Cardiovascular disease (CVD) is the most prevalent of all diseases in both industrialized and developing nations. Despite a reported 60% decrease in the age-adjusted CVD mortality rate over the past 30 years (1), disease prevalence remains largely unchanged. According to the American Heart Association, an estimated 80.7 million American adults (1 in 3) have 1 or more types of CVD (2). Mortality data show that CVD was the underlying cause of 36.3% of all deaths in 2004, or 1 of every 2.8 deaths in the U.S. (2). Globally, CVD is the number 1 cause of death; an estimated 17.5 million people died from CVD in 2005, representing 30% of all deaths worldwide (3). These grave statistics are irrespective of recent advances in understanding the pathophysiology of CVD, strong educational efforts promoting healthy life-style changes, and the development of highly effective therapies for treating the underlying risk factors and sequelae of acute ischemic events.

The process of drug development has become increasingly difficult, despite various methodologic improvements (4), and since the introduction of statins in 1987, few novel CVD therapies have emerged. Of the 16 major cardiovascular (CV) drug classes, only 4 (angiotensin receptor blockers, brain natriuretic peptide mimetics, glycoprotein IIb/IIIa inhibitors, and direct rennin inhibitors) have become available in the post-statin era (Table 1) (5,6). The development of a novel CV therapy takes ~15 years and, including the large number of failures, costs >$800 million (7). There is a high attrition rate in pre-clinical and early clinical drug development, reflecting a need for risk-based decision making. During the development of a successful therapy, many interventional targets and chemical compounds are gradually reduced to fewer candidates that are ultimately tested in clinical studies. As the number of targets and compounds decreases geometrically throughout the stages of drug development, the cost increases exponentially; hence, decisions made further along the development continuum have incrementally higher stakes.

Challenges to CV drug development include a scarcity of accepted surrogates for CV outcomes and a lack of useful platforms to assess direct vascular effects of promising new drug candidates (8). The most prominent gaps are the lack of 1) valid drug targets and pre-clinical models that predict efficacy and safety in humans; 2) biomarkers that predict CV clinical outcomes such as myocardial infarction (MI) and unstable angina; 3) responder patient segments that can increase the efficiency and probability of success of outcome trials; and 4) drug/diagnostic combinations. Genomic approaches are currently being applied in industry, biotechnology, and academic settings to address many of these gaps.
target identification and validation are considered. When pharmacologic target validation is still a useful precursor in drug development, new tools have emerged to identify, validate, and prioritize greater numbers of plausible drug targets in basic discovery. The greatest impact of genomics on drug development has been in the area of target identification and validation. Here, the past and present contributions of mouse and human genomics and the promise of systems biology to target identification and validation are considered. When interpreting genomic datasets, one must consider the concept of causation versus association. Figure 1 provides a simplified overview of where and when a genomic marker is causal or reactive. The ability to assign causality is critical to target identification.

Genomic applications to drug target identification and validation. MOUSE GENETICS. The mouse has emerged as the preferred species for the application of genomics to target identification and validation (20, 21). Various genomic methods have been applied to the mouse for the identification and validation of atherosclerosis drug targets, ranging from candidate gene approaches, such as transgenic and gene knockout, to open system approaches, such as gene expression profiling, proteomics, and genetics. One major drawback to using rodent models to study atherosclerosis is that rats and mice have a different lipoprotein profile than humans and carry most of their plasma cholesterol in the high-density lipoprotein (HDL) fraction. Rodents that have not had genetic manipulation of key lipoprotein-determining genes, such as apolipoprotein (apo) E or the low-density lipoprotein (LDL) receptor, do not develop atherosclerosis.

