to be associated with CAD. Clinicians are fully aware of the difficulty in lowering LDL-C in a significant percentage of patients despite high doses of statins. It is highly likely that these individuals have an excess of genetic variants that, through either increased synthesis or decreased degradation of LDL-C, make them less responsive to statin therapy. Such information is not yet available, and whether it will influence overall management and prevention of CAD remains to be determined. Similar genetic risk variants are also being identified for conventional risk factors such as hypertension and diabetes mellitus.

Current Benefits From the Discovery of Genetic Variants Associated With CAD

Inhibition of PCSK9: A New Potent and Safe Therapy for CAD Is Evolving

A novel therapy based on the inhibition of PCSK9 is evolving and being assessed in clinical trials. Evidence supporting cholesterol as a risk factor for CAD has been around for several decades. However, the finding of a mutation in LDL-R in the 1970s by Brown and Goldstein catalyzed the development of statin therapy. This genetic defect in affected individuals led to MI occurring in the second or third decades of their life. Within 20 years of this observation, the first drug to inhibit cholesterol synthesis was developed, which has become the main drug for the prevention and treatment of CAD, having a world budget of >$70 billion. The discovery of mutations in the PCSK9 gene has catalyzed a similar movement leading to a new therapy for reducing LDL-C and prevention of CAD.

In 2003, Seidah et al discovered an enzyme, now called PCSK9, which binds to and increases the degradation of LDL-R. This leads to increased plasma levels of LDL-C and increased incidence of MI and CAD. There are 9 genes encoding PCSK proteins (1–9). They belong to a family of serine proteases called proprotein convertase, and their chief function is to activate precursors such as growth factors and hormones. The proprotein convertases were discovered in relation to bacterial subtilisin and yeast kexin, hence the name PCSK. PCSK9, a glycoprotein with 692 amino acids, circulates in the blood. The gene encoding PCSK9 is located on the short arm of chromosome 1 (1p32). PCSK9 is expressed in several organs, particularly the liver, with the circulating blood levels derived almost exclusively from the liver. PCSK9 binds to the EGF-a (epidermal growth factor) domain of LDL-R, and the resulting LDL-R/PCSK9 complex increases lysosomal degradation through mechanisms unknown. Abifadel et al discovered a gain-of-function mutation in PCSK9, which is associated with hypercholesterolemia and increased incidents of CAD. This was followed by a discovery of loss-of-function mutation in blacks, with a frequency of 2.6%. This beneficial mutation was associated with a 28% reduction in the mean plasma LDL-C and an 88% reduction in risk of CAD. This emphasizes a protection from a lifelong exposure to a relatively minor reduction in plasma LDL-C levels. Furthermore, the lifelong exposure does not seem to be associated with any adverse effects.

It is self-evident that the plasma level of LDL-C depends on the rate of production and the rate of removal. The production of LDL-C is from the liver, and its removal is also by the liver, primarily through uptake by LDL-R, which then undergoes degradation in the liver (Figure 3). There is good evidence that PCSK9 increases plasma LDL-C by binding to LDL-R, which leads to its degradation. Normally, LDL-R after binding to LDL-C internalizes and transfers LDL-C for degradation and (150x) recycles to the surface, which is prevented by PCSK9. Inhibition of PCSK9 represents a novel approach to lowering the plasma levels of LDL-C, and the most successful has been through a monoclonal antibody. Clinical phase I and II studies have been completed, and they observed a 41% to 58% reduction in plasma LDL-C. In 1 of the studies, a further reduction in plasma LDL-C of 72% was observed despite being on atorvastatin 80 mg. Inhibition of PCSK9 increases the removal of LDL-C from the plasma and is complementary to the effect of statin therapy, which decreases its production. This is important because most individuals with increased plasma LDL-C do not achieve appropriate lowering with statins, in part because of resistance or side effects. In the United Kingdom, only 28% of individuals receiving statin therapy achieved the desired targeted level of plasma LDL-C. Thus, inhibition of PCSK9 provides a novel therapy that is much needed in the prevention of CAD, the world’s number one killer.

