Review in Advance first posted online
on June 29, 2010. (Changes may
still occur before final publication
online and in print.)

# Genetics of Coronary Artery Disease

## Kiran Musunuru[1] and Sekar Kathiresan[2]

[1]Center for Human Genetic Research and Cardiovascular Research Center, Massachusetts General Hospital, Boston, Massachusetts

[2]Department of Medicine, Harvard Medical School, Boston, Massachusetts

Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts; email: skathiresan@partners.org

## Key Words

genetics, coronary artery disease, myocardial infarction, genomics, genome-wide association studies

## Abstract

Coronary artery disease and its clinical manifestations, including myocardial infarction, are heritable traits, consistent with a role for inherited DNA sequence variation in conferring risk for disease. Knowledge of the new sequence variations in the genome that confer risk has the potential to illuminate new causal biologic pathways in humans and to thereby further improve diagnosis and treatment. Here, we review recent progress in mapping genetic loci related to coronary disease and risk factor phenotypes, including plasma lipoprotein concentrations. Genome-wide linkage (in families) and association (in populations) studies have identified more than a dozen genetic loci related to coronary disease. A key challenge now is to move from mapping loci to pinpointing causal genes and variants, and to develop a molecular understanding of how these genes lead to coronary disease.

## INTRODUCTION

Coronary artery disease and its clinical manifestations, including myocardial infarction (MI), are leading causes of death and infirmity worldwide. Although family history is a well-established independent risk factor for coronary disease and MI, the heritability of the disease is not fully understood and is the subject of intense investigation.

In this article, we review the current understanding of the genetics of coronary disease and MI, as well as the genetics of risk factors for MI. Building on a substantive body of work over the past several decades, the last few years have witnessed a significant leap forward in the characterization of genetic loci related to cardiovascular traits.

## HERITABILITY

The traditional risk factors for MI include age, plasma lipid concentrations, blood pressure, use of tobacco, and presence of type 2 diabetes mellitus. Although not included in many cardiovascular risk prediction algorithms, such as the Framingham Risk Score, family history has also been established as an important risk factor for MI. Premature MI appears to have a particularly strong genetic component, with heritability as high as 63% documented in one study (63). This may indicate the importance of inherited risk factors for early presentation of MI, as opposed to acquired risk factors that may influence the natural history of disease later in life.

In perhaps the most rigorous analysis of the inheritance of MI risk to date, the Framingham Offspring Study found that individuals with at least one parent with premature cardiovascular disease (defined as age of onset <55 years in men, <65 years in women) experienced a significant increase in risk of suffering a cardiovascular event compared to individuals with no such family history: for men, an age-adjusted odds ratio of 2.6, and for women, an odds ratio of 2.3 (51). After adjustment for the other traditional cardiovascular risk factors, the odds ratio for men remained significant at 2.0, with

women having an odds ratio of 1.7, indicating that some of the heritability of MI risk is independent of established risk factors, with the remainder attributable to heritability of the various risk factors. As much as one-half of the interindividual variability in plasma lipid concentrations is due to inherited factors (33, 40, 58, 68, 71). Type 2 diabetes mellitus (4) and blood pressure (28, 48) also have substantial heritability.

These observations have motivated investigators to undertake a variety of studies to identify genes responsible for the heritability of MI and cardiovascular risk factors. Linkage analyses and candidate gene studies were the mainstay of this work for several decades and successfully identified a handful of loci unequivocally linked to MI, as well as several additional loci with a weaker level of evidence. More recently, the completion of the Human Genome Project and the International Haplotype Map Project (34, 35) has made it possible to perform genome-scale screens for common DNA sequence variants that are associated with phenotypes of interest, an approach termed genome-wide association (GWA). This technique has expanded our knowledge of the genetic basis for coronary disease.

## EARLY FINDINGS FOR CORONARY ARTERY DISEASE

### Mendelian Lipid Disorders

Plasma lipid concentrations, particularly low-density lipoprotein cholesterol (LDL-C), have long been established as an important contributor to coronary disease. A number of monogenic, or Mendelian, lipid disorders have been characterized in individuals or families, many of which have been linked to rare variants in genes that, upon subsequent functional analysis, emerged as important regulators of lipoprotein metabolism (**Table 1**) (70).

The prototypic example is familial hypercholesterolemia (FH) due to mutations in the the *LDL receptor* gene (*LDLR*). The disease was originally characterized in individuals with

**Table 1    Selected Mendelian lipid disorders**

| Disease | Gene(s) with rare variants | Phenotype |
|---|---|---|
| Autosomal dominant hypercholesterolemia | *LDLR, APOB, PCSK9* | Elevated LDL cholesterol; premature coronary artery disease |
| Autosomal recessive hypercholesterolemia | *LDLRAP1* | Elevated LDL cholesterol; premature coronary artery disease |
| Sitosterolemia | *ABCG5, ABCG8* | Elevated plant sterols, LDL cholesterol; premature coronary artery disease |
| Primary hypoalphalipoproteinemia | *APOA1* | Low HDL cholesterol |
| Tangier disease | *ABCA1* | Low HDL and LDL cholesterol |
| Type I hyperlipoproteinemia | *LPL, APOC2* | Elevated chylomicrons, triglycerides |
| Hepatic lipase deficiency | *LIPC* | Elevated triglycerides |
| Type III hyperlipoproteinemia | *APOE* | Elevated chylomicrons, VLDL particles, IDL particles, and triglycerides |
| LCAT deficiency (Norum disease, fish-eye disease) | *LCAT* | Increased free cholesterol in tissues |
| Cholesteryl ester transfer protein deficiency | *CETP* | Elevated HDL cholesterol |

Abbreviations: LDL, low-density lipoprotein; HDL, high-density lipoprotein; VLDL, very-low-density lipoprotein; IDL, intermediate-density lipoprotein.

autosomal dominant hypercholesterolemia, tendon xanthomas, and premature coronary artery disease (25). Individuals with homozygous FH have extremely high LDL-C levels and often develop symptomatic coronary disease in childhood or early adulthood. Analysis of fibroblasts derived from these patients revealed a defect in LDL uptake and led to the identification of the LDL receptor. Loss-of-function mutations in *LDLR* explain the dominant transmission of the disease within families (25). Subsequent linkage studies in other families with autosomal dominant hypercholesterolemia uncovered causal mutations in two other genes, *APOB* and *PCSK9* (1, 80).

Other Mendelian lipid disorders confer susceptibility to premature coronary disease, including autosomal recessive hypercholesterolemia, in which mutations in both alleles of *LDLRAP1* are needed to manifest high LDL-C levels (24), and sitosterolemia, in which plant sterols and cholesterol are dramatically elevated in the bloodstream due to mutations in either of the sterol transporters encoded by *ABCG5* or *ABCG8* (7). The link between several other Mendelian lipid disorders and coronary disease is more tentative, either because there are too

few affected individuals to establish a convincing relationship or because the resulting lipid changes (low LDL-C, high HDL-C, or low triglycerides) are expected to be atheroprotective, which is harder to document than a predisposition to disease.

## Myocardial Infarction and Coronary Disease

Although a number of genome-wide linkage studies for MI and for coronary disease have been performed and have identified putative chromosomal loci related to disease, the utility of many of these studies remains in doubt (9, 20, 65). For example, a linkage study in a family with 13 members over three generations with coronary disease, nine of whom had suffered MI, narrowed a linkage signal to chromosome 15q26 (87). Of 93 genes in the locus, the *MEF2A* gene was subjected to deep resequencing in family members as a plausible candidate gene, given its expression in embryonic coronary vasculature. A 21-bp deletion in *MEF2A*, resulting in the excision of seven amino acids from the protein product, was detected in each of the affected family members for whom DNA

was available and was absent from unaffected family members. In vitro studies documented that the 21-bp deletion impairs the nuclear localization of the *MEF2A* protein product in cells, suggesting that a functional defect in the protein might be responsible for the prevalent coronary disease in family members with the mutation.

