Genomics, Transcriptional Profiling, and Heart Failure

Kenneth B. Margulies, MD,* Daniel P. Bednarik, PhD,† Daniel L. Dries, MD, MPH* Philadelphia, Pennsylvania; and Vancouver, British Columbia, Canada

After completion of the Human Genome Project, rapid and ongoing advances in genomic profiling and basic biological insights have led to a more complete and sophisticated understanding of interactions among genes, environment, and affected tissues. As a result, we are deriving increasingly comprehensive views of patients’ genetic endowment and transcriptional changes within organs and tissues affected by complex and heterogeneous conditions such as heart failure (HF). In this context, this review will summarize the ways in which progress in molecular genetics and genomics is providing new insights into mechanisms contributing to cardiomyopathy and the development of HF. We will also highlight how increasing capacity to characterize patients’ deoxyribonucleic acid (DNA) and markers of gene expression can propel new diagnostic and therapeutic opportunities for HF patients. When appropriate, these opportunities will be considered in the broader context of ongoing efforts to achieve greater individualization of prevention and treatment strategies to address the HF epidemic.

Genetics, Genomics, and Gene Expression Profiling

In their primer on genomic medicine, Guttmacher and Collins (1) defined genetics as “the study of single genes and their effects” and defined genomics as “the functions and interactions of all the genes in the genome.” In a subsequent review, Khoury (2) emphasized that this distinction implies more than just a difference in the number of genes under consideration, but also a recognition that “genetic information at 1 locus is modified by information at many other loci and their interaction with non-genetic factors.” In this view, gene expression profiling is a means of integrating the impact of DNA-based effects and a variety of other factors that modulate cellular responses. In considering the distinction between genomics and gene expression profiling, one must recognize that a patient’s DNA can be sampled from any cell in the body and is assumed to be the same throughout life, whereas gene expression profiling examines the expression (abundance) of many different ribonucleic acids (RNAs) and/or proteins in a specific circumstance. In contrast to a patient’s DNA, gene expression varies greatly on the basis of dynamic conditions, including time, stress, disease, treatments, and other factors. Moreover, gene expression is tissue specific so that defining gene expression in an organ or cell of interest requires sampling for ex vivo analysis.
Genetic cardiomyopathies. Among primary cardiomyopathies, elucidation of the genetic basis is most advanced for hypertrophic cardiomyopathies. To date, >450 different heritable mutations have been identified as capable of inducing increases in wall thickness in the absence of increased hemodynamic load with attendant risks of HF, dynamic outflow tract obstruction, and arrhythmia (3). It is currently estimated that 50% to 70% of hypertrophic cardiomyopathy cases are directly caused by single gene mutations with variable penetrance. Most of these mutations affect cardiac myocyte sarcomeric proteins, exhibit autosomal dominant inheritance, and do not directly impair intrinsic contractility. For most of the genes implicated in primary hypertrophic cardiomyopathy, several different mutations have been reported (4). Indeed, it is not unusual for affected families to share a so-called “private mutation” that is distinct from all others previously reported.

A growing number of rare mutations have been shown to produce a hereditable dilated cardiomyopathy. A primary genetic etiology has also been inferred from, and sometimes identified in, cases of arrhythmogenic right ventricular dysplasia and left ventricular noncompaction cardiomyopathy. Even in dilated cardiomyopathies with less distinctive morphologies, recent analyses suggest that 20% to 30% of cases are directly caused by single gene mutations; these are often related to cytoskeletal or metabolic function. As with hypertrophic cardiomyopathies, many different mutations have been associated with genetic dilated cardiomyopathies. Interestingly, mutations in some of the same genes can cause either a dilated or hypertrophic cardiomyopathy. For example, cardiac troponin T, titin, myosin-binding protein C, and tropomyosin mutations can be observed with either hypertrophic or dilated cardiomyopathies. Analogous phenomenon is reflected in genetically identical transgenic mouse models that strongly pre-dispose to congenital abnormalities of cardiac and great vessel structure, but do not always produce the same anomaly (5). These observations support the hypothesis that DNA mutations may not fully define the particular phenotype that results, and that so-called epigenetic factors or gene-environment interactions may alter the expression of primary mutations.

