REVIEW ARTICLE

GENOMIC MEDICINE

W. Gregory Feero, M.D., Ph.D., and Alan E. Guttmacher, M.D., *Editors*

Genomics of Cardiovascular Disease

Christopher J. O'Donnell, M.D., and Elizabeth G. Nabel, M.D.

From the National Heart, Lung, and Blood Institute and the Framingham Heart Study, National Institutes of Health, Bethesda, MD (C.J.O.); and Massachusetts General Hospital (C.J.O.), Brigham and Women's Hospital (E.G.N.), and Harvard Medical School (C.J.O., E.G.N.) — all in Boston. Address reprint requests to Dr. Nabel at Brigham and Women's Hospital, 75 Francis St., Boston, MA 02115, or at enabel@partners.org.

N Engl J Med 2011;365:2098-109.

An interactive timeline is available at NEJM.org

Cardiovascular disease is the leading cause of death in the United States. Considerable progress has been made in the past 50 years to define, identify, and modify risk factors for cardiovascular disease (e.g., hypertension, dyslipidemia, obesity, type 2 diabetes, cigarette smoking, and physical inactivity) and to develop treatments, such as coronary care units, percutaneous coronary interventions, and beta-blockers. These efforts have resulted in an age-adjusted decline in cardiovascular mortality.[1] In addition, it is now possible to detect subclinical disease by means of blood biomarker testing[2] or imaging measurements[3] years before the onset of symptoms or other clinical manifestations.[4] Despite this progress, mechanisms that underlie individual differences in the presentation and pathophysiological features of cardiovascular disease are poorly understood. In this article, we review genetic and genomic studies in cardiovascular medicine that have helped to elucidate some of these mechanisms during the past decade (Fig. 1, and interactive timeline, available with the full text of this article at NEJM.org).

Mendelian and Candidate-Gene Studies

Ten years ago, the draft sequence of the human genome, which was produced by scientists working on the Human Genome Project[5] and others,[6] was first described, leading to an expansion in the understanding of genetic contributions to cardiovascular disease. Before the Human Genome Project, many genes associated with mendelian cardiovascular disease had been identified.[7] These forms of cardiovascular disease are rare and constitute a minority of clinical cardiovascular diseases. Genetically, they are simple in that a mutation in a single gene is sufficient to cause disease, so mendelian disease is said to be monogenic. Examples include forms of premature myocardial infarction, dilated and hypertrophic cardiomyopathy, heart failure, arrhythmogenic right ventricular dysplasia, the long-QT syndrome, and aortic aneurysms.[7] Recessive mutations underlie familial forms of cardiovascular risk factors, such as hypertension, hypercholesterolemia, and type 2 diabetes. Knowledge obtained through the identification of genes associated with mendelian disease has led to breakthrough discoveries in mechanisms of cardiovascular disease and its treatment. A compelling illustration is the Nobel-Prize–winning discovery that mutations affecting the low-density lipoprotein (LDL) receptor cause hypercholesterolemia and early-onset myocardial infarction, which led to LDL cholesterol–lowering therapies that reduce the risk of cardiovascular events.[8]

Genomics of Cardiovascular Risk and Disease

Role of Genomewide Association Studies

The large majority of cardiovascular diseases, however, are polygenic, with both heritable and environmental contributions.[9] Moreover, there are heritable, polygenic com-

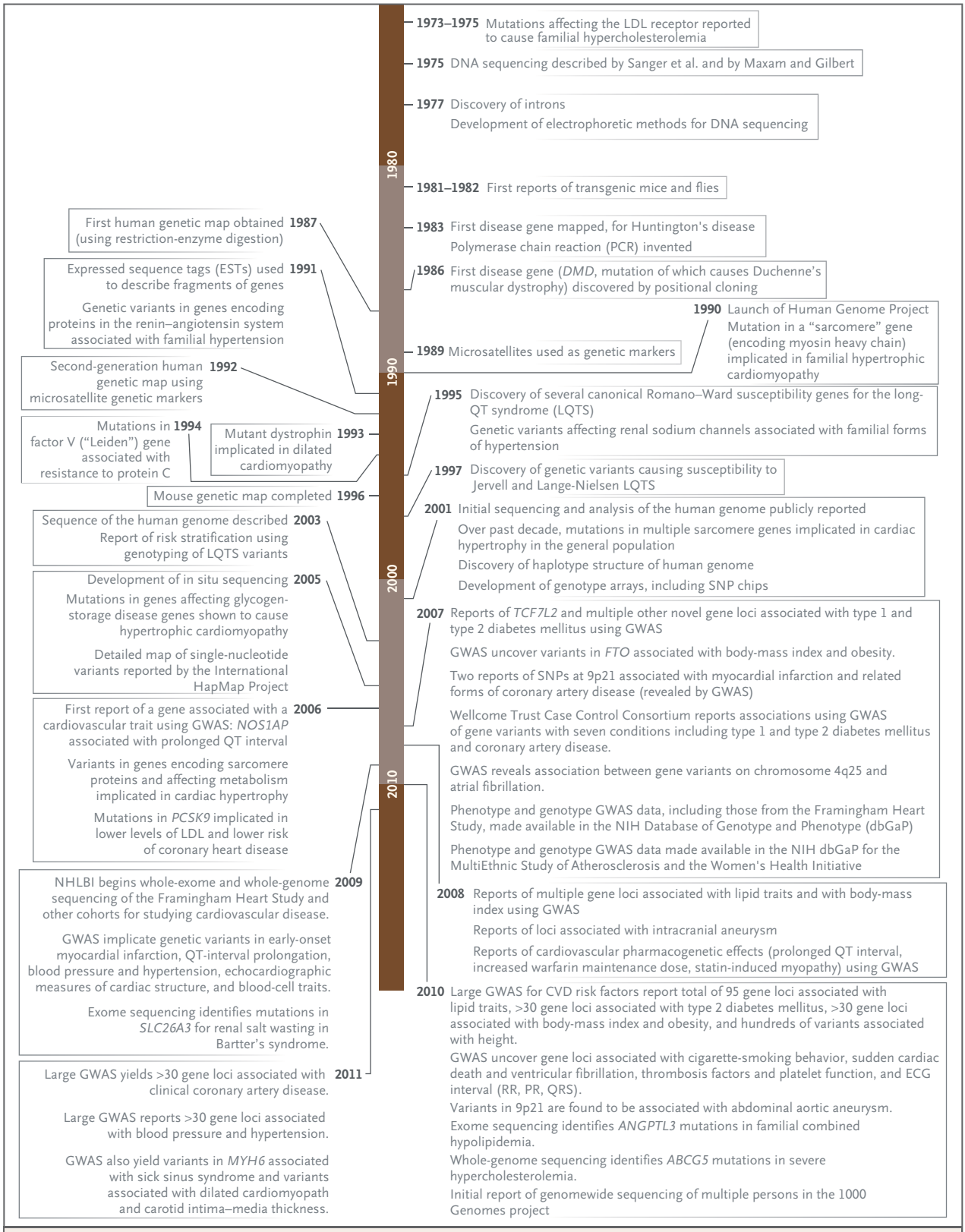

Figure 1. Timeline of Genetic and Genomic Research in Cardiovascular Medicine.