However, a large follow-up study attempting to identify additional deleterious mutations in *MEF2A* in sporadic cases of premature MI did not find any definitive mutations but did succeed in finding the 21-bp mutation in three individuals who had not suffered MI (88). By genotyping family members of these individuals, researchers confirmed that the mutation exists at very low frequency in the general population (rather than being exclusive to one family); it did not segregate with MI or coronary disease outside of the family in which the mutation had originally been described. This study suggests that *MEF2A* does not cause autosomal dominant MI and that the original family's causal mutation may have resided in one of the other 92 genes in the mapped locus on chromosome 15q26. Moreover, in a separate study, investigators were unable to replicate the original observation that the 21-bp deletion led to defective nuclear localization of *MEF2A* (27). This study illustrates the importance of replication studies in human genetics.

A subsequent linkage study of a four-generation family with 28 members with early-onset coronary disease and osteoporosis, transmitted in an autosomal dominant pattern, narrowed a linkage signal to 12p13 (54). The investigators performed detailed mapping of the locus to narrow the signal to a 750-kb interval harboring just six genes, one of which (*LDL receptor-related protein 6*, *or LRP6*) was found to have a missense variant segregating with disease. The variant was also associated with LDL-C and other features of the metabolic syndrome, known risk factors for coronary disease, thereby establishing a plausible causal pathway. The variant was not identified in the screening of 2,000 controls. Furthermore, a knockout mouse for a closely related gene

(*LRP5*) was observed to have altered plasma lipid levels (23), and the knockout mouse for *LRP6* was reported to have developed bone defects akin to osteoporosis (45). For these reasons, *LRP6* has emerged as a more credible autosomal dominant MI gene than *MEF2A*, although additional unrelated families of similar phenotype with mutations in *LRP6* will be required to verify the association. A linkage analysis of more than 400 families with members with premature coronary disease identified an associated common missense variant in the *LRP8* gene (78), although this mutation was not associated with elevated LDL-C levels.

A genome-wide linkage study performed in almost 300 Icelandic families with more than 700 individuals with a history of MI identified a set of linked variants (a haplotype) in the *ALOX5AP* gene associated with a twofold increase in risk for MI as well as stroke (30). The encoded protein product, 5-lipoxygenase activating protein (FLAP), is involved in the production of leukotriene. A subsequent candidate gene study of another actor in the same molecular pathway, *LTA4H* (encoding leukotriene A4 hydrolase), identified haplotypes in the gene associated with increased MI risk in Icelandic and other populations of European descent (29). These observations suggest that leukotriene activity may contribute to the pathobiology of MI. However, the importance of these findings remains unclear as neither locus has been replicated in subsequent GWA studies on MI performed in similar populations (see the section Genome-wide Association Studies: Myocardial Infarction and Coronary Disease, below).

## GENOME-WIDE ASSOCIATION STUDIES

In contrast with linkage studies, which have been successful in identifying rare disease-causing variants exclusive to a few individuals, GWA studies are designed to detect common variants associated with traits or diseases. Over the past few years, the GWA approach has leveraged the ability to efficiently genotype hundreds of thousands of single nucleotide

polymorphism (SNP) variants distributed across the genome in thousands of individuals. GWA is an unbiased technique because the sheer number of SNPs genotyped allows for dense coverage of the genome, and GWA can accommodate large numbers of unrelated individuals with a trait or disease of interest in a given population (with some studies now surpassing 100,000 individuals total).

GWA has now been applied to a wide variety of clinical traits and diseases, including many relevant to the cardiovascular system. Indeed, one of earliest successes of the technique emerged after its application to coronary disease. Considerable progress has also been made with several traditional and emerging risk factors for MI.

## Myocardial Infarction and Coronary Disease

The first large GWA study, published in early 2007, identified novel genetic loci associated with type 2 diabetes (79). Several months later, three independent GWA studies for coronary artery disease were published, one from the Ottawa Heart Study (55), one from deCODE Genetics (32), and one from the Wellcome Trust Case-Control Consortium (75). Each of these studies had a case-control design in which genome-wide genotyping was performed in initial cohorts, followed by genotyping of the best scoring SNPs in additional cohorts; the statistical evidence from the various cohorts was combined to yield the final results.

In spite of their use of different cohorts and different genotyping arrays, all three studies identified a significant association signal in the same locus on chromosome 9p21. No prior genetic studies had implicated this region of the genome. Moreover, the SNPs in the locus that were associated with coronary disease were not associated with any traditional cardiovascular risk factors. Thus, it appears that the genetic mechanism underlying the association signal is operating through a novel pathway. Subsequent studies established an association between the 9p21 locus with MI and other vascular pheno-

types such as abdominal aortic aneurysm, intracranial aneurysm, and peripheral arterial disease, suggesting that the sequence variations may interfere with vascular tissue development (31).

The 9p21 locus illustrates the difficulty of linking some of the genetic associations identified by GWAS with pathological mechanism. No annotated genes are present in the minimal region of association as defined by linkage disequilibrium (∼58 kb in individuals of European descent). The closest genes to the locus, *CDKN2A*, *CDKN2B*, and *ARF*, are more than 100 kb away from the index SNPs (the SNPs with the highest level of association in the GWA studies), making it unclear how the causal DNA variant(s) might influence coronary disease.

Furthermore, GWA studies for type 2 diabetes have identified a strong association signal on chromosome 9p21 that is distinct and independent from the MI locus but in close proximity (separated by a recombination hotspot), with the index diabetes SNPs several kilobases away from the index MI SNPs (17, 77, 93). As with the MI locus, there are no genes within the type 2 diabetes locus. How the causal variants in the respective loci, close together in space and far away from the closest genes, give rise to two rather different diseases remains to be determined.

One possibility is that the loci harbor nongene transcripts that regulate other genes. Within the MI 9p21 locus, there is a predicted noncoding RNA termed *ANRIL* that appears to be expressed in at least a few tissue types (66). The casual variant may affect the RNA sequence in such a way as to affect its function. Another possibility is that the causal variant lies in a regulatory element (e.g., a transcriptional enhancer) that affects the transcription of a gene or genes that are ultimately responsible for the phenotype. *ANRIL*, *CDKN2A*, *CDKN2B*, and *ARF* are logical candidates given their relative proximity to the locus, but a long-range effect on more distant genes cannot be ruled out. Multiple studies have documented that the genotype at an index SNP is correlated to expression levels of nearby genes, although

these results are conflicting and were derived in tissue types that may not be directly relevant to coronary disease (e.g., T lymphocytes) (36, 50). Nonetheless, these studies do document the presence of functional regulatory elements within the MI locus that, in the right tissue type (e.g., vascular endothelium), may modulate molecular pathways that influence MI risk. Besides the 9p21 locus, the study from the Wellcome Trust Case Control Consortium (75) identified SNPs in several additional loci associated with coronary disease. A subsequent set of GWA studies for either coronary disease or MI, each with several thousand disease cases, replicated some of these loci and identified several more loci (19, 26, 56, 83).

At present, thirteen loci now have strong statistical evidence for association with MI or coronary disease (**Table 2**). At the time of this writing, a consortium of investigators focused on coronary disease (the Coronary Artery DIsease Genome-Wide Replication and Meta-analysis or CARDIoGRAM consortium) have assembled more than 20,000 cases of coronary disease and 60,000 control samples, with the expectation of validating or discovering additional associated loci.

