Epigenomic and transcriptomic approaches in the post-genomic era: path to novel targets for diagnosis and therapy of the ischaemic heart? Position Paper of the European Society of Cardiology Working Group on Cellular Biology of the Heart

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Published Version
doi:10.1093/cvr/cvx070

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:33029766

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://
Epigenomic and transcriptomic approaches in the post-genomic era: path to novel targets for diagnosis and therapy of the ischaemic heart? Position Paper of the European Society of Cardiology Working Group on Cellular Biology of the Heart


Abstract

Despite advances in myocardial reperfusion therapies, acute myocardial ischaemia/reperfusion injury and consequent ischaemic heart failure represent the number one cause of morbidity and mortality in industrialized societies. Although different therapeutic interventions have been shown beneficial in preclinical settings, an effective cardioprotective or regenerative therapy has yet to be successfully introduced in the clinical arena. Given the complex
pathophysiology of the ischaemic heart, large scale, unbiased, global approaches capable of identifying multiple branches of the signalling networks activated in the ischaemic/reperfused heart might be more successful in the search for novel diagnostic or therapeutic targets. High-throughput techniques allow high-resolution, genome-wide investigation of genetic variants, epigenetic modifications, and associated gene expression profiles. Platforms such as proteomics and metabolomics (not described here in detail) also offer simultaneous readouts of hundreds of proteins and metabolites. Isolated omics analyses usually provide Big Data requiring large data storage, advanced computational resources and complex bioinformatics tools. The possibility of integrating different omics approaches gives new hope to better understand the molecular circuitry activated by myocardial ischaemia, putting it in the context of the human ‘diseasome’. Since modifications of cardiac gene expression have been consistently linked to pathophysiology of the ischaemic heart, the integration of epigenomic and transcriptomic data seems a promising approach to identify crucial disease networks. Thus, the scope of this Position Paper will be to highlight potentials and limitations of these approaches, and to provide recommendations to optimize the search for novel diagnostic or therapeutic targets for acute ischaemia/reperfusion injury and ischaemic heart failure in the post-genomic era.

1. Introduction

Myocardial ischaemia and reperfusion and its consequence heart failure are the leading causes of death and disability in industrialized societies. Despite intensive research in the last three decades, there is still no cardioprotective drug on the market that can alleviate myocardial ischaemia and reperfusion injury thereby reducing infarct size. Likewise, cardiac regeneration has the therapeutic potential to prevent the development of post-ischaemic heart failure. The lack of successful translation of promising therapeutic strategies to the clinical arena includes several factors such as the use of hypothesis-driven, biased target discovery and neglecting the interfering effects of common cardiovascular risk factors and their routine medications with cardioprotective pathways. Therefore, for successful development of cardioprotective therapies for acute myocardial ischaemia/reperfusion injury and ischaemic heart failure, new ways to discover valid targets must be used.

Since the completion of the Human Genome Project (HGP) a tremendous effort has been undertaken through Genetics, Genomics, and Functional Genomics studies to determine how this information might be linked to a healthy or unhealthy cardiac phenotype, particularly in myocardial ischaemia/reperfusion and the failing ischaemic heart. Despite great expectations, so far the clinical translation of basic discoveries in the areas of genetics has produced disappointing results (previously addressed by a Scientific Statement from the American Heart Association, and not discussed here). Several additional mechanisms related to epigenomic, transcriptomic, proteomic, and metabolomic regulation might be crucial to determine the pathological phenotype of myocardial ischaemia/reperfusion injury and ischaemic heart failure (for a glossary of terms, please see Table 1). The combination of omics techniques might be extremely powerful and effective to improve our understanding of the molecular mechanisms of myocardial ischaemia/reperfusion, facilitate the development of ‘tailored’ care, and in turn ameliorate the outcome of these patients. For example, a combined proteomics and metabolomics analysis of murine hearts revealed a reduction in succinate after preconditioning. Succinate was subsequently implicated in reperfusion injury through mitochondrial ROS generation. In addition, mass spectrometry is the method of choice for the analysis of post-translational modifications involved in epigenomic regulation, including histone modifications and the identification of histone-associated proteins. Although neither proteomics nor metabolomics can currently analyse the entire proteome or metabolome of mammalian tissues, with the rapid advances in mass spectrometry instrumentation the coverage is becoming increasingly more comprehensive.