Mendelian Randomization: A Novel Approach to Assess Safety and Efficacy of Genetic Variants and Their Potential as Drug Targets

The standard method required by drug agencies, such as the US Food and Drug Administration, to determine the safety and efficacy of drugs is the randomized controlled trial (RCT). RCT is performed in a predetermined adequate sample size, usually for a duration of 3 to 5 years. A new approach, referred to as Mendelian randomization, can be helpful and complementary to the RCT. The basis for Mendelian randomization is obvious. The genetic variants of an individual are randomly inherited at birth, remaining fixed throughout life, with no confounding factors. Their effects can be...
observed after decades of exposure rather than a few years as observed in RCTs. Observations in large populations are now available from GWAS, and comparison of phenotypes expressed in individuals with and without the genetic risk variant has perhaps less confounding effects than the placebo controlled trial. In blacks with loss-of-function mutation in PCSK9, the plasma LDL-C was reduced by only 28%, but the decreased risk from CAD was 88%. This observation indicated that loss of function of PCSK9 is safe and beneficial and thus a good target around which to develop drug therapies. The recent development of a monoclonal antibody inhibiting PCSK9 confirms the results of Mendelian randomization. Phase I and II clinical trials have shown the agent to be safe and effective with phase III clinical trials currently ongoing. Second, the disparity between the reduction in plasma LDL-C and the marked reduction of 88% in events from CAD would not be expected to be observed in a RCT of 3 to 5 years. The likely explanation for this disparity between the effect on LDL-C and benefit observed in Mendelian randomization studies is because of the constant exposure of individuals throughout their lifetime, rather than the short interval of the RCT. The Mendelian randomization approach is becoming a popular method to assess safety and benefit of genetic risk variants before development of drug therapy as well as providing complementary information on drugs undergoing RCTs.

The Mendelian randomization method may also be helpful in situations in which RCT has other confounding factors. The nonconfounding nature of Mendelian randomization is well elucidated in our recent study assessing the effect of genetic variants related to plasma HDL-C levels. To perform this study, we selected 3 categories of plasma lipid–related genetic variants:

1. SNP in the endothelial lipase gene (LIPG; Asn396Ser)
2. 14 common SNPs associated solely with an increase in plasma HDL-C and for a positive control
3. 13 SNPs associated solely with increased levels of plasma LDL-C

The sample size for the LIPG variant was designed with 90% power to detect a 13% reduction in risk of MI, consisting of 50,763 patients of whom 4,228 developed MI. Analysis showed that the LIPG SNP was not associated with decreased risk of MI despite significantly elevated levels of HDL-C. The LIPG SNP occurs in 2.6% of the general population and is associated with an increase of HDL-C of 0.14 mmol/L per copy with no change in plasma LDL-C or triglycerides. A replication analysis was performed in an independent population of 16,685 cases of MI and 48,872 controls. This analysis showed no association of LIPG SNP with risk of MI. This was followed by a meta-analysis of both populations (20,913 cases and 95,407 controls), again showing no association of LIPG SNP with MI. A similar analysis was performed using 14 common SNPs associated solely with increased plasma HDL-C, and again, there was no association between these SNPs and MI. As a positive control, 13 common SNPs associated solely with increased levels of plasma LDL-C showed a marked increase in risk of MI with odds ratios of 2.13 as expected from previous studies. These studies indicate that being randomly assigned at birth with a lifelong exposure to increased plasma HDL-C is not associated with any decrease in risk for MI. Although the Mendelian randomization minimizes the effects of LDL-C from other forms of interventions, we cannot exclude the possibility that other pleiotropic effects from these genetic variants may be occurring that are counterproductive to the postulated potential benefits of HDL-C. Several recent studies including RCTs suggest that plasma HDL-C levels are not associated with MI or CAD. The original observation by Gofman et al in 1966 that HDL-C is protective was based on measurements of plasma HDL-C mass using ultracentrifugation. They observed that the plasma HDL-C level correlated with the mass measurements. Given that plasma HDL-C levels are simpler to perform, it was adopted as a surrogate for HDL-C mass. Several proteins make up the HDL particle, with the majority being because of apolipoprotein A1 (70%) and apolipoprotein A2 (20%) and the remaining percentage being a variety of other proteins. Thus, there is the possibility that ≥1 of these proteins may offer protection that is not captured by simply measuring the plasma levels of HDL-C. Nevertheless, the results of these Mendelian randomization studies along with recent clinical trials strongly indicate that our simple dogma that plasma HDL-C is protective is no longer acceptable.