For each of these 13 mapped loci, the principal challenge now is to blaze a path from genomic localization to functional insights. For each associated locus, four questions require an answer: (*a*) What is the the causal DNA variant at the locus? (*b*) What is the gene regulated by the locus? (*c*) What is the mechanism by which the DNA variant affects the gene? (*d*) What is the mechanism by which the gene influences phenotype?

At one locus recently identified by GWAS for MI and LDL-C, we have shown through a series of studies in human cohorts and human-derived hepatocytes that a common noncoding polymorphism at the 1p13 locus, rs12740374, creates a C/EBP transcription factor binding site and alters the hepatic expression of the *SORT1* gene (**Figure 1**) (57). With siRNA knockdown and viral overexpression in mouse liver, we demonstrated that *Sort1* alters plasma LDL-C and very-low-density lipopro-

tein (VLDL) particle levels by modulating hepatic VLDL secretion. Thus, we provided functional evidence for a novel regulatory pathway for lipoprotein metabolism and showed that common noncoding DNA variants identified by GWASs can directly contribute to clinical phenotypes.

## Plasma Lipid Concentrations

The first reported high-density GWA study for plasma lipid concentrations was performed in the Diabetes Genetics Initiative, using genotype data from almost 3,000 individuals of European descent. This study identified an index SNP in one locus each for three lipid traits—LDL-C, HDL-C, and triglycerides—with association surpassing a statistical significance threshold that accounts for multiple testing ($P < 5 \times 10^{-8}$) (17). The SNP for LDL-C was near the *APOE* gene, and the index SNP for HDL-C was near the *CETP* gene, both well-established lipoprotein regulators; this finding provided compelling internal validation of the GWA technique. The index SNP for triglycerides was in a locus that harbored no genes previously implicated in lipoprotein metabolism, in an intron of *GCKR* (which encodes glucokinase regulatory protein). Follow-up analyses have suggested that a coding missense variant is responsible for the association with triglyceride levels (6, 64).

A second wave of GWA studies on lipid traits added the Finland–United States Investigation of NIDDM Genetics (FUSION) and SardiNIA studies to the Diabetes Genetics Initiative for a total of almost 9,000 individuals of European descent (42, 89). In order to increase the power to detect statistically significant ($P < 5 \times 10^{-8}$) associations, the most highly associated SNPs in the initial 9,000 individuals were genotyped in >18,000 individuals from additional cohorts. This staged approach resulted in the validation or identification of a total of 19 lipid-associated loci. In addition to the three loci already identified, these GWA studies added loci harboring well-known lipid regulators, including *APOA1*, *APOB*, *LDLR*,

**Table 2    Loci associated with myocardial infarction or coronary artery disease**
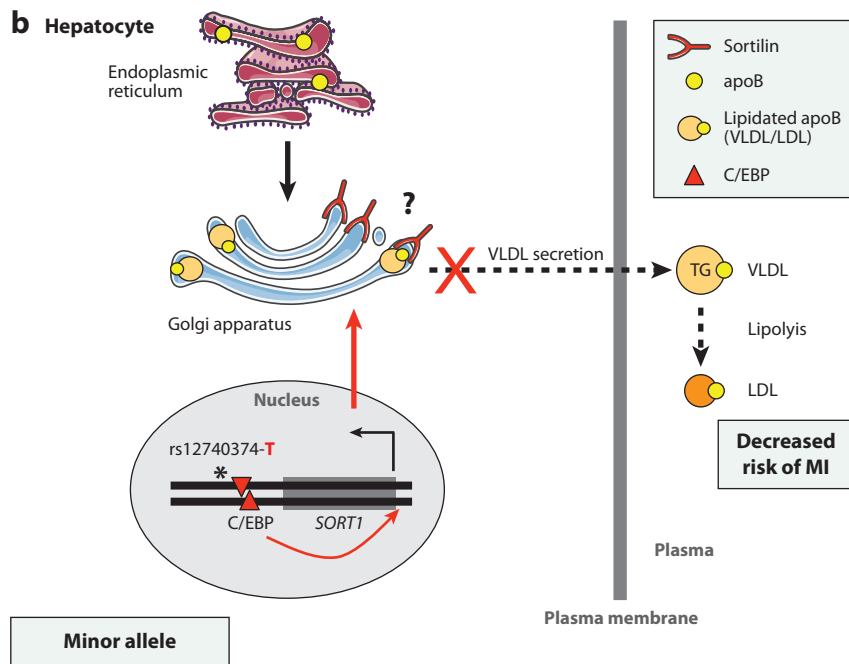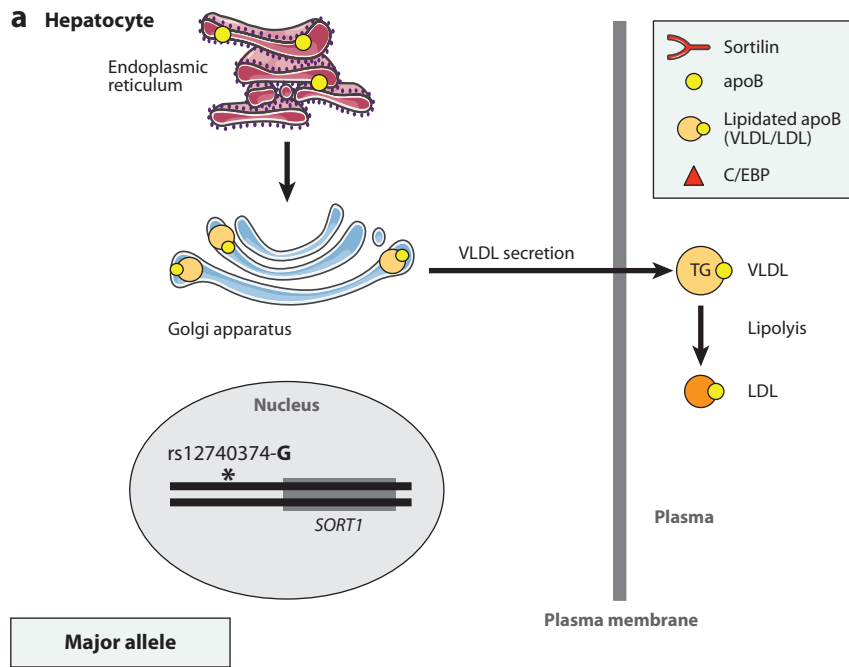
| Unique locus | Chr. | SNP | Risk allele frequency (%) | Odds ratio (95% CI) per risk allele | Gene(s) of interest within or near associated interval | Associated with low-density lipoprotein cholesterol or lipoprotein (a)? | Reference |
|---|---|---|---|---|---|---|---|
| 1 | 9p21 | rs4977574 | 56 | 1.29 (1.25–1.34) | CDKN2A-CDKN2B-ANRIL | — | 55, 32, 75, 56 |
| 2 | 1p13 | rs646776 | 81 | 1.19 (1.13–1.26) | CELSR2-PSRC1-SORT1 | Yes | 75, 56 |
| 3 | 21q22 | rs9982601 | 13 | 1.20 (1.14–1.27) | SLC5A3-MRPS6-KCNE2 | — | 56 |
| 4 | 1q41 | rs17465637 | 72 | 1.14 (1.10–1.19) | MIA3 | — | 75, 56 |
| 5 | 10q11 | rs1746048 | 84 | 1.17 (1.11–1.24) | CXCL12 | — | 75, 56 |
| 6 | 6p24 | rs12526453 | 65 | 1.12 (1.08–1.17) | PHACTR1 | — | 56 |
| 7 | 19p13 | rs1122608 | 75 | 1.15 (1.10–1.20) | LDLR | Yes | 56 |
| 8 | 2q33 | rs6725887 | 14 | 1.17 (1.11–1.23) | WDR12 | — | 56 |
| 9 | 1p32 | rs11206510 | 81 | 1.15 (1.10–1.21) | PCSK9 | Yes | 56 |
| 10 | 12q24 | rs2259816 | 37 | 1.08 (1.05–1.11) | HNF1A | Yes | 19 |
| 11 | 12q24 | rs3184504 | 40 | 1.13 (1.08–1.18) | SH2B3 | — | 26 |
| 12 | 3q22 | rs9818870 | 15 | 1.15 (1.11–1.19) | MRAS | — | 19 |
| 13 | 6q26-6q27 | rs3798220 | 2 | 1.47 (1.35–1.60) | LPA | Yes | 11, 52 |
|  |  | rs10455872 | 7 | 1.68 (1.43–1.98) | LPA | Yes | 11 |