Genetic manipulation has been used to generate mouse models of atherosclerosis. Mice deficient in apoE or the LDL receptor develop a human-like atherosclerosis that has been extensively characterized (22–24). Lesions in apoE-deficient mice initiate as fatty streaks and progress to advanced fibroproliferative plaques, containing inflammatory and HDL-C–raising therapies. Although studies have demonstrated that these lesions can rupture and lead to atherothrombosis (27–29), the mouse is not prone to plaque vulnerability. This lack of vulnerability poses a massive problem for the field. While mouse plaque rupture models have been described, the reproducibility and validity of these models remain in question. The lack of plaque rupture models has severely limited the ability to use the mouse as a model for atherosclerosis. Nevertheless, these mouse models have been instrumental in providing a high throughput system for assessing the effects of genetic manipulation and therapeutic interventions on atherosclerosis plaque burden. Given the compositional similarities of plaque in atherosclerosis-prone mice as compared to humans, there is still value in the model. Many studies have been published demonstrating the potential atherogenicity of candidate genes and the atheroprotective potential of therapeutics (30). The translation of these effects to humans has not been realized, given the long time delay in establishing human outcomes.

Of particular interest is the ability of the mouse model to elucidate the proatherogenic role of inflammation and the antiatherogenic role of high-density lipoprotein cholesterol (HDL-C). The strong causal relationships established in the mouse model underscore the therapeutic potential of anti-inflammatory and HDL-C–raising therapies. Although progress in the development of anti-inflammatory agents has been limited by the lack of necessary clinical tools for establishing proof of concept, advancements in HDL-C–based therapies have been made, and several outcomes trials are currently ongoing or planned (31, 32).

Studies in apoE-deficient mice have provided compelling evidence to suggest a causal association between inflammation, characterized by macrophage infiltration and activity in
plaque, and atherosclerosis (21). Mice with naturally-occurring mutations in granulocyte macrophage–colony stimulating factor that fail to develop differentiated tissue macrophages have markedly reduced atherosclerosis (33). These data were substantiated in a recently constructed mouse atherosclerosis model, the apoE3-Leiden mouse (34). Elevated plasma lipids and atherosclerosis developed only when this transgenic animal was fed a high-fat/high-cholesterol diet. When switched to a low-fat chow diet, the lipid profile normalized, and atherosclerosis development ceased and eventually regressed. Before the regression of atherosclerosis, macrophages completely disappeared from the plaque, suggesting a causal relationship between macrophage presence and plaque development. The role of inflammation in atherosclerosis has been further substantiated by other studies evaluating perturbations of macrophage differentiation and function through genetic and therapeutic intervention. One noteworthy example is the disruption of the chemokine monocyte chemoattractant protein-1 or its receptor CCR2 and the use of anti-CCR2 therapeutics, which significantly diminish atherosclerosis (14,15).

The mouse model has also been instrumental in understanding the relationship between atherosclerosis and HDL biology. Human epidemiologic studies have implicated HDL as atheroprotective (35); however, the atheroprotective benefits of high HDL are often confounded by concurrent low triglyceride levels (36). In addition, genetic deficiencies in HDL that are not associated with changes in triglycerides do not always associate with atherosclerosis (37), and genetic causes of HDL elevation do not always correlate with atheroprotection (38). Moreover, several drug intervention studies in humans (39–41) have failed to demonstrate that pharmacologic elevation of HDL results in atheroprotection. The HDL-raising drug niacin has been associated with CV benefit, but it also lowers triglycerides, LDL-C, and lipoprotein(a) (42), leaving uncertainty as to the exact mechanism by which niacin confers atheroprotection. The mouse models, however, provide definitive evidence that the primary HDL apolipoprotein, apoA1, is atheroprotective. Transgenic overexpression, adenoviral-mediated delivery, and intravenous infusion of apoA1 in an atherosclerosis mouse model produced dramatic and consistent reductions in atherosclerosis (30). These findings provide support for the continued pursuit of therapies that raise HDL through increasing apoA1 production.