Clinical Application of Genetic Risk Variants for the Prevention and Management of CAD

It is now proven from the discovery of genetic risk variants that many factors contribute to the pathogenesis of CAD. This
is emphasized by the observation that about two thirds of the genetic risk variants act through mechanisms independent of common known risk factors. Comprehensive prevention will have to await discovery of these pathways before one can effectively intervene. On the contrary, it would be faulty to hope for comprehensive prevention of CAD without this knowledge. It is also evident that the genetic risk burden is determined by how many of these risk variants are present rather than a particular variant. In the meantime, in assessing clinical risk, one must be cognizant of the concern for individuals with blood group A or B, who have in-dwelling artificial devices such as stents or pacemakers, of the potential additional need of antiplatelet therapy. The Nurse’s Health Study shows ≤20% increased risk associated with blood groups A and B.

One of the challenges in the management of CAD is to stratify individuals into groups according to the degree of risk. The Framingham score is the most commonly used approach. This score recognizes that a family history of CAD or its sequelae carries a significant increased risk. The availability of genetic risk variants for CAD would be expected to further improve risk stratification. However, despite several attempts using genetic risk variants, the results, although encouraging, have not been game changers. There have been many attempts to assess various configurations from the single genetic risk variant 9p21 to combinations of ≤13 variants. All of these studies show a good correlation between the genetic risk variants and the risk of CAD with some improvement in the prediction of CAD. However, the conventional criteria for significant improvement is judged on whether it changes the net reclassification index or improves discrimination by the conventional ethics, referred to as C-index. Neither of these criteria was satisfied by current attempts using genetic risk variants. One attempt by Hughes et al did find significant improvement in the C-index using genetic risk variants, but only when restricted to middle-aged men. These attempts were reviewed by Di Angelantonio and Butterworth. They discuss several confounding factors in reaching any conclusion about their role in the prevention and management of CAD. It is important to recognize that 50 genetic risk variants for CAD have been identified, but these attempts use only ≤13 variants. Second, many of these studies have been cross-sectional and not longitudinal with less than adequate sample size. It is also noteworthy that attempts assessing whether they improve prediction of the incidence of MI may be inappropriate because almost all of these variants with the exception of the blood group ABO relate to coronary atherosclerosis. It is still premature to make a definitive conclusion as to the role of genetic risk variants in the prevention and management of CAD.

**So-Called Missing Heritability Phenomenon**

The expected heritability component of susceptibility for CAD is estimated to be at 40% to 60%. The absolute heritability of the 50 genetic risk variants account for only 15% to 20%. This gap between the expected and the actual has become known as the missing heritability phenomenon. It is of note that this missing heritability and its magnitude are similar for practically all of the diseases for which genetic risk has been identified by GWAS. This represents >2000 genetic risk variants involving >200 diseases. Is this gap a correct estimate, or is it confounded by other factors? Genetic variants identified by GWAS for height account for only 10% of the expected heritability similar to that of CAD. Visscher et al estimated that common SNPs account for >45% of human height. This estimate was based on using all of the SNPs (500000) rather than only those satisfying the stringent and arbitrary statistical correction of a \( P \leq 5 \times 10^{-8} \). They claim another significant part of the remaining heritability is because of linkage disequilibrium estimates. They point to the observation that a causal SNP will be in linkage disequilibrium with several other SNPs, none of which are causal. If the genotyping collects primarily the noncausal SNPs, a likely event if the causal SNP is less frequent than the noncausal, this would give falsely low numbers for heritability estimates. Thus, they conclude common genetic risk variants account for most of the missing heritability if one uses all of the SNPs, account for the linkage disequilibrium underestimates together with recognizing that a small proportion is because of rare variants (not detectable by GWAS). This conclusion has not yet been widely accepted. The other possibility that is in keeping with Visscher et al is that there are more common genetic risk variants yet to be discovered. It is also highly likely that our estimates of heritability attributed to the common genetic risk variants may be falsely low because of epistasis. It is reasonable to conclude that epistasis, the interaction among genes, produces additional synergistic effects on the phenotype that is much greater than simply a linear addition of the individual effects of genetic risk variants. We do not at this time have a method to accurately calculate the potential epistasis effect. It is also of note that epistasis effects are both positive and negative. Another possibility is that we may have overestimated the expected heritability for disease. Lastly, some have argued that the difference is because of rare variants. Rare variants with a frequency of 0.1% to 5.0% will not be detected by GWAS even with sample sizes >200000. It is unlikely that rare variants account for more than a minimal amount of the missing heritability. For example, a rare variant occurring with a frequency of 0.1% or 1.0% is present in only 1 of 10000 or 1 of 100000, respectively. Although the effects may be great, it is clearly present in too small a percentage of the general population to account for much of the expected heritability. Exome sequencing ongoing at present suggests that rare variants will not account for the missing heritability. In a large study involving 41911 individuals, exome sequencing was performed in the pursuit of rare variants affecting autoimmune disease. Results of this study indicate that the proportion of heritability because of rare variants was negligible. A recent study performed by crossing >2 strains of yeast in which the traits are known produced an interesting and relevant finding. They were able to account for 100% of the expected heritability through primarily common risk variants each exhibiting only minimal risk effects with gene-to-gene interactions varying from 0% to 50% of the effect. Currently, it is reasonable to speculate that the likely source of missing heritability is common genetic risk variants.
Discovery of more common genetic variants with more appropriate calculations will probably account for the missing heritability.

