The Future of Cardiovascular Medicine

The Future of Genetics and Genomics
Closing the Phenotype Gap in Precision Medicine

Calum A. MacRae, MD, PhD; Ramachandran S. Vasan, MD

In this forward-facing review, we briefly summarize what we see as the state of modern cardiovascular genetics and genomics before outlining some of the key areas in which we believe there is a need for investment if we are to realize the full clinical potential of the field. There is clearly substantial overlap with the field of precision medicine, yet it may be in understanding where genetics and genomics currently fall short that we will identify how precision medicine might move forward. The recognition over the last decade that genetics or genomics alone is unlikely to offer the predictive utility necessary for precision medicine is matched by our dependence on genetics and genomics for much of the mechanistic insight necessary to achieve precision diagnostics or precision therapeutics. The overarching strategies we discuss emphasize that the future of cardiovascular genetics and genomics is inextricably linked with those of every other field of medicine. We must move beyond the silos of traditional organ-based disease and recognize fully the need for generalizable biological rules that constitute the truly mechanistic medicine or molecular medicine first envisaged >40 years ago.

The Current Impact of Cardiovascular Genetics

The cardiovascular arena has been a focus of much work on the practical application of genetics. Given the potential for coincident presentation and demise in sudden death and the availability of the implantable defibrillator (a mechanism-agnostic preventive therapy), cardiovascular medicine has been focused on the predictive utility of genetics for some time. The long tradition of physiological measurements and classic epidemiology in cardiovascular research has facilitated large genotype-phenotype correlation studies. Cardiovascular research has also had to tackle truly complex paroxysmal phenotypes such as arrhythmias that have challenged our ability to resolve traits into their constituent components.

Large kindreds, with what in retrospect seems to be relatively exceptional penetrance, have been the focus of pioneering work in the identification of causative genes and in ongoing efforts to define disease mechanisms. However, more typical families with inherited contributions to disease are smaller and considerably less informative, and there has been a proliferation of spurious “mutations” in spurious disease genes, often based on simple “guilt by association” without rigorous demonstration of mechanistic involvement. In addition, even those mutations definitively characterized in one family may have no discernable phenotype in another family, as recently demonstrated in work from Framingham and other population cohorts. These observations emphasize how important it will be to develop techniques to establish a causal role for specific variants in specific diseases in individual families or patients if genetics is to be able to truly affect clinical care directly. Although we have gleaned enormous etiologic insights, it has also become clear that using simple genotypes to specify diagnoses or to drive therapies (in individual patients) is far from reality at present.

Such problems with genotype-phenotype correlation are most obvious in disorders with profound selection pressures and thus a tendency to be associated with high de novo mutation rates, often in multiple highly conserved genes with resultant allelic heterogeneity. In some instances, the imprecise relationship between genotype and phenotype reflects real differences in the mechanism of action of different variants within the same gene. These observations also imply that the current phenotypic repertoire of clinical medicine may not suffice to allow us to understand fully the complexity of genotype-phenotype relationships.

Similar problems with prediction are found in cardiovascular disease in which negative genetic selection pressure is either more or less extreme. In the setting of severe congenital heart disease, in which the survival of the organism during the neonatal period is at risk, despite familial recurrence rates consistent with large-effect-size single genes, few if any cases have a “sufficient and necessary” genetic explanation. In this setting, gene–gene or gene–environment interactions are often raised as possible mechanisms for this discordance. Whole-genome or -exome approaches reveal the true scale of the genotype-phenotype problem while conditioning environmental stimuli (intrauterine and extrauterine) are almost invariably unknown and unmeasured. When genetic selection pressures are less acute (eg, coronary artery disease or hypertension), in addition to unmeasured environmental factors, the aggregation of different origins, reinforced...
by the success of epidemiological studies or clinical trials, and the low specificity of negative diagnoses all conspire to make clinical genetic interpretation challenging.19,20

A fundamental problem with the clinical application of genetics in cardiovascular medicine is that when there is enough genetic information to attribute causality to variants in individual families or family members, phenotype alone almost always conveys most of the information content; consequently, at present, genotype is of little incremental utility. Conversely, when there is phenotypic uncertainty, genetics, no matter how comprehensive, is unlikely to support any substantive etiologic insight.21 The inference from these generalizations is that more robust phenotyping protocols and technologies capable of detecting subtle deviations from wellness in steady state and in response to perturbations or the development of mechanism-specific therapies will be required to achieve the full potential of modern genetics.