---

**Table 1** Glossary of terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genetics</td>
<td>Study of specific, individual genes, and their role in inheritance.</td>
</tr>
<tr>
<td>Genomics</td>
<td>Study of genes, their functions, and biological effects.</td>
</tr>
<tr>
<td>Functional genomics</td>
<td>Study of changes in gene products (transcripts, proteins, metabolites) and how these changes mediate normal and abnormal biological function.</td>
</tr>
<tr>
<td>Comparative genomics</td>
<td>The study to compare the genes in one organism with those of another.</td>
</tr>
<tr>
<td>Epigenetics</td>
<td>The study of processes that lead to heritable changes in gene expression without changes in the DNA sequence.</td>
</tr>
<tr>
<td>Epigenomics</td>
<td>Study of epigenetic modifications across the genome.</td>
</tr>
<tr>
<td>Epitranscriptomics</td>
<td>Study of post-transcriptional RNA modifications not involving a change in the ribonucleotide sequence.</td>
</tr>
<tr>
<td>Gene expression</td>
<td>The study of mechanisms translating information encoded by a gene into RNA structures or proteins.</td>
</tr>
<tr>
<td>Transcriptomics</td>
<td>The study of the complete set of RNA transcripts that are produced by the genome under specific conditions.</td>
</tr>
<tr>
<td>Omics</td>
<td>Suffix added to the names of many fields to denote studies undertaken on a large or genome-wide scale.</td>
</tr>
<tr>
<td>Proteomics</td>
<td>The study of the complete set of translated proteins within a biological sample under particular circumstances.</td>
</tr>
<tr>
<td>Metabolomics</td>
<td>Quantitative study of all intermediary metabolites in a given biological state.</td>
</tr>
<tr>
<td>Phenome</td>
<td>The whole set of phenotypic entities in an organism.</td>
</tr>
<tr>
<td>Diseasome</td>
<td>All molecular or phenotype-based relationships between diseases observed in an organism.</td>
</tr>
</tbody>
</table>
Many of the stress signalling pathways activated by myocardial ischaemia and reperfusion culminate in the nucleus, leading to alterations in cardiac gene expression. A wide range of molecular players are involved in the regulation of gene expression at multiple levels, including transcriptional regulatory proteins binding specific DNA motifs in the control regions of the genes, and epigenetic modifications, inducing changes of gene expression without alterations in the gene sequence. A variety and continuously increasing number of epigenetic mechanisms modulating gene expression has been described (reviewed in), including ATP-dependent chromatin remodelling, ncRNA-based mechanisms, covalent histone modifications, and DNA methylation. Moreover, myocardial ischaemia/reperfusion injury and heart failure develop due to a combination of the individual genetic make-up and interaction with cardiovascular risk factors and comorbidities accumulating during ageing, that also affect cardiac gene expression profile and phenotype.

How variations in global cardiac gene expression profile can directly contribute to disease progression, the elucidation of the molecular signalling pathways underlying gene expression and the ensuing phenotype is extremely complex. Microarray or sequencing-based techniques now allow high-resolution, genome-wide investigation of both epigenomic and transcriptomic landscapes, and the integration of results from these approaches, with subsequent robust bioinformatics platforms to analyse ncRNA expression profile revealed a dynamic regulation of myocardial ncRNAs in failing human heart and remodelling with mechanical circulatory support. Although there is an increasing interest in protecting the