ponents of cardiovascular risk factors and subclinical disease, including coronary artery disease.[10] Approaches to identifying the genetic causes of polygenic cardiovascular diseases (and other polygenic diseases) before completion of the draft sequence of the human genome were largely unsuccessful. A decade later, hundreds of loci associated with many cardiovascular diseases and traits have been identified. This genetic bounty is the yield of genomewide association studies, which involve testing of a large set of genetic variants in case and control subjects from a population to determine which variants are associated with the disease in question (see the Glossary).[11]

The common disease–common variant hypothesis proposes that common variants, which are defined as variants with a prevalence of at least 5% in the population, have a role in the cause and

Glossary

Allele: One of two or more copies of a genetic sequence at a chromosomal location. Alleles can be considered according to their frequencies in the human population, ranging from common variants (minor allele frequency, >5%) to low-frequency variants (minor allele frequency, 0.5 to 5%) to rare or private alleles in one or a limited number of families (minor allele frequency, <0.5%). A null allele is not functional.

Candidate-gene study: An approach used in genetics research that focuses on suspected, or candidate, genes that have been selected because of a perceived match between their known or presumed function and biologic function in the disease under investigation.

Exome: All known protein-coding sequences, or exons, in the human genome, constituting approximately 1 to 2% of the 3.2 billion nucleotide base pairs in the human genome.

Gene enhancer: A short region of DNA that can be bound with proteins, such as trans-acting factors (i.e., factors acting from a different molecule), to enhance transcription levels of genes in a gene cluster. Although enhancers are usually cis-acting (i.e., acting from the same molecule), an enhancer does not need to be close to the genes it acts on and may not be located on the same chromosome.

Genomewide association study: An approach used in genetics research to look for associations between many (typically hundreds of thousands) specific genetic variations (most commonly single-nucleotide polymorphisms) and particular diseases.

Genotyping array: A technique used to study many genes at once. Thousands of gene sequences are placed in known locations on a glass slide. A sample containing DNA or RNA is deposited on the slide, or gene chip. The binding of complementary base pairs from the sample and the gene sequences on the chip can be measured with the use of fluorescence to detect the presence and determine the amount of specific sequences in the sample.

Induced pluripotent stem cell: A type of pluripotent stem cell derived from a nonpluripotent cell, typically an adult somatic cell such as a fibroblast, by transfection of a stem-cell–associated gene.

Haplotype maps: Maps of a set of DNA variations, or polymorphisms, that tend to be inherited together. Haplotype maps are used to aid gene-discovery efforts.

Haplotype structure: Patterns of haplotypes that define the structure of regions of the genome. Haplotype structure is used to compare genome regions across populations or species.

Locus: The specific chromosomal location of a gene or other DNA sequence of interest.

Mendelian inheritance: Patterns of inheritance characteristic of organisms that reproduce sexually, as described by Austrian monk Gregor Mendel in the mid-19th century.

Mendelian randomization: A method that uses measured variation in genes of known function to examine the causal effect of a modifiable exposure on a disease. It provides unbiased estimates of the effects of a putative casual variable without the need for a traditional randomized trial.

Next-generation sequencing: DNA sequencing that harnesses advances in miniaturization technology to simultaneously sequence multiple areas of the genome rapidly and at low cost.

Noncoding DNA sequence: A DNA sequence that does not encode proteins. Noncoding DNA sequences, once referred to as "junk DNA," account for the majority of genome sequences and are now known to harbor regions that regulate gene expression.

Polygenic: Produced by two or more genes.

Single-nucleotide polymorphism: A single-nucleotide variation in a genetic sequence; a common form of variation in the human genome.

Splice variants: Abnormal variations in RNA splicing that are implicated in a disease. Many genetic disorders result from splicing variants.

Whole-exome sequencing: Sequencing of the coding regions, or exons, of an entire genome.