Despite the major advances that the apoE-deficient and other mouse models have made to the study of atherosclerosis, the ultimate value of the mouse to the development of novel atherosclerosis therapies has not been defined. Unbiased classical forward mouse genetics may offer additional insights. Atherosclerosis has been studied in strains of mice
by means of dietary manipulation, establishment of recombina
tion inbred strains, or construction of F2 intercrosses 
between strains of mice with varying susceptibility. For 
instance, the C3H mouse strain is highly resistant, whereas 
the C57BL/6 is partially sensitive to diet-induced athero-
sclerosis; these 2 strains have been used as parental strains in 
F2 intercrosses to study the genetic determinants of athero-
sclerosis. Through quantitative trait locus mapping of var-
dious F2 mouse crosses, several atherosclerosis suscepti-
ble loci have been identified (43–46). These early genetic 
crosses highlight the potential to use unbiased whole ge-
ome approaches to identify novel atherosclerosis suscepti-
bility loci. However, identifying the true causal gene from 
the susceptibility loci has been challenging, because loga-

darithm of the odds (LOD) peaks defined by these inter-
crosses are typically broad, with multiple genes lying be-
neath the peak. Thus, the use of unidimensional genetic 
approaches to extrapolate from quantitative trait locus to 
quantitative trait gene has proved time consuming and often 
futile.

HUMAN GENETICS. Human genetics offers tremendous op-
portunities for target identification and validation. The 
ability to identify causal genes for diseases with Mendelian 
genetic inheritance patterns has led to significant gains in 
drug development. More than 30 years ago, the association 
between familial hypercholesterolemia (FH) and the LDL
receptor was identified through the study of fibroblasts derived from FH patients (47,48). The research demonstrated that fibroblasts deficient in LDL receptors had significantly up-regulated 3-hydroxy-3-methylglutaryl co-enzyme A (HMG CoA) reductase, the rate-limiting enzyme in cholesterol biosynthesis (48). The high risk of coronary artery disease (CAD) in patients with FH, together with the growing body of epidemiologic data associating LDL-C with CAD, validated HMG CoA reductase as an antiatherosclerosis drug target. This led to the development of statin drugs, which are responsible for reducing the morbidity and mortality from CAD by ~30% (49).

Recent successful family-based studies, such as the discovery of associations between genetic variation in LRP6 with metabolic syndrome and CAD and between PCSK9 or ARH and LDL and CAD, continue to highlight the power of Mendelian genetics for drug target identification and validation (13,50,51). Despite the potential of Mendelian genetic approaches, such cohorts are rare and small in size (47,48,50,52), offering valuable but limited information. Population-based human genetic approaches offer a significantly greater opportunity than do Mendelian genetics, but until recently, have been technically challenging. Efforts over the past 10 years have focused on biased, population-based genetic approaches, in which candidate genetic variants from 1 to several genes were evaluated for association with human disease. These approaches have offered little to drug discovery, as the veracity of associations in almost all cases has been unclear because of lack of reproducibility.

The recent completion of the human genome sequence and the HapMap, in addition to substantial reductions in the cost of genotyping and sequencing, have made genome-wide association studies (GWAS) a reality (53). Several GWAS for metabolic and CV diseases have been recently published (9,10,54). Although data from these studies require further evaluation and the ultimate value of GWAS remains unknown, several loci have been associated with LDL, HDL, and CAD. While these associations are encouraging, the physical proximity of a single nucleotide polymorphism (SNP) and gene is only a starting point for considering causality. In drug development, strong evidence for association of the SNP with disease, the identification of the causal gene, and a high probability that a therapeutic agent can be developed to agonize or antagonize a given target (i.e., “druggability”) are necessary to establish a valid drug target. In addition, for a target to be “druggable,” its function must be well understood, as is common with certain enzymes or cell surface receptors.