Abbreviations: Chr, chromosome; SNP, single nucleotide polymorphism; CI, confidence interval
Table adapted from References 11, 19, 26, 56.

*LPL*, *PCSK9*, and *HMGCR*. The last is of particular note because it encodes the enzyme (3-hydroxy-3-methylglutaryl-coenzyme A reductase) inhibited by the widely used statin class of LDL-C-lowering drugs. These studies also identified six novel loci, two of which were validated in simultaneously published, independent GWA studies on LDL-C (a locus on chromosome 1p13) and triglycerides (a locus on chromosome 7q11) (46, 76, 86).

A third wave of even larger GWA studies, incorporating up to 40,000 individuals of European descent, have validated or identified more than 30 lipid-associated loci, of which approximately half have established lipid regulators (2, 44, 74). A notable finding of these studies is that genes in 11 of the loci are the same genes that harbor rare mutations that cause Mendelian lipid disorders (**Table 1**). Thus, an important lesson from the GWA studies on lipids is that

**a  Hepatocyte**

Endoplasmic
reticulum

Golgi apparatus

**Nucleus**

rs12740374-**G**

*

*SORT1*

VLDL secretion

TG    VLDL

Lipolyis

LDL

Sortilin

apoB

Lipidated apoB
(VLDL/LDL)

C/EBP

**Plasma**

**Plasma membrane**

**Major allele**

**b  Hepatocyte**

Endoplasmic
reticulum

Golgi apparatus

**?**

**Nucleus**

rs12740374-**T**

*

C/EBP    *SORT1*

VLDL secretion

TG    VLDL

Lipolyis

LDL

**Decreased
risk of MI**

Sortilin

apoB

Lipidated apoB
(VLDL/LDL)

C/EBP

**Plasma**

**Plasma membrane**

**Minor allele**

the same genes that cause Mendelian disorders also have common variants that result in smaller but significant changes in lipid levels.

Criticisms that GWA studies can only discover common variants that have little clinical importance, since the effect sizes are small, ignore the possibility that a GWA gene can prove to be clinically relevant if the gene's activity is modulated by a large degree—whether by a naturally occurring, rare mutation in an individual or in a family, as is the case in a Mendelian disorder, or by deliberate targeting of gene activity by a drug. *HMGCR* is a case in point. If statins had not been discovered prior to the GWA era, the finding that common intronic variants in *HMGCR* cause small changes in LDL-C would have suggested pharmacological inhibition of 3-hydroxy-3-methylglutaryl-coenzyme A reductase as a novel therapeutic strategy. By this reasoning, some of the novel GWA loci discovered to date for plasma lipids may harbor clinically useful drug targets and are now the focus of intense functional investigation.

Furthermore, GWA studies can identify common variants that "tag" unrecognized gene coding variants that have large effects on gene function. These coding variants can be discovered by deep resequencing of the genes in a GWA locus. In one example, a GWA on postprandial triglyceride levels in Old Order Amish individuals identified a strongly associated SNP near a gene cluster with *APOA1*, *APOC3*, *APOA4*, and *APOA5*; subsequent exon sequencing identified a nonsense mutation in *APOC3* that proved to be the source of the association signal (69). Carriers of the nonsense mutation had reduced plasma apoC-III protein levels, reduced triglyceride levels, and increased HDL-C levels. Notably, they also had significantly decreased coronary disease as assessed by coronary artery calcification measurement.

At the time of this writing, the Global Lipids Genetics Consortium has assembled data on plasma lipid concentrations for more than 100,000 individuals of European descent and had identified 95 significantly associated loci ($P < 5 \times 10^{-8}$), with 59 showing genome-wide significant association with lipid traits for the first time (82). The newly reported associations include SNPs near known lipid regulators (e.g., *CYP7A1*, *NPC1L1*, and *SCARB1*) as well as in scores of loci not previously implicated in lipoprotein metabolism. The 95 loci contribute not only to normal variation in lipid traits but also to extreme lipid phenotypes and impact lipid traits in three non-European populations (East Asians, South Asians, and African Americans). Three of the novel genes—*GALNT2*, *PPP1R3B*, and *TTC39B*—were validated with experiments in mouse models. Several novel loci associated with serum lipids were also associated with coronary disease. Taken together, these findings provide the foundation to develop a broader biological understanding of lipoprotein metabolism and to identify new therapeutic opportunities for the prevention of CAD.

←-----------------------------------------------------------------------------------

**Figure 1**

Proposed model for 1p13-sortilin pathway. Hepatic apoB synthesis and lipidation begins in the endoplasmic reticulum. Bulk lipid addition continues in the Golgi apparatus to form triglyceride-rich VLDL. VLDL that is secreted undergoes extrahepatic lipolysis into LDL. (*a*) In hepatocytes with the major allele (G) at rs12740374, this DNA base disrupts a C/EBP consensus site; in the absence of C/EBP binding, there is low transcription of the *SORT1* gene in the nucleus. (*b*) In hepatocytes with the minor allele (T) at rs12740374, this DNA base creates a consensus site on which a C/EBP dimer binds, which in turns activates transcription of the *SORT1* gene in the nucleus. This results in increased production of sortilin protein, which results in decreased VLDL secretion, decreased plasma LDL-C levels, and decreased MI risk. Sortilin traffics through the Golgi apparatus, among other cellular compartments; the exact mechanism(s) by which sortilin inhibits VLDL secretion remain(s) to be determined. Abbreviations: apoB, apolipoprotein B; C/EBP, CCAAT-enhancer-binding protein; LDL, low-density lipoprotein; MI, myocardial infarction; VHDL, very-high-density lipoprotein.

## Other Risk Factors

GWA studies have now been undertaken for a number of cardiovascular risk factors besides plasma lipid concentrations. As with lipids, several successful waves of GWA studies have been performed for type 2 diabetes mellitus, with more than 20 associated loci identified (8, 21, 53, 85, 90, 92). Similarly, GWA studies on blood pressure have identified more than a dozen loci associated with systolic blood pressure, diastolic blood pressure, or hypertension (49, 60). In most cases, the functional connections between loci and phenotypes remain to be determined.

The GWA approach has also been applied to several emerging risk factors for MI. For example, plasma levels of C-reactive protein (CRP) and fibrinogen, two inflammatory biomarkers that predict disease in prospective cohort studies, are each significantly associated with several genetic loci (14, 16, 18, 73). Among the associated SNPs for the respective traits are variants within the *CRP* (encoding CRP) and *FGB* (encoding fibrinogen β chain) genes themselves. Each of the two biomarkers is associated with loci harboring a variety of metabolic, inflammatory, and immunity genes, suggesting that their plasma concentrations reflect inputs from multiple metabolic, inflammatory, and immune pathways.