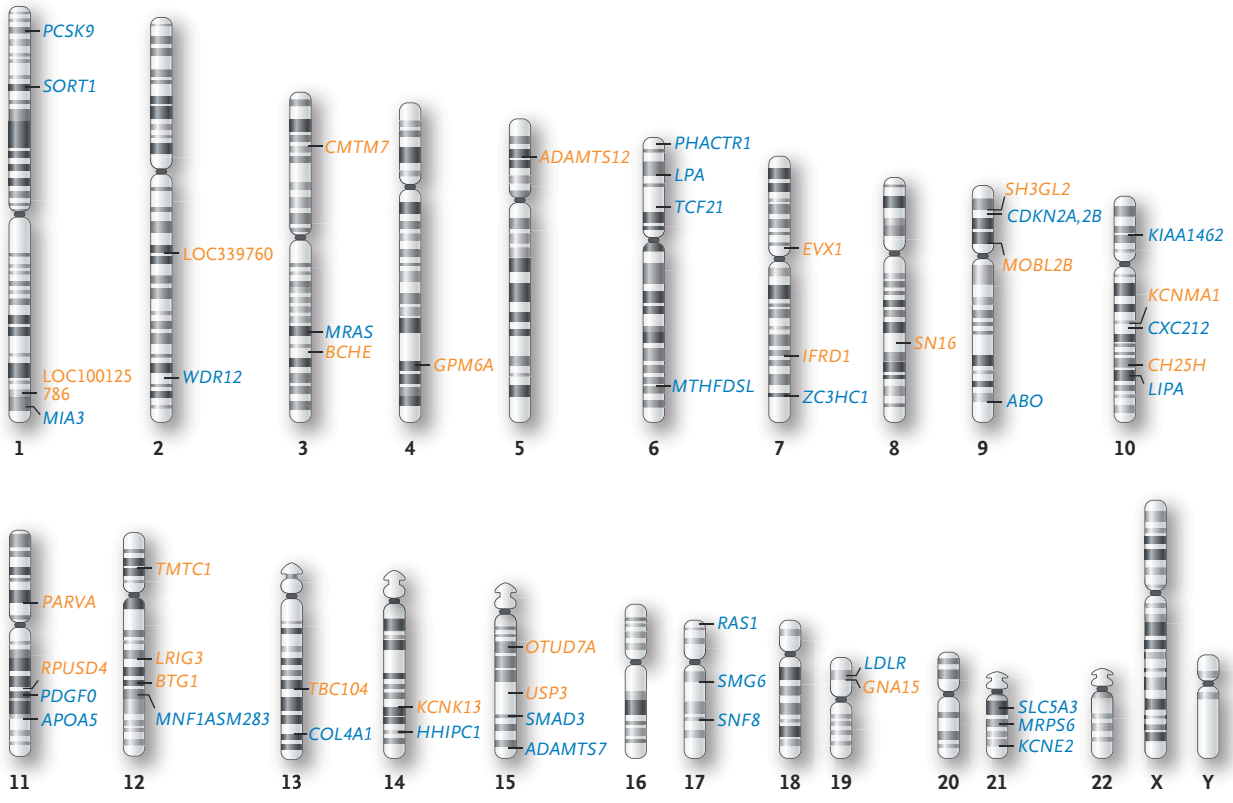

Figure 2. Genomic Locations of Genetic Variants Associated with the Risk of Myocardial Infarction and Heart Failure.
Genetic variants that are associated with myocardial infarction (blue) and heart failure (orange) are shown according to their chromosomal location. LOC denotes locus.

pathophysiology of common diseases. It is on this premise that the genomewide association study is based, since it comprises tests of association between disease and common variants spread throughout the genome. On the basis of this approach, researchers have assembled catalogues of cardiovascular variants, using genotyping arrays, haplotype maps, and statistical methods. Changes in data-sharing policies have led to the creation of publicly available genome databases. International collaborations combining study cohorts, often including tens of thousands of research participants, have been formed.[12]

The astonishingly large number of new loci associated with cardiovascular risk factors, subclinical indexes, and disease end points have provided insights into the biologic pathways that underlie disease (Fig. 2). The application of such findings to the prediction of risk and to the prevention and treatment of disease is premature and awaits considerable research.

CORONARY ARTERY DISEASE AND MYOCARDIAL INFARCTION

Genomewide association studies have identified about 30 loci associated with myocardial infarction and coronary artery disease (Table 1).[13-35] A meta-analysis of 14 such studies that involved 22,233 case subjects with coronary artery disease and 64,762 control subjects of European descent and that were followed by replication studies involving 56,682 case and control subjects identified 13 new loci associated with coronary artery disease, in addition to confirming 10 of 12 previously reported loci.[14] *ABO* and *ADAMTS7* were found to be associated with angiographically confirmed coronary atherosclerosis, *CNNM2* with high blood pressure, and the *APOA5* gene cluster with elevated levels of triglycerides and cholesterol subfractions. The majority of loci associated with myocardial infarction reside in genomic regions that have not previously been implicated in coronary artery disease; only a minority of loci mediate an

Table 1. Representative Genomewide Association Studies of Common Cardiovascular Diseases.

Study	Clinical Outcome	Sample Size		Major Ethnic Group	Selected Major Findings
		Genomewide Association	Replication		
		number of subjects			
Kathiresan et al., 2009[13]	Myocardial infarction	6,042	19,492	European	9 Loci, including 9p21 and *SORT1*
Schunkert et al., 2011[14]	Coronary artery disease	86,995	56,682	European	>20 Loci, including *ABO* and *ADAMTS7*
C4D Consortium, 2011[15]	Coronary artery disease	30,482	40,593	European, East Asian	Novel loci, including *LIPA*
Smith et al., 2010[16]	Heart failure incidence	23,821	NA	European, African	Suggestive loci, including *LRIG3*
Morrison et al., 2010[17]	Heart failure mortality	2,992	NA	European, African	Suggestive loci, including *ADAMTS12*
Villard et al., 2011[18]	Dilated cardiomyopathy	2,287	2,467	European	Top loci: *HSPB7* and *BAG3*
Ikram et al., 2009[19]	Ischemic stroke	19,602	7,269	European, African	Top locus: *NINJ2*
Yamada et al., 2009[20]	Ischemic stroke	267	5,981	Japanese	Top locus: *CELSR1*
Akiyama et al., 2010[21]	Intracranial aneurysm	482	1,398	Japanese	Suggestive loci, including *ARHGEF11* and *TMEM195*
Bilguvar et al., 2008[22]	Intracranial aneurysm	7,856	1,171	European, Japanese	Top loci: 9p21 and *SOX17*
Yasuno et al., 2010[23]	Intracranial aneurysm	15,295	4,777	European, Japanese	Associations with 9p21 and *CNNM2*
Koriyama et al., 2010[24]	Peripheral arterial disease	1,553	2,239	Japanese	Top loci: *OSBPL10* and 10p12.31
Thorgeirsson et al., 2008[25]	Peripheral arterial disease	10,995	4,848	European	Top locus: *CHRNA3*
Elmore et al., 2009[26]	Abdominal aortic aneurysm	235	1,421		Suggestive locus: 3p12.3
Gretarsdottir et al., 2010[27]	Abdominal aortic aneurysm	31,795	10,718	European	Top loci: 9p21 and *DAB2IP*
Benjamin et al., 2009[28]	Lone atrial fibrillation	40,518	6,218	European	Top loci: *PITX2* and *ZFHX3*
Gudbjartsson et al., 2007[29]	Lone atrial fibrillation	36,137	5,806	European	Top loci: 4q25 and *ZFHX3*
Ellinor et al., 2010[30]	Lone atrial fibrillation	14,179	4,771	European	Top loci: *KCNN3* and 20q13.13
Holm et al., 2011[31]	Sick sinus syndrome	38,384	1,654	European	Top locus: *MYH6*
Bezzina et al., 2010[32]	Ventricular fibrillation	972	537	European	Top locus: *CXADR*
Arking et al., 2010[33]	Sudden cardiac arrest	650	19,611	European	Top locus: *GPC5*
Arking et al., 2011[34]	Sudden cardiac death	22,055	14,265	European	Top locus: *BAZ2B*
Trégouët et al., 2009[35]	Venous thromboembolism	1,647	3,237	European	Top locus: *AB0*

effect through known risk factors. A second meta-analysis of genomewide association studies involving more than 30,000 case and control subjects showed an additional four new loci associated with coronary artery disease in multiple ethnic groups.[15]

Studies of early-onset myocardial infarction have identified more than 10 risk loci.[13] The most strongly associated locus is 9p21.[36,37] This region harbors genes (*CDKN2A* and *CDKN2B*) that are implicated in cell cycling and cancer, although the single-nucleotide polymorphisms (SNPs) associated with myocardial infarction are located not near these or other protein-coding genes but rather within a noncoding RNA molecule called ANRIL (for antisense noncoding RNA in the *INK4* locus). The expression of ANRIL splice variants, but not *CDKN2B* or other nearby genes, is associated with atherosclerosis.[38] SNPs at the 9p21 locus are also associated with other cardiovascular diseases, including stroke[39] and aortic aneurysm.[40]

Loci associated with coronary artery disease harbor genes known to be important in lipid variation, including *SORT1*, *PCSK9*, *HNF1A*, *MRAS*, and *LPA*. The position of other SNPs implicates inflam-

matory processes in conferring a risk of coronary atherosclerosis.