A recent association was made for a locus on chromosome 1p13 with LDL–C and CAD (9,10,54). The 1p13 SNP lies in close proximity to 3 genes: cadherin-EGF-LAG 7-pass G-type receptor 2 (CELSR2), proline-serine-rich coiled-coil 1 (PSRC1), and sortilin 1 (SORT1). The transformation of a genetic association to causality requires the functional consequence of an SNP that associates with disease. Since most genetic variation occurs in intragenic or noncoding sequences, recent attention has been on “functionating” SNPs by their association with gene expression. One mechanism for ascribing function to an SNP is by establishing its association to gene expression. This is termed an expression SNP (55–60). Functional SNPs (i.e., those that change gene expression levels) can be identified by assessing variations in whole genome gene expression by microarray analyses in a given tissue across many people and then associating expression with genetic variation (61). Interestingly, the SNP at 1p13 that associates with LDL and CAD also has a highly significant association with expression levels of the 3 proximal genes, CELSR2, PSRC1, and SORT1 (9,54,61), strongly implicating this SNP as the causal genetic variation and 1 or all of these genes as causal for CAD. This breakthrough analysis has now reduced the potential underlying gene(s) from many in that region to 3. This narrowing of the likely candidates now allows more focused analyses using knockdown and knockout technologies in vitro and in vivo, which will enable further in-depth investigations around these 3 genes and the pathway they impact, resulting in resolution of the region and perhaps a druggable novel target.

The potential of GWAS was demonstrated by the recent association of multiple SNPs located in a gene-poor region on chromosome 9p21 with type 2 diabetes mellitus (DM) and the risk of MI (54,62,63). These SNPs, although all located very close to each other in the same region, have independent associations to DM and the risk of MI (64). The association of the 9p21 region to MI has been replicated in several studies, all showing an association independent from known risk factors. This finding could lead to the identification of novel nonlipid antiatherosclerosis drug targets. Although the 9p21 SNP is close to several genes (cyclin-dependent kinase inhibitor 2B, cyclin-dependent kinase inhibitor 2A, and methylthioadenosine phosphorylase), none has an obvious connection to CAD nor has any been implicated as causal for the association. Two recent studies suggest that a noncoding ribonucleic acid (RNA), antisense noncoding RNA in the INK4 locus (ANRIL), may be the causal gene underlying the association (64,65), but these data do not provide conclusive evidence of an association. If ANRIL should prove to be the causal gene, additional data will be needed to elucidate how a noncoding RNA transcript is driving risk. Moreover, because ANRIL is not “druggable,” potential drug targets would need to be identified through an understanding of the ANRIL genetic network. Human genetics offers extremely powerful tools for identifying disease causal regions of the genome, but transforming genetic association into disease causality will require significant effort and, in many cases, turning these associations into druggable targets will require more sophisticated systems biology approaches. If a druggable gene is identified in this region through additional analyses, then the therapeutic potential of this region is high. The SNPs in this region could be tested and used as markers to segment
a population according to their risk of MI or type 2 DM, and those patients can be more aggressively treated.

**SYSTEMS BIOLOGY.** Genetics offers a unique opportunity to associate heritable variation with disease; however, the challenges in defining causality (Fig. 1) and the potential that the causal genes are not druggable underscores the need for better genome-mining strategies. One approach that is gaining momentum is systems biology, the science of establishing multiple genomic datasets which, when intersected and queried using sophisticated informatics and statistical modeling tools, can establish causal associations of gene pathways with disease. Systems biology uses not just genetics, but also such datasets as those from multiplexed immunoassay phenotyping, RNA profiling, proteomics, and metabolomics.

A proof-of-principle experiment for a systems biology approach for target identification demonstrated the power of coupling gene expression with genetics (16,55,61,66,67). In an F2 mouse cross with multiple metabolic phenotypes, hydroxysteroid (11-beta) dehydrogenase 1 (Hsd11b1), zinc finger protein 90 (Zfp90), and other genes were found to be causal for metabolic traits. Reconstructed gene networks were used to assign function and position the genes of interest to a specific biological pathway, thereby strengthening the argument of causality for a phenotype (68–70). Causal genes of interest were later tested in either knockout or transgenic mice, or by pharmacologic intervention, confirming the hypotheses established by causality and network analysis. Hypotheses generated from this approach allowed testable studies. For example, when mice overexpressing the gene Zfp90 were analyzed, they had significantly higher fat mass than the controls, as predicted by the systems approach (16). Thus, the use of causality and networks has allowed for the identification of key genes likely to drive disease, separating them from merely reactive ones.