## MENDELIAN RANDOMIZATION AND IMPLICATIONS OF CAUSALITY

The availability of large cohorts of individuals in whom genome-wide genotyping has been performed makes it feasible to probe the relationships between epidemiological risk factors for a disease and the pathogenesis of the disease. In a technique termed Mendelian randomization, DNA variants are used to address the question of whether an epidemiological association between a risk factor and disease reflects a casual influence of the former on the latter (15). In principle, if a DNA variant is known to directly affect an intermediate phenotype (e.g.,

a variant in the promoter of a gene encoding a biomarker, affecting its expression) and the intermediate phenotype is truly causal for the disease, then the DNA variant should be associated with the disease to the extent predicted by (*a*) the size of the effect of the variant on the phenotype and (*b*) the size of the effect of the phenotype on the disease. If in an adequately powered sample the predicted association between the variant and disease were not observed, it would argue against a purely casual role for the intermediate phenotype in the pathogenesis of the disease.

The study design is akin to a prospective randomized clinical trial in that the randomization for each individual occurs at the moment of conception—genotypes of DNA variants are randomly "assigned" to gametes during meiosis, a process that should be impervious to the typical confounders observed in observational epidemiological studies. For example, a parent's disease status or socioeconomic status should not affect which of the parent's two alleles at a given SNP is passed to a child, with each allele having an equal (50%) chance of being transmitted via the gamete to the zygote. Thus, Mendelian randomization should be unaffected by confounding or reverse causation. Mendelian randomization has potential shortcomings, including (*a*) the technique is only as reliable as the robustness of the estimates of the effect sizes of the variant on the phenotype and of the phenotype on disease, and (*b*) it assumes that the DNA variant does not influence the disease by means other than the intermediate phenotype being studied (pleiotropy), which may not be true. Nevertheless, Mendelian randomization has the potential to be as informative as a traditional randomized clinical trial.

Several Mendelian randomization studies have confirmed a causal relationship between LDL-C and coronary disease. Nonsense variants in the *PCSK9* gene that significantly reduce plasma LDL-C concentrations were observed to be associated with reduced incidence of coronary heart disease in an African American cohort (12, 13). Similarly, in European Americans

a common missense variant in *PCSK9* associated with lower LDL-C levels was also found to be associated with lower risk of MI (13, 43). In a more systematic GWA study, SNPs in 11 loci found to be associated with LDL-C were also reported to be associated with coronary disease (89).

Similar to LDL-C, four recent genetic studies have confirmed prior observations that plasma lipoprotein(a) causally relates to coronary disease (11, 38, 39, 52, 83). Together, these studies have provided the three key pieces of evidence required in Mendelian randomization studies. First, they each showed that in individuals of European ancestry, polymorphisms at the *LPA* locus [the gene encoding for apolipoprotein(a)] are associated with increased plasma lipoprotein(a) levels. Second, higher plasma lipoprotein(a) levels were associated with increased risk for coronary events. Finally, the studies demonstrated that individuals carrying *LPA* variants associated with higher plasma lipoprotein(a) levels were, in turn, at higher risk for coronary events and that this increased risk for coronary disease was abolished after adjustment for plasma lipoprotein(a).

Unlike the results with plasma LDL-C concentrations and plasma lipoprotein(a), three recent, large Mendelian randomization studies of *CRP* variants that affect plasma CRP concentrations, performed in thousands of individuals, did not show an association between these variants and ischemic vascular disease or coronary disease (18, 47, 91). It is unlikely that these studies were confounding or affected by pleiotropy, since the tested variants were within the *CRP* gene itself rather than being in other loci that might secondarily affect CRP levels. Although these studies cannot definitively rule out a causal role of CRP in cardiovascular disease, they strongly suggest that high CRP levels are indirectly rather than directly related to coronary disease. A parallel line of genetic evidence also casts doubt on the notion that inflammatory biomarkers such as CRP are critical mediators of MI and coronary disease. Of the 13 loci most highly associated with MI and coronary disease (**Table 2**), five are related to

plasma LDL-C or lipoprotein(a), arguing for a strong causal relationship between LDL [or a modified LDL particle such as lipoprotein(a)] and disease.

None of the other 8 loci are clearly related to inflammation, suggesting that inflammatory molecules are of less pathobiological importance to MI than either plasma lipid concentrations or the as yet uncharacterized risk mechanisms represented by the 8 non-lipid-related loci. Of critical importance, this pattern cannot be attributed to a negative bias of GWA studies toward inflammatory gene SNPs, since GWA studies on diseases such as rheumatoid arthritis and Crohn's disease have identified numerous SNPs associated with inflammation and immune system functioning (3, 72).

Genetic and therapeutic data supporting causal relevance are also much less extensive for plasma HDL-C and triglycerides. For example, patients with Tangier disease have extremely low HDL-C due to mutations in *ABCA1* (encoding adenosine triphosphate–binding cassette transporter A1), but it has been controversial whether *ABCA1* mutations that lower HDL-C increase risk for coronary disease (22). In addition, treatment with torcetrapib, a drug that raised plasma HDL-C level to a very substantial degree, failed to achieve the predicted benefit on risk of MI (5). Several Mendelian randomization studies have been conducted for HDL-C or triglycerides. However, these studies have not been definitive, because of limited sample size, use of a small number of SNPs at a few genes, or use of SNPs that affected multiple lipid fractions (22, 37, 89, 84).

As more robust statistical frameworks underlying genetic techniques such as Mendelian randomization are developed in the next few years, we can expect fresh new insights to emerge regarding the pathogenetic contributions of risk factors to coronary disease. For example, Mendelian randomization studies focused on variants in inflammatory genes or variants that specifically affect HDL-C may ultimately clarify the extent of the contribution of inflammation and HDL-C to coronary disease.

## MULTIMARKER GENOTYPE SCORES FOR CORONARY RISK PREDICTION

Commonly used cardiovascular risk algorithms such as the Framingham Risk Score, which includes several traditional risk factors and is limited to 10-year predictions, do not accurately predict many coronary events. Much effort is directed towards identification of novel risk factors that, when combined with conventional risk algorithms, will better predict who will develop disease. Accordingly, there is considerable interest in determining whether the use of genetics will improve risk prediction.

A genetic risk score combining nine SNPs associated with either LDL-C or HDL-C (with the score ranging from 0 to 18, for the number of unfavorable alleles) was associated with incident cardiovascular disease in a prospective cohort study, with each unfavorable allele conferring a 15% increase in risk after adjustment for traditional risk factors, including plasma lipid concentrations (41). Upon stratification into groups with high risk scores or low risk scores, individuals with high risk scores had a 63% increase in risk. However, addition of the genotype score to traditional risk factors did not significantly improve risk discrimination, with no change in a commonly used metric (C statistic). Notably, all of the SNPs used in this genetic risk score predated the GWA era; thus, the genetic risk score did not include most of the recently identified lipid-associated SNPs, which upon inclusion might significantly improve the predictive value of the risk score.

A comprehensive genetic risk score would include not just lipid-associated SNPs but also SNPs not associated with traditional risk factors, such as index SNPs in the chromosome 9p21 locus identified in GWA studies. The 9p21 genotype by itself confers up to a 60% increase in risk (in individuals with two unfavorable alleles) (55, 56, 81). A risk score comprising nine GWA SNPs associated with early-onset MI, including a 9p21 SNP and three SNPs associated with LDL-C, is even more highly associated with disease, with a twofold difference in risk for MI between extreme quintiles of risk score (56).