Disease-susceptibility genes. In addition to the relatively uncommon disease-causing mutations discussed in the preceding text, more common DNA sequence variations may have substantial impact on patients’ susceptibility to disease. Most of these variations consist of so-called single nucleotide polymorphisms (SNPs) that occur when a single base pair in the genome differs among individuals. Although some SNPs are relatively rare, many have a more balanced frequency in certain populations. In fact, SNP frequencies often vary between populations, so an SNP common in 1 geographical or ethnic group may be much rarer in another. In a given population, the term “minor allele” is often applied to the less common variant. So-called nonsynonymous SNPs are variants in the coding sequences that affect the polypeptide sequence once a gene is transcribed and translated. However, SNPs in the noncoding regions of genes can still affect gene splicing and transcription factor binding. Most importantly, variations in the DNA sequences can affect disease occurrence and responses to pathogens, chemicals, and drugs.

With respect to disease-modifying SNPs relevant to HF, the case of corin is a particularly instructive example. Corin is a transmembrane serine protease, located within cardiac myocytes, that processes the natriuretic peptide precursor molecules into biologically active forms. The gene encoding corin contains 2 separate SNPs (T555I and Q568P) that are always found together (referred to as complete linkage disequilibrium). The corin minor allele is inherited in ~14% of black subjects and <1% of white subjects. In 2 separate large human cohorts, the corin minor allele was associated with higher blood pressure and an increased risk for prevalent hypertension (6). Subsequent studies further demonstrated that for any given level of blood pressure, left ventricular mass was higher among patients with the corin minor allele (7). These findings indicate that the corin minor allele pre-disposes to hypertension and additionally increases the likelihood and severity of cardiac hypertrophy in patients with systemic hypertension.

Perhaps most exciting, technologies affording a rapid, comprehensive analysis of the roughly 500,000 gene variants defining each patient’s genetic individuality are now propelling new and powerful efforts to identify the genetic contributions to disease susceptibility in the absence of an identified familial inheritance pattern. Surveys of genetic variation in patients with and without a given disease have been termed genome-wide association studies (GWAS). As recently reviewed, GWAS have already identified new genetic contributions to a patient’s susceptibility to coronary artery disease and myocardial infarction (8). Greater disease frequency or severity in patients with 2, rather than 1, copies of a particular polymorphism may further validate its pathophysiological role. Beyond elucidating genetic biomarkers of disease risk, the specific genetic variants identified by GWAS also provide new biological insights by identifying pathophysiological processes whose contributions had been underappreciated and by clarifying how known risk factors pre-dispose to disease development. These insights, in turn, may be exploited by intensifying primary or secondary prevention efforts using existing interventions. In this regard, ongoing progress in elucidating the genetic basis of susceptibility to myocardial infarction, hypertension, and type 2 diabetes mellitus will allow clinicians to apply appropriate interventions that should delay or prevent the
development of cardiomyopathies resulting from these conditions (8). Conversely, with most cases of nonischemic dilated cardiomyopathy still classified as idiopathic on the basis of current criteria, this condition seems particularly well suited for identifying genetic contributions to patient susceptibility that might be exploited for prevention and treatment. Ultimately, GWAS-derived insights into pathophysiology will enable new therapeutic strategies that will specifically target new interventions to the subset of patients whose genetic pre-disposition to cardiomyopathy also renders them most likely to benefit from the intervention.

**Lessons About Disease Biology**

**From Genomics and Expression Profiling**

**Insights into molecular integration.** The growing capacity for broad-based characterization of gene expression has, in many ways, informed our understanding of HF pathophysiology. For example, the combination of transgenic mouse models and transcription profiling has provided important insights that are relevant to both disease pathogenesis and transcriptional profiling in other settings. Often a single transcriptional change in the mouse heart triggers adaptations within several functionally different pathways that may or may not be intuitively related to the overexpressed or underexpressed molecule. For example, Tang et al. (9) reported that cardiac-specific tumor necrosis factor-alpha overexpression induces dysregulation of >1,000 genes even at a stage of model development characterized by compensatory hypertrophy without significant contractile dysfunction or left ventricular dilation. The consistent observation that a single molecular manipulation induces multiple and varied transcriptional adaptations supports the general concept that myocardial cellular biology is highly integrated. At the same time, novel informatics analyses of the diverse genes activated in a given disease state may provide insights into which signaling pathways (9) or transcription factors (10) are exerting dominant effects on the diseased heart.