HEART FAILURE

Genomewide association studies have identified many possible loci associated with heart failure and death from heart failure, although few of such studies have been replicated.[16,17] Genomewide association studies for heart failure have been limited by modest numbers of cases of heart failure (relative to the number of such studies for coronary artery disease) and the heterogeneous nature (and thus heterogeneic sets of cases) of heart failure. A recent genomewide association study for idiopathic dilated cardiomyopathy identified variants in *HSPB7*, encoding a heat-shock protein previously implicated in heart failure, and *BAG3*; a marked myopathy develops in mice deficient in *Bag3*.[18]

ARRHYTHMIAS

Genomewide association studies have uncovered genetic variants for arrhythmias, including atrial fibrillation,[28-30] ventricular fibrillation,[32] sudden cardiac death,[33,34] and the sick sinus syndrome[31] (Table 1). *MYH6*, a previously unidentified gene associated with susceptibility to the sick sinus syndrome, encodes the alpha heavy-chain subunit of cardiac myosin,[31] suggesting that myosin proteins may regulate cardiac conduction in addition to myocyte function.

PERIPHERAL AND CEREBRAL VASCULAR DISEASES

Genomewide association studies have yielded evidence of varying strength for genetic loci associated with ischemic stroke,[19,20] intracranial aneurysm,[21-23] peripheral arterial disease,[24,25] aortic aneurysm,[25-27] venous thromboembolism,[35] and erythrocyte phenotypes[41] (Table 1). In some cases, these loci are common to coronary artery disease and myocardial infarction, suggesting a common genetic contribution to multiple vascular beds.

MODIFIABLE RISK FACTORS AND SUBCLINICAL DISEASE

Genome consortia with sample sizes exceeding 10,000 cases have investigated major modifiable risk factors (e.g., hypertension, dyslipidemia, type 2 diabetes, and cigarette smoking) and obesity, using quantitative measures and clinically relevant extreme end points (Table 2). These studies point to loci associated with hypertension (as defined by the use of antihypertensive therapy or elevations in blood pressure),[42] LDL and high-density lipoprotein (HDL) cholesterol and triglycerides,[43] the number of cigarettes smoked,[44] type 2 diabetes (with the use of quantitative traits of fasting glucose and glycated hemoglobin levels),[45,46] and obesity (with the use of body-mass index and adiposity-related traits).[47,48] Studies have also identified multiple loci associated with other presumed quantitative biomarkers of cardiovascular risk and presymptomatic coronary artery disease (e.g., fibrinogen,[49] C-reactive protein,[50] intercellular adhesion molecule 1,[51] plasma homocysteine,[52] and carotid-artery intima–media thickness and plaque[53]).

Newer approaches for establishing evidence of causality use human populations as a model system. Mendelian randomization analysis takes advantage of the lifelong association between a risk allele and a quantitative measure of a biomarker for clinical coronary artery disease to estimate whether there is evidence of a causal association between the biomarker and the disease. Mendelian randomization studies have used genetic variants that have previously been implicated by genomewide association studies to provide evidence for a causal association between LDL cholesterol (variants in *LDLR*)[54] or lipoprotein(a) (variants in *LPA*)[55] and coronary artery disease and evidence against a causal association between C-reactive protein (variants in *CRP*) and coronary artery disease.[56]

Current genomewide association studies focus on large populations in order to strengthen the evidence for association and replication. For example, a genomewide association study of more than 100,000 research participants identified 95 loci associated with at least one of three lipids: LDL and HDL cholesterol and triglycerides.[43] Although each individual variant had only a modest effect, the combined effect of the 95 loci explained approximately 25% of the genetic variance in LDL and HDL cholesterol levels. Similarly, large meta-analyses of variants that are associated with type 2 diabetes[45] and myocardial infarction[14] suggest that these variants, in aggregate, account for 25% and 10%, respectively, of the inherited variation in disease outcomes.

We have learned several lessons from genomewide association studies. Although some risk loci that have been discovered through this approach encompass genes that had been previously impli-

Table 2. Representative Large-Scale Genomewide Association Studies of Risk Factors for Cardiovascular Disease.*

Study	Risk Factor or Clinical Trait	Sample Size		Major Ethnic Group	Selected Major Findings
		Genomewide Association	Replication		
		number of subjects			
Ehret et al., 2011[42]	Systolic and diastolic blood pressure, hypertension	69,395	133,661	European	>25 Loci ,including *CACNB2* and *SH2B3*
Teslovich et al., 2010[43]	Total and LDL cholesterol	100,184	39,875	European, South Asian, East Asian, African	>35 Loci, including *SORT1* and *HMGCR*
Teslovich et al., 2010[43]	HDL cholesterol	100,184	39,875	European, South Asian, East Asian, African	>35 Loci, including *SCARB1* and *CETP*
Teslovich et al., 2010[43]	Triglycerides	100,184	39,875	European, South Asian, East Asian, African	>20 Loci, including *ANGPTL3* and *JMJD1C*
Thorgeirsson et al., 2008[25]	Quantity of cigarettes smoked	31,266	54,731	European	Top loci: *CHRNB3* and 15q25
TAGC, 2010[44]	Quantity of cigarettes smoked	74,053	68,988	European	Top loci: *DBH* and *CYP2A6*
Voight et al., 2010[45]	Type 2 diabetes mellitus	47,117	94,337	European	>25 Loci, including *TCF7L* and *IRS1*
Dupuis et al., 2010[46]	Fasting glucose level	46,186	76,558	European	>20 Loci, including *GCKR* and *ADRA2A*
Speliotes et al., 2010[47]	Body-mass index	123,865	125,931	European	>30 Loci, including *FTO* and *TMEM18*
Heard-Costa et al., 2009[48]	Waist circumference	31,373	38,641	European	Top loci: *NRXN2* and *MC4R*

* Genomewide association studies were considered to be large in scale if they include a sample of more than 10,000 subjects. HDL denotes high-density lipoprotein, LDL low-density lipoprotein, and TAGC Tobacco and Genetics Consortium.

cated, most of such loci implicate genes that had not been thought to have a role in conferring disease risk. Many loci show association with cardiovascular disease across groups of differing ancestry, and most cardiovascular traits are influenced by a large number of loci. However, limitations of genomewide association studies have prevented immediate translation of these findings into clinical practice, since each variant has a very small effect and is therefore not useful for prediction. Moreover, the implicated variants are rarely themselves the causal variants; rather, they are linked to the true causal variants, and identification of the latter usually warrants a great deal of additional work.