Despite this promising data, attempts to incorporate 9p21 SNPs into risk prediction models have been disappointing to date. As with the lipids-only genetic risk score, the addition of the 9p21 genotype to traditional risk factors in prospective cohort studies with men (81) and women (67) resulted in no improvement in risk discrimination (as judged by C statistic).

As of the time of this writing, a comprehensive genotype score incorporating all of the SNPs discovered to be strongly associated with cardiovascular disease had yet to be evaluated. Finally, as described at the beginning of this review, having a family history of early-onset MI in at least one parent more than doubles the personal risk of having a cardiovascular event (51). As more disease-associated genetic variants are discovered and added to a comprehensive genetic risk score, it will be critical to assess whether the score adds any predictive value above and beyond simply asking a patient about his/her family history; it may emerge that the score has little value for most individuals. For this reason, a genetic risk score may ultimately prove most useful in children or young individuals to gauge lifetime cardiovascular risk and guide early interventions.

## FUTURE DIRECTIONS

In the past few years, the GWA methodology has yielded more candidate regulators of cardiovascular traits and diseases than all of the genetics studies of the preceding era, although the relative importance of these new loci to disease pathogenesis awaits future studies. We now have in hand a long list of genetic loci—together, harboring hundreds of genes—that are associated with phenotypes related to coronary disease. The principal challenge for the next few years will be to define which of the candidate genes are truly causal and to delineate the molecular mechanisms by which they influence atherosclerosis.

One strategy to identify causal genes is to perform deep resequencing of positional

candidate genes in the hopes of uncovering "smoking-gun" mutations (nonsense mutations that yield truncated protein products or missense mutations that alter amino acids critical to protein function) that are clearly linked to phenotype. In one early success for this strategy, rare mutations in *IFIH1* were identified via the sequencing of 10 positional candidate genes for type 1 diabetes (nominated by GWA studies) in 480 diabetic patients and 480 control individuals (59). The variants were genotyped in a total of 30,000 individuals, which confirmed that each of the variants were protective against disease. These findings established *IFIH1* as a causal gene for type 1 diabetes, though the mechanism remains unclear.

Maximizing the chances of success for this strategy will entail careful selection of individuals for gene sequencing. For quantitative traits such as plasma lipid concentrations, choosing individuals with extreme values may increase the odds of finding smoking-gun mutations. For example, focusing exclusively on the lowest stratum (∼3%) of LDL-C concentrations in African Americans in the Dallas Heart Study yielded two nonsense variants in the *PCSK9* that proved to be protective against coronary heart disease in the Atherosclerosis Risk in Communities study (13). Future efforts along these lines might involve choosing the top and bottom strata of a prospective cohort study, or recruiting individuals who present to clinics with extremely high or low lipid levels, and sequencing all of the novel GWA-nominated positional candidate genes in these individuals, followed by genotype replication in a full prospective cohort study to confirm association with lipid traits. For qualitative traits such as MI, focusing on extreme presentations of disease or health—e.g., MI in young individuals or absence of coronary disease in very elderly individuals—may prove to be a fruitful approach.

One important shortcoming of the resequencing strategy is that the failure to find smoking-gun mutations in a gene does not rule out its being a causal gene but may simply reflect that there are no naturally occurring mutations in the gene to be found in the study population.

This could be because the gene is so important to normal development and function that a rare variant greatly perturbing the gene's function would not be tolerated in a viable organism. Another possibility is that variants do exist but in populations different from the study population (e.g., a different ethnic group).

Furthermore, causal variants likely exist in genes not identified by GWA studies. We expect that this last point will soon be addressed by next-generation sequencing technology that will allow for whole-exome sequencing (i.e., all exons of all genes in the genome) and, ultimately, whole-genome sequencing. In proof-of-concept experiments, whole exome sequencing of a handful of individuals affected by Mendelian disorders has identified candidate causal genes (10, 61, 62). Application of this approach to complex traits such as MI and coronary disease is likely to expand the identification of genes and variants.

Undertaking functional validation in appropriate model systems is a parallel strategy to identify causal genes. For loci where there are obvious gene candidates, investigators can use mice as a model in which to study gene function and determine whether the genes influence atherosclerotic plaque formation or MI risk factors (assuming that the genes have mouse orthologs). For example, because lipoprotein metabolism is largely centered on the liver, methods to reliably overexpress and knock down candidate genes in mouse liver (viral vectors and antisense/double-stranded RNAs, respectively), followed by measurement of plasma lipid changes within a few days, offer the opportunity to screen GWA-nominated lipid genes in a high-throughput fashion. This somatic approach is preferable to the prohibitively time-intensive strategy of generating transgenic or knockout mice for each candidate gene. Furthermore, the ability to modulate plasma lipids in directions predicted to be favorable for MI risk (e.g., decreasing LDL-C) by a somatic approach in mice forecasts a parallel strategy in humans that could be of therapeutic value.

Loci in which there are no clear gene candidates—e.g., the 9p21 locus for MI, for

which there is no annotated coding gene within the ~58-kb span—may not be suitable for investigation in mice or other animal models. These loci most likely harbor nongenic regulatory elements such as transcriptional enhancers, repressors, microRNAs, and the like, that have long-range effects on distant genes; these elements may function quite differently in mice. It may ultimately prove necessary to study these loci in a human model system, such as human embryonic stem cells, to determine how they influence atherosclerotic disease.

## CONCLUSIONS

The last few years have witnessed a dramatic advance in mapping genetic loci for coronary disease and MI. Whereas our prior knowledge was largely limited to rare variants and low-frequency variants in lipid genes supplemented by a handful of findings from linkage and candidate genes studies, the GWA approach has identified numerous loci with common variants associated with MI and/or its risk factors. As the exact nature of many of these associations remains unclear, a key challenge now is to move from mapping loci to pinpointing causal genes and variants and then to the development of a molecular understanding of how these genes lead to coronary disease. We anticipate that such study will ultimately yield important new genetic, epidemiological, and functional insights into the development of coronary disease.

## LITERATURE CITED

1. Abifadel M, Varret M, Rabes JP, Allard D, Ouguerram K, et al. 2003. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat. Genet.* 34:154–56

2. Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, et al. 2009. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat. Genet.* 41:47–55

3. Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, et al. 2008. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat. Genet.* 40:955–62

4. Barroso I. 2005. Genetics of type 2 diabetes. *Diabet. Med.* 22:517–35

5. Barter PJ, Caulfield M, Eriksson M, Grundy SM, Kastelein JJ, et al. 2007. Effects of torcetrapib in patients at high risk for coronary events. *N. Engl. J. Med.* 357:2109–22

6. Beer NL, Tribble ND, McCulloch LJ, Roos C, Johnson PR, et al. 2009. The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Hum. Mol. Genet.* 18:4081–88

7. Berge KE, Tian H, Graf GA, Yu L, Grishin NV, et al. 2000. Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. *Science* 290:1771–75

8. Bouatia-Naji N, Bonnefond A, Cavalcanti-Proenca C, Sparso T, Holmkvist J, et al. 2009. A variant near *MTNR1B* is associated with increased fasting plasma glucose levels and type 2 diabetes risk. *Nat. Genet.* 41:89–94

9. Broeckel U, Hengstenberg C, Mayer B, Holmer S, Martin LJ, et al. 2002. A comprehensive linkage analysis for myocardial infarction and its related risk factors. *Nat. Genet.* 30:210–14

10. Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR, et al. 2009. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc. Natl. Acad. Sci. USA* 106:19096–101

11. Clarke R, Peden JF, Hopewell JC, Kyriakou T, Goel A, et al. 2009. Genetic variants associated with Lp(a) lipoprotein level and coronary disease. *N. Engl. J. Med.* 361:2518–28

12. Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH. 2005. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in *PCSK9*. *Nat. Genet.* 37:161–65

13. Cohen JC, Boerwinkle E, Mosley TH, Jr, Hobbs HH. 2006. Sequence variations in *PCSK9*, low LDL, and protection against coronary heart disease. *N. Engl. J. Med.* 354:1264–72

14. Danik JS, Pare G, Chasman DI, Zee RYL, Kwiatkowski DJ, et al. 2009. Novel loci, including those related to Crohn disease, psoriasis, and inflammation, identified in a genome-wide association study of fibrinogen in 17,686 women: the Women's Genome Health Study. *Circ. Cardiovasc. Genet.* 2:134–41

15. Davey Smith G, Ebrahim S. 2003. 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* 32:1–22

16. Dehghan A, Yang Q, Peters A, Basu S, Bis JC, et al. 2009. Association of novel genetic loci with circulating fibrinogen levels: a genome-wide association study in 6 population-based cohorts. *Circ. Cardiovasc. Genet.* 2:125–33

17. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research, Saxena R, Voight BF, Lyssenko V, Burtt NP, et al. 2007. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 316:1331–36

18. Elliott P, Chambers JC, Zhang W, Clarke R, Hopewell JC, et al. 2009. Genetic loci associated with C-reactive protein levels and risk of coronary heart disease. *JAMA* 302:37–48

19. Erdmann J, Grosshennig A, Braund PS, Konig IR, Hengstenberg C, et al. 2009. New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat. Genet.* 41:280–82

20. Francke S, Manraj M, Lacquemant C, Lecoeur C, Lepretre F, et al. 2001. A genome-wide scan for coronary heart disease suggests in Indo-Mauritians a susceptibility locus on chromosome 16p13 and replicates linkage with the metabolic syndrome on 3q27. *Hum. Mol. Genet.* 10:2751–65

21. Frayling TM. 2007. Genome-wide association studies provide new insights into type 2 diabetes etiology. *Nat. Rev. Genet.* 8:657–62

22. Frikke-Schmidt R, Nordestgaard BG, Stene MC, Sethi AA, Remaley AT, et al. 2008. Association of loss-of-function mutations in the *ABCA1* gene with high-density lipoprotein cholesterol and risk of ischemic heart disease. *JAMA* 299:2524–32

23. Fujino T, Asaba H, Kang MJ, Ikeda Y, Sone H, et al. 2003. Low-density lipoprotein receptor-related protein 5 (LRP5) is essential for normal cholesterol metabolism and glucose-induced insulin secretion. *Proc. Natl. Acad. Sci. USA* 100:229–34

24. Garcia CK, Wilund K, Arca M, Zuliani G, Fellin R, et al. 2001. Autosomal recessive hypercholesterolemia caused by mutations in a putative LDL receptor adaptor protein. *Science* 292:1394–98

25. Goldstein JL, Brown MS. 2009. The LDL receptor. *Arterioscler. Thromb. Vasc. Biol.* 29:431–8

26. Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadottir A, Sulem P, et al. 2009. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat. Genet.* 41:342–47

27. Guella I, Rimoldi V, Asselta R, Ardissino D, Francolini M, et al. 2009. Association and functional analyses of MEF2A as a susceptibility gene for premature myocardial infarction and coronary artery disease. *Circ. Cardiovasc. Genet.* 2:165–72

28. Havlik RJ, Garrison RJ, Feinleib M, Kannel WB, Castelli WP, McNamara PM. 1979. Blood pressure aggregation in families. *Am. J. Epidemiol.* 110:304–12

29. Helgadottir A, Manolescu A, Helgason A, Thorleifsson G, Thorsteinsdottir U, et al. 2006. A variant of the gene encoding leukotriene A4 hydrolase confers ethnicity-specific risk of myocardial infarction. *Nat. Genet.* 38:68–74

30. Helgadottir A, Manolescu A, Thorleifsson G, Gretarsdottir S, Jonsdottir H, et al. 2004. The gene encoding 5-lipoxygenase activating protein confers risk of myocardial infarction and stroke. *Nat. Genet.* 36:233–39

31. Helgadottir A, Thorleifsson G, Magnusson KP, Gretarsdottir S, Steinthorsdottir V, et al. 2008. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat. Genet.* 40:217–24

32. Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Blondal T, et al. 2007. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* 316:1491–93

33. Heller DA, de Faire U, Pedersen NL, Dahlen G, McClearn GE. 1993. Genetic and environmental influences on serum lipid levels in twins. *N. Engl. J. Med.* 328:1150–56

34. International HapMap Consortium. 2005. A haplotype map of the human genome. *Nature* 437:1299–320

35. International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, et al. 2007. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449:851–61

36. Jarinova O, Stewart AF, Roberts R, Wells G, Lau P, et al. 2009. Functional analysis of the chromosome 9p21.3 coronary artery disease risk locus. *Arterioscler. Thromb. Vasc. Biol.* 29:1671–77

37. Johannsen TH, Kamstrup PR, Andersen RV, Jensen GB, Sillesen H, et al. 2009. Hepatic lipase, genetically elevated high-density lipoprotein, and risk of ischemic cardiovascular disease. *J. Clin. Endocrinol. Metab.* 94:1264–73

38. Kamstrup PR, Tybjaerg-Hansen A, Steffensen R, Nordestgaard BG. 2009. Genetically elevated lipoprotein(a) and increased risk of myocardial infarction. *JAMA* 301:2331–39

39. Kathiresan S. 2009. Lp(a) lipoprotein redux—from curious molecule to causal risk factor. *N. Engl. J. Med.* 361:2573–74

40. Kathiresan S, Manning AK, Demissie S, D'Agostino RB, Surti A, et al. 2007. A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. *BMC Med. Genet.* 8(Suppl 1):S17

41. Kathiresan S, Melander O, Anevski D, Guiducci C, Burtt NP, et al. 2008. Polymorphisms associated with cholesterol and risk of cardiovascular events. *N. Engl. J. Med.* 358:1240–49

42. Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, et al. 2008. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat. Genet.* 40:189–97

43. Kathiresan S, Myocardial Infarction Genetics Consortium. 2008. A *PCSK9* missense variant associated with a reduced risk of early-onset myocardial infarction. *N. Engl. J. Med.* 358:2299–300

44. Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, et al. 2009. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat. Genet.* 41:56–65

45. Kokubu C, Heinzmann U, Kokubu T, Sakai N, Kubota T, et al. 2004. Skeletal defects in ringelschwanz mutant mice reveal that Lrp6 is required for proper somitogenesis and osteogenesis. *Development* 131:5469–80

46. Kooner JS, Chambers JC, Aguilar-Salinas CA, Hinds DA, Hyde CL, et al. 2008. Genome-wide scan identifies variation in *MLXIPL* associated with plasma triglycerides. *Nat. Genet.* 40:149–51

47. Lawlor DA, Harbord RM, Timpson NJ, Lowe GD, Rumley A, et al. 2008. The association of C-reactive protein and CRP genotype with coronary heart disease: findings from five studies with 4,610 cases among 18,637 participants. *PLoS One* 3:e3011

48. Levy D, DeStefano AL, Larson MG, O'Donnell CJ, Lifton RP, et al. 2000. Evidence for a gene influencing blood pressure on chromosome 17. Genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the Framingham Heart Study. *Hypertension* 36:477–83

49. Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, et al. 2009. Genome-wide association study of blood pressure and hypertension. *Nat. Genet.* 41:677–87

50. Liu Y, Sanoff HK, Cho H, Burd CE, Torrice C, et al. 2009. INK4/ARF transcript expression is associated with chromosome 9p21 variants linked to atherosclerosis. *PLoS ONE* 4:e5027