**Time and severity effects.** Another lesson from profiling transgenic mouse models concerns the time dependence of the transcriptional adaptations to the altered transgene. For example, Aronow et al. (11) showed that there were no dysregulated genes common to 4 separate cardiac hypertrophy models profiled when their phenotypes remained relatively well compensated, but there were more dysregulated genes and greater overlap as the phenotypes of the transgenic animals worsened with age. This transcriptional convergence with more advanced disease includes many of the “typical” transcriptional adaptations often observed in acquired models of HF, including up-regulation of fetal genes such as atrial natriuretic peptide, B-type natriuretic peptide, beta-myosin heavy chain (MHC), and vascular smooth muscle actin. Such findings support the general contention that many of the transcriptional changes within the severely failing myocardium, including the human heart, are a consequence, rather than a cause, of HF. In recent years, pre-clinical studies have also identified previously underappreciated, but functionally important, levels of control over gene expression that are enhancing our understanding of myocardial biology and that affect the types of expression profiling that can be utilized for clinical diagnostic and therapeutic purposes. Three important levels of regulation are discussed in the following sections.

**Epigenetics.** The term “epigenetics” refers to the heritable regulation of gene expression through modification of chromosomal components without an alteration in the nucleotide sequence of the genome. Often epigenetic modifications are acquired structural changes in DNA and chromatin that may affect disease susceptibility, responses to pathological stress, and associated levels of messenger ribonucleic acid (mRNA). Although there are varied types of epigenetic influences, all represent ways in which extrinsic factors can modulate the expression (transcription) of genes, even though the protein-encoding regions of the genes themselves do not change. Work by Kaneda et al. (12,13) probed gene regulation and their modifications including CpG methylation of genomic DNA as well as acetylation, phosphorylation, ubiquitination, and SUMOylation of core histone proteins. One type of epigenetic change clearly shown to modify gene expression and phenotype is the balance between histone acetylation and deacetylation. For example, studies by Olson et al. (14) have revealed key roles for histone deacetylases in the control of pathological cardiac growth. Investigations focusing on the network of genes that are under abnormal epigenetic regulation in failing hearts will ultimately facilitate the development of new epigenetics-based therapies (12,13). Particularly for DNA methylation, the evolution of rapid and direct methods to profile methylation patterns, CpG islands, and their relationships to histone structure, and expression should expedite the epigenetic profiling of pivotal HF genes.

**Transcriptional splice variation and functional protein isoforms.** The complexity of the genome resides not only in the many genes comprising it, but also in the many levels of processing that influence the proteins that are produced and their abundance. Indeed, the nearly complete coverage of the human genome indicates that it encodes only 20,000 to 25,000 protein-coding genes. This number is significantly below that estimated from the complexity observed at the mRNA level through analyses of expressed sequence tags (between 100,000 and 150,000). This “gene shortfall” has raised a great interest in alternative RNA splicing, which could account for the discrepancy between the number of genes and protein isoforms (15). In terms of cardiovascular genomics, splicing variants have been reported for several transcripts, including phosphodiesterases and troponin T (16–18). Figure 1 illustrates the basic dynamics of alternative mRNA splicing. An alternatively spliced transcript may encode a protein that is missing a specific functional domain, thereby altering its activity. The frequency of specific splice variants can be defined by disease
states, routes of cellular differentiation and phenotype, shifts in metabolic pathways, occurrence of SNPs in splice donor/acceptor motifs, and other physiological stimuli.