Another recent application of systems biology has strongly implicated oxidative phosphorylation and mitochondrial dysfunction in the pathogenesis of atherosclerosis. In an F2 mouse cross from C57BL/6 and C3H parental strains in an apoE-deficient background, in which atherosclerotic lesions were among the phenotypes scored, gene expression profiling showed that gene expression variation in liver and gonadal fat pads had a major role in atherosclerosis (71). Many gene expression quantitative trait loci in liver and adipose were linked both in trans and cis to atherosclerotic lesions. The association of pathways to atherosclerosis provides a powerful tool for the identification of novel targets. In contrast to this integrative systems approach, the classical approach utilizing genetics alone identifies only loci, not genes or pathways.

Network analysis is another approach that is gaining attention. Networks are informatics-driven reconstructions of complex interactions (68,72–74). Many network analyses for cardiovascular diseases have been restricted to networks derived solely from the literature or whole genome gene expression profiling (75–77); these approaches can provide useful information about how a pathway might be regulated in a given disease state, but do not discern causality, which is critical to the identification of a drug target. Information-rich networks are being developed to overcome the limitations of literature-only or gene-expression-only networks. In particular, Bayesian and gene–gene correlation networks reconstructed from F2 mouse crosses using both genetic and gene expression data have been useful in providing context to a gene of interest and assessing whether it is causal and, thus, a disease drug target (16,66,68,69,78). One study showed that a macrophage-enriched network derived from an F2 mouse cross contained many genes with causal relationships to metabolic phenotypes such as body weight (66); however, additional data from closed systems were needed to elevate these genes from candidate to validated targets. The 3 genes in this network, lipoprotein lipase, lactamase beta, and protein phosphatase 1-like, were validated as obesity genes, through either targeted gene knockdown or transgenic approaches, strengthening the association between this network and metabolic disease traits.

The systems biology and network approaches have been steadily gaining ground for use in mice; however, progress in human samples has been slow, likely due to the lack of human samples available for expression profiling. Nevertheless, recent studies have demonstrated the potential power of using gene expression to identify transcripts that are associated with SNPs in humans. For instance, expression profiles from blood or tissues (e.g., liver) have been used to identify expression SNPs. In addition, studies are under way to combine networks derived from both mouse F2 crosses and humans with liver-derived expression SNPs to identify disease genes (61).

A systems biology approach has also been informative in excluding a potential candidate drug target, C-reactive protein (CRP), which is a biomarker of inflammation and has a strong association with CAD (79). There has been substantial debate as to whether CRP associates with CAD as a disease marker or as a causal factor. If the latter, CRP would be a good drug target. Genetic variation in the CRP locus has effects on circulating CRP levels. Systems biology approaches have been used, leveraging epidemiologic associations of CRP and metabolic syndrome phenotypes together with Mendelian randomization modeling to exclude causality (80–82). In these Mendelian randomization studies, patients are, in a sense, randomly allocated to high versus low CRP arms based on CRP genetic variation known to affect CRP expression. The genetic anchoring provides a basis for segmenting causal (genetic) versus reactive secondary (e.g., inflammatory) causes of CRP level. Because of a less than expected difference in the prevalence of clinical phenotype between the high and low CRP genotype patients, the studies concluded that CRP is not causal (the concept of Mendelian randomization is schematized in
In brief, these powerful studies strongly implicate CRP as a marker of disease, but not a causal factor.

Systems biology will help us to understand complex disease. The identification of SNPs or loci associated with disease is an important first step. The functionation of that SNP is a critical next step and, in many cases, will require a systems biology approach (Fig. 3). Networks will allow us to reduce the information-rich, high-density data that emerge from systems biology into bite-size pieces that are relevant and can be properly processed. Networks and pathways can also provide a priori knowledge about a disease, allowing a pre-selected set of SNPs to be tested in a given cohort, thereby reducing the likelihood of false discovery due to multiplicity. Such innovative genomic approaches will be necessary to maximize the value of GWAS.