TARGETED AND GENOMEWIDE DNA SEQUENCING

The next generation of genomic approaches is upon us, thanks to the increasing efficiency and decreasing cost of sequencing technology. Deep DNA sequencing is performed with the use of miniaturized technology that simultaneously sequences multiple areas of the genome. This approach is used to sequence candidate regions and specific components of the human genome, such as exons or noncoding DNA. Sequencing studies in community-based cohorts have shown that at least 1 of every 64 persons carries a functional mutation in one of three genes (*NCCT*, *NKCC2*, or *ROMK*) that is associated with clinically significant alterations in blood pressure[57] and that there is a striking excess of nonsynonymous variants in the gene encoding adipokine (*ANGPTL4*) in persons with low triglyceride levels.[58] Targeted sequencing in a population with extreme values of LDL cholesterol identified variants in *PCSK9* that occur in up to 3% of the population. Such variants are associated with a low LDL cholesterol level and a decreased risk of incident coronary artery disease.[59] A sequencing study of *FAAH* and *MGLL* at extremes of the body-mass index uncovered several rare variants in promoter and enhancer regions of these genes, suggesting that sequence variation may regulate extreme obesity.[60]

Target genes for sequencing also include those identified by genomewide association studies that

are associated with cardiovascular disease and its risk factors. Resequencing of exons in *APOA5*, *GCKR*, *LPL*, and *APOB* — genes that are implicated in hypertriglyceridemia — revealed twice as many rare missense or nonsense variants in persons with high triglyceride levels than in control subjects, corresponding to a carrier frequency of 28.1% in affected persons and 15.3% in control subjects.[61]

Whole-exome sequencing has uncovered genetic variants in persons with rare forms of diseases that affect blood pressure and circulating blood lipids, including a missense mutation in *SLC26A3* (a congenital chloride diarrhea locus) in a patient with a suspected diagnosis of the renal salt-wasting disease Bartter's syndrome,[62] a variant in *ANGPTL3* in a family with familial combined hypolipidemia,[63] and variants in *BAG3* in families with familial dilated cardiomyopathy.[64] Whole-exome sequencing in large populations may help to identify rare genetic variants in relatively common cardiovascular diseases, such as early myocardial infarction, that confer a major risk of complications and death and for which there is good evidence of heritability; indeed, this is a major focus of whole-exome sequencing at present. Whole-genome sequencing has also confirmed single-gene disease variants in families with rare diseases,[65] including mutations in *ABCG* in an infant with severe hypercholesterolemia.[66]

CARDIOVASCULAR DISEASE PATHWAYS

The best yield of genomewide association studies is the provision of insights into the biologic pathways — often previously unsuspected — that underlie causes of disease. Such insights have led to hypothesis-driven investigations of implicated pathways with the use of molecular, genetic, biochemical, and cellular approaches. The genetic control of myocardial infarction and lipids provides a case in point.

The initial findings that connected genetic variation in the 9p21 region with atherosclerotic progression and myocardial infarction caught cardiovascular scientists by surprise, because this region is devoid of genes that had previously been associated with coronary artery disease. Subsequent research has revealed possible mechanisms by which genes in this region may contribute to atherosclerosis. A recent report implicated the 9p21 risk interval in regulating the expression of cardiac *CDKN2A/B* expression genes.[38] Another report implicated inflammatory pathways through 33 gene enhancers located in the 9p21 region.[67] Two SNPs located in one of these enhancers disrupt a binding site for STAT1, a signal transduction protein that regulates inflammation. This enhancer locus physically interacts with the *CDKN2A/B* locus and an interval downstream of *IFNA21*, the gene that encodes interferon-γ in human vascular endothelial cells. The activation of interferon-γ affects transcriptional regulation of the 9p21 locus, including STAT1 binding, suggesting a link between genetic susceptibility to coronary artery disease and response to inflammatory signaling in vascular cells. Yet another driver of disease may be the aforementioned noncoding RNA element ANRIL.

Loci that are discovered by genomic approaches are also relevant to lipid biology and potential therapeutic targets. The gene encoding sortilin (*SORT1*) contains a common variant that creates a binding site for a transcription factor that when bound to this site alters liver expression of LDL cholesterol in humans. Transgenic studies have shown that mice with *Sort1* that contains the transcription-factor binding site have altered plasma levels of LDL cholesterol, suggesting a previously unknown regulatory pathway for LDL cholesterol.[68] A potentially therapeutic target was implicated by the discovery of a pair of common nonsense variants in *PCSK9* in patients of African descent. An estimated 2.6% of persons of African descent carry one of these variants, and each variant results in comparatively low lipid levels and a reduced susceptibility to myocardial infarction.[69] Humans who are homozygous for this null allele are healthy and have a reduced risk of myocardial infarction, suggesting that the absence of PCSK9 may be well tolerated and thus rendering PCSK9 an attractive drug target.

SNPs that more recently have been found to be associated with cardiovascular disease and its risk factors implicate components of pathways previously identified with risk factors for or mechanisms of the disease. For example, the 95 loci associated with levels of LDL and HDL cholesterol or triglycerides in more than 100,000 persons of European ancestry implicate nearly all the 18 genes that have previously been shown to be mutated