The Current Impact of Cardiovascular Genomics

Large cardiovascular population studies and clinical trial data sets have formed the basis of numerous genome-wide association studies (GWASs), exploiting the utility of modern genomic technologies for cheap and efficient genotyping of common alleles. These approaches have identified hundreds of loci contributing to a range of continuous phenotypes from plasma lipids to ECG parameters.6,7 Binary phenotypes such as atrial fibrillation have also been successfully studied, and the chromosome 4q25 locus for this arrhythmia is among the largest-effect-size loci identified for any trait.22 Importantly, for the vast majority of identified GWAS loci, the underlying mechanism of their contribution to disease is unknown currently, at least partially reflecting the fact that in most instances these alleles rarely contribute more than a small proportion of the observed heritable variance for the trait.23 Here again, gene–gene or gene–environment interactions are rationally invoked, but there have been few systematic efforts to address the relevance of each of these mechanisms for individual traits. Ever larger GWASs seem unlikely to address these concerns because the scale of study required to understand interacting small effect sizes in the setting of unmeasured conditioning environmental variables is prohibitively costly. Once again, the inference is that we require more rigorous, proximate, quantitative phenotypes and an approach to the objective measurement of environmental factors at a scale and degree of comprehensiveness compatible with modern genomics.

Genomic measures of common variation have considerably greater rigor for prediction than the highly pleiotropic Mendelian variation, but current limitations for this purpose are the lack of lifetime risk prediction models, the incremental nature of the genetic information over traditional risk factors, the very modest effect sizes, and the consequent absence of mechanistic insight. It is also often posited that because loci encoding therapeutic targets have been identified by GWASs, these studies will offer insights into the choice of drug targets. The instances in which this has occurred appear to derive from biological roles for these genes that were discovered as a result of large-effect rare variants in the same genes.24 GWAS loci may reflect the role of common variation in these genes in population risk, but it will be difficult to discern which of the hundreds of GWAS loci for each trait encodes such a gene without comprehensive understanding of the genetic and environmental architecture of each trait. Nevertheless, GWAS data may be useful for the evaluation of potential therapeutic targets identified by other means or for the assessment of potential pharmacological safety concerns.25

Genomic technologies have spawned other large and unbiased data sets, including transcriptomics, metabolomics, proteomics, and lipidomics. Despite some remarkable insights from these technologies, the cardiovascular community has been slow to incorporate these new profiling technologies into the clinical arena. This contrast with the emerging and relatively uncritical approach to “biomarker as a diagnosis” status for traditional biomarkers (such as troponin T) is difficult to explain. Defining cost-effective ways to move granular genomic approaches to the clinic will be a vital component of the phenotyping revolution that will be necessary to execute the vision of precision medicine.

The Major Challenges

Several consistent themes appear to be emerging from failed efforts to reduce genetic and genomic data to clinical utility over the last few years. In the main, these represent an understandable reliance on the traditional disease syndromes, the success of many current phenotypic paradigms in other study designs such as clinical trials, and equally successful therapeutic strategies. Ironically, the success of prior paradigms may become a major driver for change as the acute syndromes around which modern cardiology has been built plateau or slowly dwindle in importance, chronic disease becomes more prevalent,26 and earlier detection and definitive cure become the benchmarks against which medical science is judged. Some of what we believe are the important elements for realizing a genomic vision are outlined below, but, these are all intimately related to the need for a dramatic and step-wise increase in the personal information content assembled and deployed for medical benefit.