**Biomarkers and Proof of Concept**

A major challenge in developing novel CV therapies is the ability to assess clinical efficacy or proof of concept in a...
small, short-duration phase II clinical trial. The ultimate goal of most CV therapeutics is to reduce the risk of future clinical events, such as MI or stroke. Because the assessment of CV event reduction requires expensive, long-term outcome trials, a high probability of success is necessary for a clinical research organization to perform the study and for investigators and patients to consent to the years of clinical testing. Despite assurances, recent CV clinical outcome trials have shown no benefit or even harm, including the ARISE (Aggressive Reduction of Inflammation Stops Events) trial for the direct vessel wall acting therapy, AGI-1067 (succinobucal) (84), and the ILLUMINATE (Investigation of Lipid Level Management to Understand Its Impact in Atherosclerotic Events) trial for the HDL-C-raising cholesteryl ester transfer protein inhibitor, torcetrapib (41). These failures underscore the need for biomarkers that can predict outcomes, reducing the risks associated with testing novel CV therapies.

Genomic approaches are currently under way to develop novel, predictive biomarkers, particularly for plaque vulnerability and HDL function. Plaque vulnerability is an area in which genomics may facilitate proof-of-concept decisions. Given the absence of predictive plaque vulnerability biomarkers, the decision to perform a clinical outcome study on the efficacy of AGI-1067 in reducing plaque vulnerability was made based on data from intravascular ultrasonography studies. Because intravascular ultrasound measures plaque burden, not vulnerability, it may not have been the appropriate tool for assessing the beneficial effects of an agent designed to stabilize plaque. A variety of genomic markers are being considered as predictors of CV events. For instance, gene expression of circulating white blood cells has led to the identification of RNA markers of CAD (85,86). Additionally, correlations between gene expression in mouse atherosclerotic plaque and human coronary disease have been studied (87). These markers will require validation, ideally by longitudinal association with outcome, and then be used to help establish proof of concept for novel therapeutics.

High-density lipoprotein is a highly predictive CAD risk marker; however, the biology of HDL appears to be more complex than that of LDL, and HDL function may be more important than HDL level in driving CAD risk. Proteomic approaches are being used to understand the composition of HDL and identify markers that render HDL atheroprotective. Using proteomics, proteins have been identified on HDL in differential amounts in patients with CAD versus healthy controls, and a subset of these proteins change after statin therapy (88). These observations suggest that the HDL proteome may be useful in identifying novel HDL-raising compounds for proof of concept and movement into clinical outcome studies.

The High Risk Plaque (HRP) initiative is a private-public partnership that was recently formed to advance the understanding, recognition, and management of coronary events (89). The HRP initiative has implemented an important research study known as BioIMAGE, a biomarker validation and discovery study initiated in 2008 to identify novel risk markers for heart attack and stroke (Fig. 4). The

### Figure 4  BioIMAGE Study Design

BioIMAGE is a payor-based observational study to identify imaging and/or circulating biomarkers that predict 3-year cardiovascular (CV) events with incremental improvement over traditional risk assessment (Framingham risk scoring). The study will enroll 7,300 Humana members ages 55 to 80 years (male) or 60 to 80 years (female) from Chicago, South Florida, and Kentucky without known CV disease or other serious medical conditions, of whom 6,000 will undergo imaging studies and 1,300 will serve as nonimaging controls. Physical measurements include electrocardiogram and ankle-brachial index; imaging measurements include carotid artery ultrasonography and abdominal aorta computed tomography (CT) coronary artery calcium scoring in all 6,000 imaging enrolled subjects and carotid artery magnetic resonance imaging (MRI), aorta CT, and carotid/aorta FDG-positron emission tomography/CT in a subset of the 6,000 imaging participants. Plasma, ribonucleic acid (RNA), and deoxyribonucleic acid (DNA) will be collected for genomic and candidate biomarker studies in all participants. The study is expected to complete enrollment in the second quarter (2Q) of 2009 and longitudinal follow-up in 2012. 1Q = first quarter.
study began enrolling patients (~7,300 targeted) in early 2008 and has a target completion date in the first quarter of 2009 (90). Patients undergo an extensive set of baseline biomarker analyses (imaging and circulating) and are monitored longitudinally to track clinical events. This multimodality, information-rich study will provide valuable data on an extensive set of imaging, proteomic, metabolomic, gene expression, and genetic markers that should facilitate proof-of-concept decision making for future CV therapies and novel diagnostics to assess patient risk and therapy response.