The significance of splice variation is well illustrated by alternative splicing of the phosphodiesterase gene family (18). Here, alternative splicing alone allows a single gene to have multiple different mRNA transcripts producing proteins that may differ substantially from each other, even to the extent of having opposing effects (15). Other examples where alternatively spliced mRNAs and their products influence, or are the result of, HF pathology have been reported (19). Recently, Qiu et al. (20) demonstrated that a pro-proliferative and anti-apoptotic splice variant of the cell-cycle–related kinase transcript that is present during conditions of compensation is down-regulated during the transition to heart failure subsequent to myocardial infarction or 3 weeks of tachypacing. In this case, 2 splice variants differ from each other in terms of protein-protein interactions, substrate specificity, and regulation of the cell cycle. The ability to perform genome-wide splice variant profiling using existing commercial platforms will likely extend our understanding of the functional importance of alternative splicing in the development and regulation of HF pathology (21,22).

Microribonucleic acid (miRNA) profiling. Profiling of miRNAs represents another new facet of gene expression profiling that also shows promise for enhancing our understanding of human HF while providing patient-specific insights that may inform clinical prognostication and therapeutic strategies. miRNAs are recently-discovered RNA molecules containing 21 to 22 nucleotides that regulate the expression of target mRNA within virtually all cell types. Because individual miRNAs may affect the expression of many mRNA targets (23), they are capable of inducing coordinated and potent effects on cell function. In fact, miRNAs have been implicated in developmental processes, cell proliferation, differentiation, stress-responses, apoptosis, and oncogenesis.

Although relatively few studies have specifically examined the regulation of cardiac structure and function by miRNA, the results have been compelling. For example, in mouse models of chronic LV pressure overload and genetic cardiomyopathies, 3 separate studies (24–26) have demonstrated altered regulation of miRNAs that appear to be necessary for the development of cardiac hypertrophy in response to pathophysiological stress. In 2 of the studies, manipulation of a single miRNA was sufficient to produce or suppress cardiac myocyte hypertrophy in vitro and in vivo (24,26). From a mechanistic standpoint, particularly intriguing studies have demonstrated that a highly conserved miRNA (miR-208) encoded by an intron of the alpha-MHC gene plays a pivotal role in regulating the expression of beta-MHC in states long known to be associated with a change in the relative abundance of alpha- and beta-MHC, such as pressure overload and hypothyroidism (25). Interestingly, miRNA-208 knockout mice unable to up-regulate beta-MHC were resistant to hypertrophy and fibrosis during chronic pressure overload, but age-dependent defects in sarcomere structure and cardiac performance developed. These studies demonstrate a functionally significant role for
a cardiac-specific miRNA in the elegant coordination of contractile protein expression in normal and stressed hearts.

**Diagnostic Applications of Genomics and Transcriptional Profiling**

**DNA-based screening.** Multiplexed assays are already available to screen for known cardiomyopathy-causing mutations in affected patients and their families. In hypertrophic cardiomyopathy, the utility of genetic screening for genetic counseling and to complement clinical screening of family members has been established in previous reports (27,28). Because most cases have a primary genetic basis, the available screening methods have a 60% likelihood of detecting the genetic basis for a patient with hypertrophic cardiomyopathy; the diagnostic yield will gradually increase as more disease-associated mutations are identified (29). Although screening of patients with dilated cardiomyopathy currently has a <30% yield, rates will likely increase as more mutations are identified. Compared with echocardiography, an advantage of screening for disease-causing mutations is the identification of genotypes, such as troponin T mutations, associated with particularly high risk of premature death, HF, or lethal arrhythmias (29).

Although identification and validation of disease-modifying mutations and polymorphisms is still in its infancy, it is likely that comprehensive screening for variants that may affect disease susceptibility will ultimately be accessible to clinicians. As with disease-causing mutations, these panels will have the capability to identify patients and family members at increased risk for adverse outcomes using only a blood sample. In contrast to more limited genomic analyses, these comprehensive screens will also represent a productive platform for discovering new associations between DNA variations and disease susceptibility, severity, and prognosis. The existence of effective disease-modifying therapies for pre-symptomatic patients, such as beta-blockers and angiotensin-converting enzyme inhibitors, will provide an impetus for earlier identification of genetic risk.