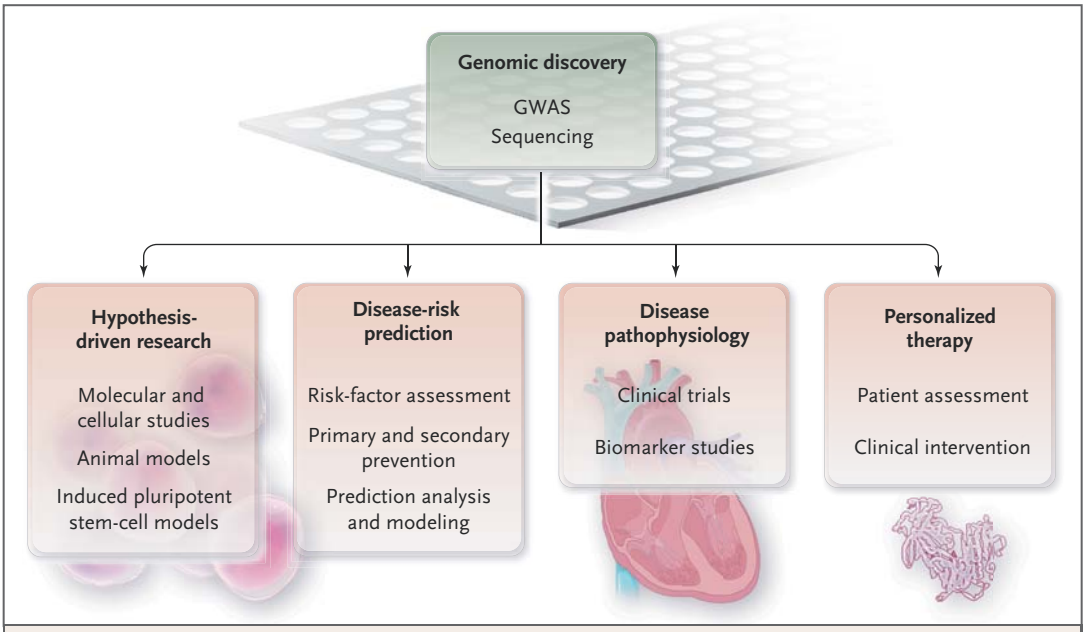

Figure 3. Investigative Pathways Leading from Gene Discovery to Clinical Application.

Genomewide association studies (GWAS) and gene sequencing represent discovery research. Findings from these studies stimulate further investigations in hypothesis-driven research and disease-risk prediction. Applications to clinical cardiovascular care will follow.

in rare mendelian lipid disorders.[43] Most of these loci were also associated with cardiovascular disease in persons of African descent and in persons of Asian descent. Some loci housed both common and rare variants; in such cases, the common variants explained more of the heritability.

An exciting approach is the use of induced pluripotent stem cells to model cardiovascular disease and test hypotheses generated from genomic studies. Two lines of induced pluripotent stem cells, which were generated from dermal fibroblasts in patients who had the long-QT syndrome with specific genetic mutations, have been differentiated into cardiomyocytes and are being used to investigate electrophysiological properties, such as action potentials and ion fluxes.[70,71]

PREDICTION, PREVENTION, AND TREATMENT

The improved understanding of cardiovascular pathophysiology that has been achieved through genetic discovery provides new opportunities for prediction, prevention, and treatment (Fig. 3). Genetic risk prediction is at an early stage, and insufficient evidence exists at present to warrant the use of a genetic risk score on the basis of SNPs identified through genomic approaches.[72,73] Additional research is needed to prospectively assess the utility of genetic risk scores in the prediction of cardiovascular disease, such as myocardial infarction and coronary artery disease, before clinical use. Some observers have suggested that 150 genes with odds ratios of 1.5 or 250 genes with odds ratios of 1.25 will be needed.[74] Future studies of genetic risk scores will probably require hundreds of associated SNPs (identified through genomewide association studies), combined with low-frequency risk alleles discovered through whole-genome sequencing, to provide evidence-based prediction.

Studies are under way to investigate the use of next-generation sequencing for the screening of rare forms of cardiovascular disease[75] and to annotate all mutations within an individual's genome.[76] Although the costs of accurate whole-genome sequencing are dropping dramatically, the costs of data analysis and storage remain high and are barriers to clinical implementation.

Loci that are uncovered by genomewide association studies, despite their modest effects, may

have therapeutic implications. For example, a common variant at the *HMGCR* locus (with a prevalence of 40% among persons of European ancestry) is associated with a very modest elevation in the LDL cholesterol level (2.8 mg per deciliter [0.07 mmol per liter]).[43] There are no known rare mutations at this locus, presumably because they would be lethal. Yet the encoded protein, 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase, is the target of statin drugs, which reduce LDL cholesterol levels and the risk of myocardial infarction.

Pharmacogenomics has potential near-term application. Discoveries that are provided by genomewide association studies have strengthened the evidence for pharmacogenetic interactions in a number of commonly used cardiovascular drugs. Variants in *CYP2C9* and *VKORC1* explain up to 40% of the variation in the adjusted dose of warfarin, and the Food and Drug Administration (FDA) recently revised the label on warfarin to allow for genotype-specific dose ranges.[77] Variations in a cytochrome P-450 enzyme, *CYP2C19*, are associated with decreased antiplatelet efficacy and an increased risk of cardiovascular disease among patients taking the antiplatelet drug clopidogrel, which prompted an FDA warning and recommendations for a dose adjustment or use of alternative drugs in patients with this variant.[77]

Evidence of other pharmacogenetic interactions that may be clinically important is accumulating. Variation in the β_1-adrenergic–receptor gene, *ADRB1*, is associated with altered responsiveness to beta-blockade in heart failure.[78] A variant in *SLC01B1* has been implicated in statin-related myopathy.[79] Deep sequencing of variants and genes related to drug absorption, distribution, metabolism, and excretion may identify specific variants that contribute to the heterogeneity of responsiveness to cardiovascular drugs.

FUTURE DIRECTIONS

The field of cardiovascular genomics has two distinct goals: understanding biologic mechanisms and applying that knowledge to personalized medicine. Knowledge of molecular pathways can lead to improved therapeutics on a broad basis (regardless of the individual genotype) or at an individualized level (targeted specifically to the genotype). During the past 5 years, the discovery of hundreds of cardiovascular loci is a start. In the years to come, we will require studies of tens of thousands of patients with cardiovascular disease that combine tests of genomewide association (involving missense, rare, and common variants) with sequencing.[80] Functional studies are essential to characterize molecular and cellular pathways and to develop appropriately targeted therapies. Genomics is permeating biomedical research, and these genomic advances must proceed through basic discoveries, functional characterization, preclinical proof-of-principle studies, first-in-human studies, clinical trials, and regulatory approval. Cardiovascular science and medicine have made enormous strides over the past century, beginning with the brilliant elucidation of cardiovascular physiology and leading to molecular and cellular studies, with concurrent epidemiologic determinations of risk factors. The cardiovascular field is now primed for genomic medicine to make equal, if not greater, contributions.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