**Outcome Trials and Responder Populations**

“Personalized medicine” refers to the use of information and data from a patient’s phenotype, genotype, or level of gene expression to stratify disease, select a medication, provide a therapy, or initiate a preventative measure that is specifically suited to that patient at the time of administration. With approval rates for novel therapies dropping precipitously and the challenges and costs associated with developing novel CV therapies rising steeply, the hope of developing a personalized approach to medicine is a distant reality. A more realistic goal is that of “segmentalized medicine” and, to a certain extent, reflects the current state, in which segments are defined by clinical or biochemical phenotypes. An obvious example is the segmentation of patients with elevated LDL-C or known CAD for treatment with statins. A hypothetical promise of genomics is the ability to identify responder segments based on a molecular, as opposed to phenotypic, understanding of disease.

A potential application of the segmentalized medicine approach is the identification of patients at elevated risk for an adverse effect or who would be more likely to derive benefit from a given therapy. Recent experience with warfarin represents an example in which markers have identified patients at risk for adverse effects. Polymorphisms in a warfarin-metabolizing enzyme, CYP2C9, and its target, vitamin K epoxide reductase complex 1 (VKORC1), significantly affect the warfarin therapeutic window (91); these data have become incorporated into the warfarin product label (92). However, the use of genotype to guide warfarin dosing has not become standard of care because pharmacogenetic determinants of warfarin action account for only ~50% of its variability. Moreover, prospective clinical studies to demonstrate outcome and safety benefit with pharmacogenetic-based prescribing are lacking.

A more recent example of a pharmacogenomic safety biomarker has been the discovery of an association between common variants in solute carrier organic anion transporter family, member 1B1 (SLCO1B1) and statin-induced myopathy (93). SLCO1B1 encodes organic anion transporting polypeptide 1B1, a transporter, for the hepatic uptake of statins. Several studies have demonstrated that genetic variation in this transporter leads to increased circulating statin (94). More stringent guidelines are emerging for LDL-C targets, particularly in high-risk patient populations, driving the prescription of higher statin doses, which are associated with an increased incidence of myopathy. The SLCO1B1 high-risk alleles could be used to identify patients at risk for myopathy with high-dose statins, for whom LDL-C–lowering approaches beyond statins might be preferred.

Methodologically, there are 2 approaches to identify patient segments: retrospective and prospective. In the retrospective approach, a clinical trial is conducted in a broad population and then markers, whether genetic, genomic, or other, are studied to identify in a hypothesis-generating manner a responder population (i.e., patients with the greatest therapeutic index of efficacy over safety). Putative markers are subsequently validated in a prospective, hypothesis-testing clinical trial. This approach was used for the warfarin CYP2C9 and VKORC1 associations; observations were first made in retrospective studies and then, since the genetic variants are frequent and the prothrombin time end point continuous, prospective studies were performed to validate the biomarkers. For the SLCO1B1 variants, the data demonstrating an association to myopathy were retrospective. Although the variants are frequent, the myopathy event is not, occurring in ~0.2% of patients on lower statin (95), making prospective confirmation difficult. Further, given the strength of the association and the mechanistic rationale, it would not be ethical to study this association prospectively.

While the prospective approach appears straightforward, in areas outside of oncology, it has not proven successful. In CVDs, for which it is increasingly necessary to demonstrate outcome benefit to warrant the use of novel mechanisms, this approach is time consuming and costly as it involves going from a hypothesis-generating trial to a hypothesis-testing trial in which the outcome is CV events. Nevertheless, the retrospective approach has resulted in interesting hypotheses that require prospective validation. In a statin outcome trial, PROVE IT–TIMI 22 (Pravastatin or Atorvastatin Evaluation and Infection Therapy–Thrombolysis In Myocardial Infarction 22), a variation in a kinesin microtubule transport gene, KIF6, was identified as a putative marker for statin response. Carriers of a common (~50% carrier frequency) KIF6 719Arg allele were found to have an ~30% increased risk of CAD and a significantly greater response to statin treatment versus noncarriers (18,19,96).