**Transcriptional profiling in transplantation.** Because endomyocardial biopsy (EMB) is routinely performed to monitor cardiac allografts for rejection, it has provided a convenient starting point for demonstrating the utility of myocardial transcriptional profiling for defining disease presence and prognosis. Early studies by Zhang et al. (30) demonstrated that detection of cell-adhesion molecules by competitive polymerase chain reaction was predictive of allograft rejection, and Shulzhenko et al. (31) found that detection of genes encoding cytotoxic T-cell activation precedes histological evidence of rejection and adds diagnostic sensitivity to histological criteria. Subsequently, Bayliss et al. (32) showed that graded increases in vascular endothelial growth factor mRNA expression in EMB specimens were associated with both a more severe histological score and a substantially greater risk of cardiac-related death. Finally, altered myocardial expression of specific transcripts (allograft inflammatory factor or angiotensin II receptor) has been identified as a biomarker of allograft coronary arteriopathy (33,34), the leading cause of graft failure beyond the first year after transplant. Thus, transcriptional profiling of allograft biopsies provides important pathophysiological and prognostic clues beyond the simple detection or exclusion of cellular rejection provided by the histological grading currently employed.

Application of EMB to detect allograft rejection in cardiac transplant recipients also provided an opportunity to explore and validate peripheral blood biomarkers of allograft rejection. In peripheral blood mononuclear cells from blood samples obtained at the time of surveillance EMB, Horwitz et al. (35) identified 40 transcripts with both a favorable dynamic range and differential expression in patients with and without clinically significant allograft rejection. These studies also demonstrated regression of pathologic expression patterns in blood leukocytes in association with histological regression of rejection after treatment of rejection episodes. Subsequently, Deng et al. (36) reported that an 11-gene polymerase chain reaction–based assay of peripheral blood cells was able to distinguish biopsy-proven moderate/severe rejection with excellent sensitivity and good specificity, while others suggested that peripheral blood transcriptional profiling may improve diagnostic accuracy when histological grading alone may underestimate the severity of rejection (37). Accordingly, peripheral blood cell transcriptional profiling shows great promise for reducing the need for EMB while enhancing diagnostic accuracy in heart transplant recipients. This approach of early invasive discovery and validation permitting a subsequent less-invasive diagnostic strategy is clearly relevant beyond the transplant population.

**Transcriptional profiling in cardiomyopathy.** The fact that transcriptional profiling of the myocardium requires the relatively invasive procedure of native heart EMB tends to reduce its diagnostic application. Nevertheless, recent consensus recommendations highlight a variety of clinical scenarios in which diagnostic native heart EMB is clinically appropriate (38). Because several of the strongest indications for EMB are in circumstances with relatively recent onset of unexplained HF and cardiomyopathy, the potential for enhanced diagnostic and/or prognostic specificity with transcriptional profiling is particularly great. Beyond distinguishing the transcriptional profiles of ischemic and non-ischemic cardiomyopathy (39–42), other studies have reported distinctive myocardial transcriptional profiles from patients with inflammatory cardiomyopathies (43) and giant cell myocarditis (44). Perhaps most importantly, profiling of early myocardial disease by mRNA and miRNA analyses have already begun to provide insight into the prognosis, severity, and pathophysiology of nonischemic cardiomyopathies (45,46). Given that RNA amplification technology now permits a full microarray analysis from a single EMB specimen (40,47), the added diagnostic value provided by transcriptional profiling could eventually shift the risk-
benefit ratio further in favor of early diagnostic EMB, particularly for patients with nonischemic cardiomyopathy and recent onset of disease. Consistent with the example provided by cardiac allografts, early studies have begun to explore the utility of peripheral blood cell profiling as a means of identifying the presence and severity of cardiomyopathy (48–50). Recent data from other fields suggest that miRNAs may have particular potential for peripheral blood sampling that reflects myocardial adaptations to disease (51–53).