51. Lloyd-Jones DM, Nam BH, D'Agostino RB S, Levy D, Murabito JM, et al. 2004. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. *JAMA* 291:2204–11

52. Luke MM, Kane JP, Liu DM, Rowland CM, Shiffman D, et al. 2007. A polymorphism in the protease-like domain of apolipoprotein(a) is associated with severe coronary artery disease. *Arterioscler. Thromb. Vasc. Biol.* 27:2030–36

53. Lyssenko V, Nagorny CL, Erdos MR, Wierup N, Jonsson A, et al. 2009. Common variant in *MTNR1B* associated with increased risk of type 2 diabetes and impaired early insulin secretion. *Nat. Genet.* 41:82–88

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

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

56. Myocardial Infarction Genetics Consortium, Kathiresan S, Voight BF, Purcell S, Musunuru K, et al. 2009. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat. Genet.* 41:334–41

57. Musunuru K, Strong A, Frank-Kamenetsky M, Lee NE, Ahfeldt T, et al. From noncoding variant to phenotype via *SORT1* at the 1p13 cholesterol locus. *Nature.* In press

58. Namboodiri KK, Kaplan EB, Heuch I, Elston RC, Green PP, et al. 1985. The Collaborative Lipid Research Clinics Family Study: biological and cultural determinants of familial resemblance for plasma lipids and lipoproteins. *Genet. Epidemiol.* 2:227–54

59. Nejentsev S, Walker N, Riches D, Egholm M, Todd JA. 2009. Rare variants of *IFIH1*, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science* 324:387–89

60. Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, et al. 2009. Genome-wide association study identifies eight loci associated with blood pressure. *Nat. Genet.* 41:666–76

61. Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, et al. 2010. Exome sequencing identifies the cause of a mendelian disorder. *Nat. Genet.* 42:30–35

62. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, et al. 2009. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* 461:272–76

63. Nora JJ, Lortscher RH, Spangler RD, Nora AH, Kimberling WJ. 1980. Genetic-epidemiologic study of early-onset ischemic heart disease. *Circulation* 61:503–8

64. Orho-Melander M, Melander O, Guiducci C, Perez-Martinez P, Corella D, et al. 2008. Common missense variant in the glucokinase regulatory protein gene is associated with increased plasma triglyceride and C-reactive protein but lower fasting glucose concentrations. *Diabetes* 57:3112–21

65. Pajukanta P, Cargill M, Viitanen L, Nuotio I, Kareinen A, et al. 2000. Two loci on chromosomes 2 and X for premature coronary heart disease identified in early- and late-settlement populations of Finland. *Am. J. Hum. Genet.* 67:1481–93

66. Pasmant E, Laurendeau I, Heron D, Vidaud M, Vidaud D, Bieche I. 2007. Characterization of a germ-line deletion, including the entire *INK4/ARF* locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res.* 67:3963–69

67. Paynter NP, Chasman DI, Buring JE, Shiffman D, Cook NR, Ridker PM. 2009. Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3. *Ann. Intern. Med.* 150:65–72

68. Pilia G, Chen WM, Scuteri A, Orru M, Albai G, et al. 2006. Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet.* 2:e132

69. Pollin TI, Damcott CM, Shen H, Ott SH, Shelton J, et al. 2008. A null mutation in human *APOC3* confers a favorable plasma lipid profile and apparent cardioprotection. *Science* 322:1702–5

70. Rader DJ, Cohen J, Hobbs HH. 2003. Monogenic hypercholesterolemia: new insights in pathogenesis and treatment. *J. Clin. Invest.* 111:1795–803

71. Rao DC, Laskarzewski PM, Morrison JA, Khoury P, Kelly K, et al. 1982. The Cincinnati Lipid Research Clinic Family Study: cultural and biological determinants of lipids and lipoprotein concentrations. *Am. J. Hum. Genet.* 34:888–903

72. Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Guiducci C, et al. 2008. Common variants at *CD40* and other loci confer risk of rheumatoid arthritis. *Nat. Genet.* 40:1216–23

73. Ridker PM, Pare G, Parker A, Zee RY, Danik JS, et al. 2008. Loci related to metabolic-syndrome pathways including *LEPR, HNF1A, IL6R,* and *GCKR* associate with plasma C-reactive protein: the Women's Genome Health Study. *Am. J. Hum. Genet.* 82:1185–92

74. Sabatti C, Service SK, Hartikainen AL, Pouta A, Ripatti S, et al. 2009. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat. Genet.* 41:35–46

75. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, et al. 2007. Genomewide association analysis of coronary artery disease. *N. Engl. J. Med.* 357:443–53

76. Sandhu MS, Waterworth DM, Debenham SL, Wheeler E, Papadakis K, et al. 2008. LDL-cholesterol concentrations: a genome-wide association study. *Lancet* 371:483–91

77. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, et al. 2007. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 316:1341–45

78. Shen GQ, Li L, Girelli D, Seidelmann SB, Rao S, et al. 2007. An LRP8 variant is associated with familial and premature coronary artery disease and myocardial infarction. *Am. J. Hum. Genet.* 81:780–91

79. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, et al. 2007. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 445:881–85

80. Soria LF, Ludwig EH, Clarke HR, Vega GL, Grundy SM, McCarthy BJ. 1989. Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. *Proc. Natl. Acad. Sci. USA* 86:587–91

81. Talmud PJ, Cooper JA, Palmen J, Lovering R, Drenos F, et al. 2008. Chromosome 9p21.3 coronary heart disease locus genotype and prospective risk of CHD in healthy middle-aged men. *Clin. Chem.* 54:467–74

82. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, et al. Biological, clinical, and population relevance of 95 loci for blood lipids. *Nature*. In press

83. Tregouet DA, Konig IR, Erdmann J, Munteanu A, Braund PS, et al. 2009. Genome-wide haplotype association study identifies the *SLC22A3-LPAL2-LPA* gene cluster as a risk locus for coronary artery disease. *Nat. Genet.* 41:283–85

84. Triglyceride Coronary Disease Genetics Consortium and Emerging Risk Factors Collaboration. 2010. Triglyceride-mediated pathways and coronary disease: a collaborative analysis of 101 studies. *Lancet* 375:1634–39

85. Unoki H, Takahashi A, Kawaguchi T, Hara K, Horikoshi M, et al. 2008. SNPs in *KCNQ1* are associated with susceptibility to type 2 diabetes in East Asian and European populations. *Nat. Genet.* 40:1098–102

86. Wallace C, Newhouse SJ, Braund P, Zhang F, Tobin M, et al. 2008. Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. *Am. J. Hum. Genet.* 82:139–49

87. Wang L, Fan C, Topol SE, Topol EJ, Wang Q. 2003. Mutation of *MEF2A* in an inherited disorder with features of coronary artery disease. *Science* 302:1578–81

88. Weng L, Kavaslar N, Ustaszewska A, Doelle H, Schackwitz W, et al. 2005. Lack of *MEF2A* mutations in coronary artery disease. *J. Clin. Invest.* 115:1016–20

89. Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, et al. 2008. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat. Genet.* 40:161–69

90. Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, et al. 2008. Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. *Nat. Genet.* 40:1092–97

91. Zacho J, Tybjaerg-Hansen A, Jensen JS, Grande P, Sillesen H, Nordestgaard BG. 2008. Genetically elevated C-reactive protein and ischemic vascular disease. *N. Engl. J. Med.* 359:1897–908

92. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, et al. 2008. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.* 40:638–45

93. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, et al. 2007. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316:1336–41

Changes may still occur before final publication online and in print.