Genetic Architecture

To be able to accurately model a trait and thus predict its clinical behavior and downstream consequences requires a refined understanding of its underlying genetic and nongenetic architecture. Combining genetic epidemiology, molecular genetics, and model organism insights, Chakravarti and colleagues transformed our insights into the causes of Hirschsprung disease from a presumed environmental insult to an oligogenic disorder in which almost all of the trait can be attributed to variation in a few genes acting and interacting in a single pathway.27,24 In comparison, for very few other human conditions are we able to describe even the most basic attributes of the genetic architecture. Indeed, when such information is available, it often runs counter to current strategies for that specific trait in human genetics.29 For many disorders in which the focus is on defining the mechanism of small effect GWAS signals, what extant genetic epidemiology exists suggests that major Mendelian genes are at play and that common alleles possibly play a modest role.
Gene–Gene and Gene–Environment Interactions

The power of human genetic studies to define gene–gene interactions is limited by the sheer size of the populations that would be required to encompass representative alleles of different effect sizes and to characterize quantitatively the effects of discrete allelic combinations on the ultimate phenotype. It would seem reasonable to assume that only by designing studies that include multiple independent orthogonal phenotypes will scientists be able to accomplish such an undertaking. This strategy would also have the benefit of offering potential phenotypic classifiers that would identify unsuspected disease subsets, with consequent improvement in the etiologic homogeneity that renders family-based genetics so powerful. Finding ways to incorporate such a gargantuan project into routine preventive or clinical care is critical for the success of precision medicine and will involve leveraging the electronic health record for truly meaningful use.

Without conditioning environmental variables, including infectious exposures, many major genetic effects are simply inaccessible. The investment required to bring environmental insight to disease has lagged behind genomics in remarkable ways. Although enormous data sets exist in almost every venue of our lives and electronic inventory of millions of items traveling across the planet is part of the supply chain in many companies, the ascertainment of environmental factors for understanding health and disease has been slow to grow in comparison. Exploiting social and other network data, maps of exposures for individual geospatial locations, and potentially even transactional or other federated data from numerous contexts in our lives may seem like a forceful intrusion, but these approaches are commonplace in commerce or financial security and will be essential to dissect disease mechanisms even in a modern genomic context. There is no doubt that security and will be essential to dissect disease mechanisms these approaches are commonplace in commerce or financial industries in which it has occurred to date. Importantly, digital phenotypes have universal standards that would meet such simple criteria. Electrocardiography is one notable exception.

At the core of this redesign of the phenotyping enterprise are several key requirements. Unless we attempt to move our assessment of disease more proximately in the pathobiological chain, we will not achieve the change in resolution that is necessary to elucidate the genetic architecture of cardiovascular traits. The contrast between deep phenotyping and broad phenotyping is often drawn, yet these distinctions are artificial. We may need deep phenotyping to understand mechanism in a specific disorder, but we will need broad phenotyping to understand the heterogeneity of disease and to capture the full potential of genomics. The ideal phenotypes will be standardized, continuous, and linear over several log orders; will have cellular representation; and will explicitly balance depth and breadth. It is perhaps salutary to note how few modern clinical phenotypes have universal standards that would meet such simple criteria. Although not all phenotypes will exhibit these attributes, to integrate the full translational cycle, it will be important to be able to measure identical end points in the full spectrum of cellular or animal models and to do so both under basal conditions and with a standardized perturbation. We also must consider the power of longitudinal data sets extending from baseline wellness into the very earliest deviations from that state into the spectrum of disease. Avoiding the trap of the “final common pathway” is perhaps the most obvious imperative for genetics and genomics in the next few years.

Drug exposures are perhaps the most embarrassingly underexplored resource in our current phenotyping armamentarium. The development of arrays of stimulus-response pairs to describe the state of a system will itself begin to structure phenotypic data into a more comprehensive description of the entire organism. With such an approach layered on top of traditional or binary traits, orthogonal relationships may become apparent that will efficiently stratify disease long before there is mechanistic understanding of the underlying differences. An early example from other settings is the addition of seizures or hearing loss to autism classifications, but we are limited by the scope of current clinical data sets.