REFERENCES

1. NHLBI fact book, fiscal year 2009. Bethesda, MD: National Heart, Lung, and Blood Institute, February 2010 (http://www.nih.gov/about/FactBook2009_final.pdf).
2. Kucharska-Newton AM, Couper DJ, Pankow JS, et al. Hemostasis, inflammation, and fatal and nonfatal coronary heart disease: long-term follow-up of the Atherosclerosis Risk in Communities (ARIC) cohort. Arterioscler Thromb Vasc Biol 2009; 29:2182-90.
3. Kathiresan S, Larson MG, Keyes MJ, et al. Assessment by cardiovascular magnetic resonance, electron beam computed tomography, and carotid ultrasonography of the distribution of subclinical atherosclerosis across Framingham risk strata. Am J Cardiol 2007;99:310-4.
4. Greenland P, Alpert JS, Beller GA, et al. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. Circulation 2010;122(25):e584-e636.
5. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. Nature 2001;409:860-921. [Errata, Nature 2001;411:720, 412:565.]
6. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. Science 2001;291:1304-51.
7. Nabel EG. Cardiovascular disease. N Engl J Med 2003;349:60-72. [Erratum, N Engl J Med 2003;349:620.]
8. Goldstein JL, Brown MS. The LDL receptor. Arterioscler Thromb Vasc Biol 2009; 29:431-8.
9. Lee DS, Pencina MJ, Benjamin EJ, et al. Association of parental heart failure with risk of heart failure in offspring. N Engl J Med 2006;355:138-47.
10. Parikh NI, Hwang SJ, Larson MG, et al. Parental occurrence of premature cardio-

vascular disease predicts increased coronary artery and abdominal aortic calcification in the Framingham Offspring and Third Generation cohorts. Circulation 2007; 116:1473-81.
11. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet 2003;33:177-82.
12. O'Donnell CJ, Nabel EG. Cardiovascular genomics, personalized medicine, and the National Heart, Lung, and Blood Institute: part I: the beginning of an era. Circ Cardiovasc Genet 2008;1:51-7.
13. Kathiresan S, Voight BF, Purcell S, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet 2009;41:334-41. [Erratum, Nat Genet 2009;41:762.]
14. Schunkert H, Konig IR, Kathiresan S, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet 2011; 43:333-8.
15. Coronary Artery Disease (C4D) Genetics Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. Nat Genet 2011;43:339-44.
16. Smith NL, Felix JF, Morrison AC, et al. Association of genome-wide variation with the risk of incident heart failure in adults of European and African ancestry: a prospective meta-analysis from the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium. Circ Cardiovasc Genet 2010;3: 256-66.
17. Morrison AC, Felix JF, Cupples LA, et al. Genomic variation associated with mortality among adults of European and African ancestry with heart failure: the cohorts for heart and aging research in genomic epidemiology consortium. Circ Cardiovasc Genet 2010;3:248-55.
18. Villard E, Perret C, Gary F, et al. A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. Eur Heart J 2011;32:1065-76.
19. Ikram MA, Seshadri S, Bis JC, et al. Genomewide association studies of stroke. N Engl J Med 2009;360:1718-28.
20. Yamada Y, Fuku N, Tanaka M, et al. Identification of CELSR1 as a susceptibility gene for ischemic stroke in Japanese individuals by a genome-wide association study. Atherosclerosis 2009;207:144-9.
21. Akiyama K, Narita A, Nakaoka H, et al. Genome-wide association study to identify genetic variants present in Japanese patients harboring intracranial aneurysms. J Hum Genet 2010;55:656-61.
22. Bilguvar K, Yasuno K, Niemelä M, et al. Susceptibility loci for intracranial aneurysm in European and Japanese populations. Nat Genet 2008;40:1472-7.
23. Yasuno K, Bilguvar K, Bijlenga P, et al. Genome-wide association study of intracranial aneurysm identifies three new risk loci. Nat Genet 2010;42:420-5.
24. Koriyama H, Nakagami H, Katsuya T, et al. Identification of evidence suggestive of an association with peripheral arterial disease at the OSBPL10 locus by genome-wide investigation in the Japanese population. J Atheroscler Thromb 2010;17: 1054-62.
25. Thorgeirsson TE, Geller F, Sulem P, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature 2008;452:638-42.
26. Elmore JR, Obmann MA, Kuivaniemi H, et al. Identification of a genetic variant associated with abdominal aortic aneurysms on chromosome 3p12.3 by genome wide association. J Vasc Surg 2009;49: 1525-31.
27. Gretarsdottir S, Baas AF, Thorleifsson G, et al. Genome-wide association study identifies a sequence variant within the DAB2IP gene conferring susceptibility to abdominal aortic aneurysm. Nat Genet 2010;42:692-7.
28. Benjamin EJ, Rice KM, Arking DE, et al. Variants in ZFHX3 are associated with atrial fibrillation in individuals of European ancestry. Nat Genet 2009;41:879-81.
29. Gudbjartsson DF, Arnar DO, Helgadottir A, et al. Variants conferring risk of atrial fibrillation on chromosome 4q25. Nature 2007;448:353-7.
30. Ellinor PT, Lunetta KL, Glazer NL, et al. Common variants in KCNN3 are associated with lone atrial fibrillation. Nat Genet 2010;42:240-4.
31. Holm H, Gudbjartsson DF, Sulem P, et al. A rare variant in MYH6 is associated with high risk of sick sinus syndrome. Nat Genet 2011;43:316-20.
32. Bezzina CR, Pazoki R, Bardai A, et al. Genome-wide association study identifies a susceptibility locus at 21q21 for ventricular fibrillation in acute myocardial infarction. Nat Genet 2010;42:688-91.
33. Arking DE, Reinier K, Post W, et al. Genome-wide association study identifies GPC5 as a novel genetic locus protective against sudden cardiac arrest. PLoS ONE 2010;5(3):e9879.
34. Arking DE, Junttila MJ, Goyette P, et al. Identification of a sudden cardiac death susceptibility locus at 2q24.2 through genome-wide association in European ancestry individuals. PLoS Genet 2011;7(6):e1002158.
35. Trégouët DA, Heath S, Saut N, et al. Common susceptibility alleles are unlikely to contribute as strongly as the FV and ABO loci to VTE risk: results from a GWAS approach. Blood 2009;113:5298-303.
36. McPherson R, Pertsemlidis A, Kavaslar N, et al. A common allele on chromosome 9 associated with coronary heart disease. Science 2007;316:1488-91.
37. Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 2007;316:1491-3.
38. Visel A, Zhu Y, May D, et al. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. Nature 2010;464:409-12.
39. Anderson CD, Biffi A, Rost NS, Cortellini L, Furie KL, Rosand J. Chromosome 9p21 in ischemic stroke: population structure and meta-analysis. Stroke 2010;41: 1123-31.
40. Helgadottir A, Thorleifsson G, Magnusson KP, et al. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. Nat Genet 2008;40:217-24.
41. Ganesh SK, Zakai NA, van Rooij FJ, et al. Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. Nat Genet 2009;41:1191-8.
42. International Consortium for Blood Pressure Genome-Wide Association Studies. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature 2011;478:103-9.
43. Teslovich TM, Musunuru K, Smith AV, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010;466:707-13.
44. Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nat Genet 2010;42:441-7.
45. Voight BF, Scott LJ, Steinthorsdottir V, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 2010;42:579-89.
46. Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet 2010;42: 105-16.
47. Speliotes EK, Willer CJ, Berndt SI, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet 2010;42:937-48.
48. Heard-Costa NL, Zillikens MC, Monda KL, et al. NRXN3 is a novel locus for waist circumference: a genome-wide association study from the CHARGE Consortium. PLoS Genet 2009;5(6):e1000539.
49. Dehghan A, Yang Q, Peters A, et al. Association of novel genetic loci with circulating fibrinogen levels: a genome-wide association study in 6 population-based cohorts. Circ Cardiovasc Genet 2009;2: 125-33.
50. Elliott P, Chambers JC, Zhang W, et al. Genetic loci associated with C-reactive protein levels and risk of coronary heart disease. JAMA 2009;302:37-48.
51. Barbalic M, Dupuis J, Dehghan A, et al.