The identification of the KIF6 719Arg allele was made through a candidate gene rather than a genome-wide approach (96), leaving speculation as to whether an unbiased GWAS would yield more and better predictive markers. Although statins produce considerable reductions in CV risk, the residual risk of CAD has been estimated as high as 70% (49). Could the KIF6 “nonresponder” or a yet to be discovered variant define a responder segment for other nonstatin therapies?

The prospective approach may be a more practical one. Here, genomics are applied to dissect disease pathways to identify not only targets but also biomarkers. The biomarkers can then be used to prospectively identify responder
segments and response to therapy. A risk associated with this approach is mismatching of the segment with the drug, but this possibility must be considered against the alternative risk of failing to achieve a viable therapeutic window in the general population.

Summary and Future Landscape

Genomics is impacting CVD drug development today and offers the potential to provide significant future impact. Genomics is already providing demonstrable value for target identification and validation. Genetic efforts in rodent models and humans have provided valuable information with respect to CV drug targets. Progress in using genomics as a tool for discovering biomarkers and improving outcome trial design has been made and, over the next 10 years, the routine application of genomics to clinical decision making and responder segment identification will emerge. For genetics to fulfill its ultimate value proposition, a systems biology approach will be necessary to differentiate causality from association and through pathway analysis to help identify druggable targets. Systems biology will significantly improve our ability to bring well-validated novel targets into the clinic.

The successful application of genomics to CV drug development will require aligned and collaborative efforts of academia, governments, cross-functional industries, and regulatory bodies. Future eras will be dominated by massive advances in technology, information handling, and computational power. Maximizing the value offered by these new technological capabilities will require access to patients, data, samples, and patient consent for genomic studies. This underscores the need for comprehensive biobanks, which are repositories for biological samples (including deoxyribonucleic acid) from a representative portion of a human population with associated information describing the people to whom the samples belong.

Several biobanks exist and others are under development. The United Kingdom Biobank intends to archive data and samples on 500,000 patients (97). The Nordic countries have pioneered biobanking and have established several, including 1 in Iceland that began as part of the Reykjavik Heart Study utilized by deCODE Genetics to successfully identify several CVD genes (98). The Japanese Biobank started in 2003 with the purpose of collecting deoxyribonucleic acid, serum, and clinical information from ~300,000 subjects in Japan (99). While these represent an important first step, efforts are needed to establish larger, information-rich biobanks, in which systems-based genomic approaches will help identify new CVD targets and biomarkers with which prospectively enrolled clinical trials can be conducted in targeted patient segments. Indeed, the ability to screen a biobank for a response or safety marker in advance of recruitment will help patient segments. Indeed, the ability to screen a biobank for a response or safety marker in advance of recruitment will help patient segments.

which we will be able to apply advancing genomic tools and leverage systems biology approaches to samples and data, we will find novel and well-validated targets to pursue new CVD therapies. If established correctly with the appropriate consent, these same biobanks will serve as a source for identifying and studying the patient segments that will benefit the most from these novel therapies. To work, this undertaking will require time and the collaborative efforts of governments, academia, patients, patient advocacy groups, payers, and industry.

Acknowledgments

The authors gratefully acknowledge Patricia Masarachia and Allyson Debes of the Merck Information Management Department for their help with retrieval of information on approval dates for currently marketed therapies. The authors also thank Christine McCrary Sisk for editing and critical review of the manuscript and Joerg Koglin for helpful comments.

Reprint requests and correspondence: Dr. Andrew S. Plump, Cardiovascular Diseases, Merck Research Laboratories, 126 East Lincoln Avenue, Rahway, New Jersey 07065. E-mail: andrew_plump@merck.com.

REFERENCES


33. Boden WE, Pearson TA. Raising low levels of high-density lipoprotein cholesterol is an important goal of therapy. Am J Cardiol 2000;85:645–50.