Central to the entire notion that we describe is the idea that phenotypes by design must be translatable and computable. That this type of data collection would be useful or feasible in certain situations is not in question, but its general utility and the means by which this approach might be tested will doubtless be subject to skepticism in traditional biomedical circles. As a result, much of this transformation will likely emerge from outside conventional professional activity, as has happened in the industries in which it has occurred to date. Importantly, digital phenotypes are part of the missing link in the implementation of the electronic health records for detecting disease associations (sometimes referred to as phenome-wide association studies).

The Phenotype Gap: Broad or Deep

As noted earlier, without a more granular phenotypic resolution, it will not be feasible to deconvolute genomes, transcriptomes, or proteomes in any systematic way. Layering in other levels of complexity such as exosomes and cellular connectomes will be feasible only if we change the approach to clinical and personal phenotyping. A focus on “legacy” traits such as blood pressure or heart size, which offer limited stratification but have deep historical roots, has essentially fixed human phenotyping in a past era. Although we have been able to improve the resolution of noninvasive assessment of many of these metrics, the scope of our phenotyping vision has lagged far behind. Only with the advance of external economic forces such as the availability of personal sensors in every smart phone and the quantified-self movement have medical professionals begun to reassess the scope and rigor of phenotyping. Medicine requires a truly revolutionary reappraisal of what we measure to monitor wellness and health.

Rigorous Models With Predictive Utility for Efficient Translation

An interesting phenomenon in contemporary biology is that many of the genomic data sets collected in cells or model organisms over the last few decades have yet to be rigorously integrated into clinically meaningful use. This appears to be a result of barriers that escape rationale, including the
Mechanism and Therapy

As we understand mechanism, our ability to discover and develop therapies is transformed. Genetic models have had remarkable success in rare diseases but have yet to penetrate most of the common disorders that afflict humans. A major investment in phenotyping might accelerate the application of genetics and genomics and the identification of unsuspected disease mechanisms, which then lead to the discovery of novel therapies. To achieve this goal at a truly personal or precise level will mandate a change in the paradigms of drug discovery. When integrated with the evolving pharmaceutical fields of high-content, phenotype-driven screening and target identification, it will become increasingly feasible to use disease itself, rather than one specific pathway downstream in the causal chain, as the effective target in drug discovery.40,41 Although not supplanting traditional drug discovery and development paradigms, these and other innovations will have to change the scale and efficiency of drug discovery by close to an order of magnitude if society is to be able to afford precision medicine. Without precision therapy, precision medicine will be of limited utility.28,29 Changing the phenotypic resolution of non cell-autonomous disease would allow the concepts to be extended to many more disorders, including common cardiovascular disorders, with the potential to accelerate translation.

Conclusions

Genetics and genomics have transformed our understanding of biomedicine but, despite large investments, have yet to directly affect the vast majority of patients. This appears to be a result, at least in part, of barriers to the translation of fundamental insights both to and from patient care. In this review, we have suggested that the current alignment between forces inside and outside medicine to redesign clinical care can be harnessed to reinvent genetics, genomics, and translational biology if the appropriate insights are incorporated. This effort will require shared vision, shared technology, and shared vocabulary, but perhaps most important, it will require the recognition that research and development are an integral part of a learning health system. Cardiology has a unique legacy, including remarkable efforts in prevention and wellness, data-driven care, quantitative modeling, and even remote monitoring, that position us well to be at the forefront of the revolutions in both clinical care and translational genomic science that precision medicine promises.

Sources of Funding

This work was supported by the National Health, Lung, and Blood Institute, Framingham Heart Study (National Health, Lung, and Blood Institute/National Institutes of Health contracts N01-HC-25195 and HHSN268201500001I), and National Human Genome Research Institute (U01 HG007690, U01 HG006500).