Large-scale genomic studies reveal central role of ABO in sP-selectin and sICAM-1 levels. Hum Mol Genet 2010;19:1863-72.
52. Mälarstig A, Buil A, Souto JC, et al. Identification of ZNF366 and PTPRD as novel determinants of plasma homocysteine in a family-based genome-wide association study. Blood 2009;114:1417-22.
53. Bis JC, Kavousi M, Franceschini N, et al. Meta-analysis of genome-wide association studies from the CHARGE consortium identifies common variants associated with carotid intima media thickness and plaque. Nat Genet 2011;43:940-7.
54. Linsel-Nitschke P, Götz A, Erdmann J, et al. Lifelong reduction of LDL-cholesterol related to a common variant in the LDL-receptor gene decreases the risk of coronary artery disease — a mendelian randomisation study. PLoS ONE 2008; 3(8):e2986.
55. Kamstrup PR, Tybjaerg-Hansen A, Steffensen R, Nordestgaard BG. Genetically elevated lipoprotein(a) and increased risk of myocardial infarction. JAMA 2009;301:2331-9.
56. C Reactive Protein Coronary Heart Disease Genetics Collaboration (CCGC). Association between C reactive protein and coronary heart disease: mendelian randomisation analysis based on individual participant data. BMJ 2011;342:d548.
57. Ji W, Foo JN, O'Roak BJ, et al. Rare independent mutations in renal salt handling genes contribute to blood pressure variation. Nat Genet 2008;40:592-9.
58. Romeo S, Pennacchio LA, Fu Y, et al. Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL. Nat Genet 2007;39:513-6.
59. Cohen JC, Boerwinkle E, Mosley TH Jr, Hobbs HH. Sequence variations in *PCSK9*, low LDL, and protection against coronary heart disease. N Engl J Med 2006; 354:1264-72.
60. Harismendy O, Bansal V, Bhatia G, et al. Population sequencing of two endocannabinoid metabolic genes identifies rare and common regulatory variants associated with extreme obesity and metabolite level. Genome Biol 2010;11:R118.
61. Johansen CT, Wang J, Lanktree MB, et al. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. Nat Genet 2010;42: 684-7.
62. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A 2009;106:19096-101.
63. Musunuru K, Pirruccello JP, Do R, et al. Exome sequencing, ANGPTL3 mutations, and familial combined hypolipidemia. N Engl J Med 2010;363:2220-7.
64. Norton N, Li D, Rieder MJ, et al. Genome-wide studies of copy number variation and exome sequencing identify rare variants in BAG3 as a cause of dilated cardiomyopathy. Am J Hum Genet 2011;88: 273-82.
65. Roach JC, Glusman G, Smit AF, et al. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. Science 2010;328:636-9.
66. Rios J, Stein E, Shendure J, Hobbs HH, Cohen JC. Identification by whole-genome resequencing of gene defect responsible for severe hypercholesterolemia. Hum Mol Genet 2010;19:4313-8.
67. Harismendy O, Notani D, Song X, et al. 9p21 DNA variants associated with coronary artery disease impair interferon-γ signaling response. Nature 2011;470:264-8.
68. Musunuru K, Strong A, Frank-Kamenetsky M, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature 2010;466:714-9.
69. Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. Nat Genet 2005;37: 161-5.
70. Itzhaki I, Maizels L, Huber I, et al. Modelling the long QT syndrome with induced pluripotent stem cells. Nature 2011;471:225-9.
71. Yazawa M, Hsueh B, Jia X, et al. Using induced pluripotent stem cells to investigate cardiac phenotypes in Timothy syndrome. Nature 2011;471:230-4.
72. Palomaki GE, Melillo S, Neveux L, et al. Use of genomic profiling to assess risk for cardiovascular disease and identify individualized prevention strategies — a targeted evidence-based review. Genet Med 2010;12:772-84.
73. Ripatti S, Tikkanen E, Orho-Melander M, et al. A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. Lancet 2010;376:1393-400.
74. Pepe MS, Gu JW, Morris DE. The potential of genes and other markers to inform about risk. Cancer Epidemiol Biomarkers Prev 2010;19:655-65.
75. Bell CJ, Dinwiddie DL, Miller NA, et al. Carrier testing for severe childhood recessive diseases by next-generation sequencing. Sci Transl Med 2011;3:65ra4.
76. Ashley EA, Butte AJ, Wheeler MT, et al. Clinical assessment incorporating a personal genome. Lancet 2010;375:1525-35.
77. Wang L, McLeod HL, Weinshilboum RM. Genomics and drug response. N Engl J Med 2011;364:1144-53.
78. Liggett SB, Mialet-Perez J, Thaneemit-Chen S, et al. A polymorphism within a conserved beta(1)-adrenergic receptor motif alters cardiac function and beta-blocker response in human heart failure. Proc Natl Acad Sci U S A 2006;103:11288-93.
79. The SEARCH Collaborative Group. *SLCO1B1* variants and statin-induced myopathy — a genomewide study. N Engl J Med 2008;359:789-99.
80. International HapMap 3 Consortium. Integrating common and rare genetic variation in diverse human populations. Nature 2010;467:52-8.