**Therapeutic Applications of Genomics and Transcriptional Profiling**

**Predicting efficacy.** Pharmacogenetics is the study of how the clinical responses to drugs vary with the patient’s genes. In HF therapeutics, studies elucidating how beta-1-adrenergic receptor polymorphisms affect the clinical actions of beta-blockers represent a particularly example of pharmacogenetic impact. After demonstrating that the presence of an arginine (rather than glycine) at position 389 confers a greater myocardial response to beta-adrenergic stimulation and blockade (54), Liggett et al. (55) examined clinical outcomes in a randomized trial of bucindolol in chronic HF. When patients were stratified on the basis of their haplotype, patients with the glycine haplotype exhibited no evidence of improved survival with bucindolol, but there was a significant survival benefit with active treatment among patients with the Arginine 389 haplotype. Thus, this beta-1 receptor polymorphism clearly modulates the response to a proven therapeutic agent in HF patients. Analogous demonstrations of treatment-modifying SNPs have been reported for the angiotensin-converting enzyme polymorphisms (56), aldosterone synthase CYP11B2 polymorphisms (57), the alpha-2C-adrenergic receptor polymorphisms (58), and GRK5 polymorphisms (59) in patients with HF. Collectively, these demonstrations of how genotype affects existing HF therapeutics foreshadow how clinicians will be able to rationally prioritize these and other interventions and move toward personalized therapeutics.

**Predicting toxicity.** Genotype may also modify the odds of an adverse response to certain therapeutic interventions in patients with HF. For example, Terra et al. (60) demonstrated that the presence or absence of 1 or both of 2 particular SNPs had a profound effect on the likelihood of patients requiring up-titration of other HF medications during treatment with a long-acting beta-blocker. Likewise, Wojnowski et al. (61) identified several SNPs that altered the likelihood of developing acute and chronic cardiotoxicity after treatment with the anthracycline doxorubicin. These emerging data concerning the effects of SNPs upon treatment efficacy and toxicity highlight the need to identify and validate other functionally important genetic variations that affect the development and treatment of HF while devising strategies to make these data available to clinicians in a manner that allows informed individualization of therapeutic strategies. To enhance identification of therapeutically relevant associations, the prospective collection of genetic material from all stages of clinical trials should be encouraged and the resulting analysis shared with regulatory agencies and academic collaborators in a consortium setting. With respect to utilization of these data, it is likely that advances in bioinformatics and increasingly prevalent electronic medical record systems will be essential enabling technologies. As in all genetic epidemiology studies, it is imperative that findings derived from a single small cohort be validated by replication in larger independent cohorts of patients.

**Conclusions and Future Prospects**

The clinical purpose of characterizing patients’ genotype and expression profiles is to enhance disease identification, prevention, and treatment while minimizing adverse effects. From this perspective, identification of disease-causing mutations has the potential to identify particularly high-risk genotypes in patients with overt disease and identify or exclude genetic risk in their family members. At the same time, accelerating identification of more common disease- and treatment-modifying variants will continue to provide new pathophysiological insights and also allow more rational prioritization and application of therapeutic options. Applications of clinical genomics will be enabled by the ongoing decreases in the cost of genome sequencing from a simple blood sample, routine inclusion of genotyping in clinical trials, ongoing GWAS efforts, and well-conceived protections against discriminatory use of genetic information by payers.

Complementing relatively static genomic characterizations, transcriptional profiling of animal models and human specimens will continue to provide new insights into disease mechanisms. Attention to characterizing splice variants and miRNAs, especially if combined with myocardial sampling relatively early during disease progression, will provide new clues to myocardial pathobiology while also identifying novel biomarkers for diagnostic and prognostic purposes. In keeping with the paradigm established in cardiac allografts, these same approaches are also likely to identify circulating biomarkers that can guide therapeutic decisions. Although the more extensive characterizations provided by transcriptional profiling and other new biomarkers will challenge clinicians’ abilities to utilize complex and diverse information, advances in information technology will help manage this complexity to permit greater individualization of prevention and treatment strategies in the setting of a growing HF epidemic.

Reprint requests and correspondence: Dr. Kenneth B. Margulies, University of Pennsylvania School of Medicine, BRB II/III, 421 Curie Boulevard, Philadelphia, Pennsylvania 19104. E-mail: ken.margulies@uphs.upenn.edu.
REFERENCES


23. Boyd SD. Everything you wanted to know about small RNA but were afraid to ask. Lab Invest 2008;88:569–78.


Key Words: genomics • transcription • heart failure.