**Future Directions**

The success of GWAS in the past 6 years has catapulted the search for genetic risk for disease into a new era. CAD is a highly preventable disease as shown by results from clinical trials modifying known risk factors. Comprehensive prevention for CAD has been postulated for the 21st century with expectations that CAD will be markedly attenuated if not eliminated. For the prevention to be comprehensive as with many complex diseases, it must also prevent genetic risk. It is time to plan for a global assessment and develop biobanks with millions of samples from well-defined phenotypes. This would facilitate annotating millions of SNPs to their associated disease across all ethnic groups. There are 20 diseases responsible for 80% of all deaths in the world with heart disease and cancer accounting for two thirds of those deaths. It is appropriate and feasible with today’s technology in the next 10 years to annotate thousands of SNPs to these diseases. The second challenge is to annotate genetic risk variants to what I shall refer to as subphenotypes. CAD like many complex diseases has multiple factors contributing to its pathogenesis, namely monocytes, adhesion molecules, cholesterol oxidation, formation of bone cells, inflammatory cytokines, and many others. With existing technology and biobanking of millions of samples, one can obtain these phenotypes and relate them to their corresponding genotype. This will lead to more specific targets for the development of novel drugs.

Coupled with the success of GWAS and related discoveries and innovations, the pricing of DNA sequencing has decreased 1 million-fold, leading to the sequencing of the human genome within a week for <$1000. However, our ability to store, retrieve, and analyze data has only increased 16-fold. The data on human genes and the annotation of genetic risk variants to disease are proliferating, and with the inexpensive routine DNA sequencing, the public will demand its use and application. Physicians will have to be educated and the knowledge compacted to have clinical application. Recognizing this dilemma, the various societies and credentialing groups have convened to form the Inter Society Coordinating Committee for genomic physician education. The Inter Society Coordinating Committee has defined genetic clinical competencies and is developing a single genetic information Website, available to all physicians and allied health personnel. The next decade promises to be a tipping point for genetic application and personalized medicine.

**Acknowledgments**

I thank Peggy Offley and Heather Stevenson for their assistance in the preparation of the article.

**Sources of Funding**

Dr Roberts has received grant support from Canadian Institutes of Health Research, grant #MOP82810, and from Canada Foundation for Innovation, grant #11966.

**Disclosures**

Dr Roberts is a consultant to Cumberland Pharmaceuticals and confirms no conflicts.

**References**


84. Reilly MP, Li M, He J, et al; Myocardial Infarction Genetics Consortium; Wellcome Trust Case Control Consortium; Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) Investigators; Genetic Investigation of ANthropometric Traits (GIANT) Consortium; Asian Genetic Epidemiology Network—Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; 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;44:981–990.


97. Davies RW, Dandona S, Stewart AF, Chen L, Ellis SG, Tang WH, Hazen SL, Roberts R, McPherson R, Wells GA. Improved prediction...


Genetics of Coronary Artery Disease
Robert Roberts

Circ Res. 2014;114:1890-1903
doi: 10.1161/CIRCRESAHA.114.302692

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/114/12/1890