## 2. Epigenomics and transcriptomics in the ischaemic heart

### 2.1 Epigenomics

Although a large number of studies have investigated the role of epigenetic mechanisms in cardiac development, postnatal maturation and adverse cardiac remodelling in response to pressure overload, only few studies have specifically investigated genome-wide epigenetic modifications in cardiomyocytes induced by myocardial ischaemia or in ischaemic heart failure. It has been recently demonstrated, using a combination of chromatin immunoprecipitation followed by sequencing (ChiPSeq) and microarray transcriptome profiling, that ischaemic preconditioning promotes epigenetic repression of the Mtor gene and induction of autophagy by enhancing dimethylation on lysine 9 on histone H3. In cardiac samples from end-stage heart failure patients, differences in DNA methylation and distinct patterns of histone H3 methylation (lysine 26) could also be identified. Methylations differences in several genes related to heart disease or with unknown function were also identified in samples from patients with idiopathic dilated cardiomyopathy. As small molecule inhibitors of histone deacetylases are available and approved for clinical use, there is much interest in repurposing these compounds for cardiovascular indications in hopes of translating this biology into the clinical domain. Although such compounds are already used in clinical trials or daily practice in cancer patients, their potential cardiac effects have been poorly described. Future studies will be needed to address their effects in the ischaemic heart.

### 2.2 Transcriptomics

Transcriptomics studies in animal models of myocardial ischaemia/reperfusion and ischaemic heart failure have clearly demonstrated the importance of altered expression of genes involved in cardiac metabolism, cell growth and survival, inflammation, cytoskeleton structure, and extracellular matrix remodelling (recently reviewed in). However, only a few studies performed global gene expression profiling in ischaemic heart failure. Some studies identified gene expression profiles differentiating between ischaemic and non-ischaemic cardiomyopathy and others compared transcriptomics of failing and non-failing human tissue samples. However, Kuner et al. revealed poor separation of human cardiomyopathies of ischaemic and non-ischaemic aetiologies by gene expression profile. ncRNA expression profile revealed a dynamic regulation of myocardial ncRNAs in failing human heart and remodelling with mechanical circulatory support. Although there is an increasing interest in protecting the
ischaemic heart by ischaemic conditioning and related cardioprotective mechanisms, only a small fraction of conditioning studies assessed global gene expression profile of the heart in response to ischaemic conditioning and most of these studies are rather descriptive. The first studies in the literature using the pioneer DNA-chips showed that cardioprotection by ischaemic preconditioning affected gene expression profile of the heart in rabbits and rats more than was previously thought. These results have been confirmed by several papers showing that ischaemic conditioning triggers a cardioprotective gene expression profile in the heart at the transcript level. Moreover, changes in the expression of the post-transcriptional regulators of gene expression, i.e. ncRNAs, have recently been shown by ischaemic conditioning in rat hearts. However, little is known about whether remote ischaemic conditioning may also affect the global gene expression profile of the heart. Nevertheless, the study of Simkhovich et al. showed that regional ischaemia led to changes in the gene expression profile in the remote non-ischaemic area of the heart. So far one paper has confirmed that remote ischaemic preconditioning leads to alteration of gene expression profile in the heart in mice. In line with this assumption, extracellular vesicles, potential carriers of microRNAs (miR), have been shown to mediate the cardioprotective effect of remote ischaemic preconditioning in rat hearts.

Identification of global gene expression profile of the heart due to ischaemic conditioning stimuli may help to identify key cardioprotective pathways to be targeted, e.g. by gene therapy (see for a recent review). As an example, by systematic comparisons looking at the direction of miR expression changes due to ischaemia/reperfusion with or without conditioning stimuli, potential cardioprotective miR targets termed ‘protectomiRs’ such as miR-125b, miR-139-3p, miR-320, miR-532-3p, and miR-188 have been identified. Similarly to myocardial ischaemia/reperfusion, changes in the global cardiac gene expression profile including miRNAs has been shown during the development of heart failure in different preclinical animal models as well as in human endomyocardial biopsies of heart failure patients. Moreover, transcriptic studies have identified some miRNA clusters (miR15, miR17/92, miR 302/367) and individual miRs (including miR-99/Let-7c, miR-100/Let-7a, miR-199a, and miR-590), that control cardiomyocyte proliferation and may exert potent cardiac regeneration of the adult myocardium (reviewed in reference).

2.3 Influence of risk factors, comorbidities, and their pharmacological treatment

Several risk factors, ageing, comorbidities, and their drug treatments have been shown to modulate gene expression profile of the normal and diseased heart thereby modulating cellular signalling by either direct modification of transcription factors, post-transcriptional, and/or post-translational modifications or epigenetic mechanisms. Most of the studies determined specific epigenetic patterns or transcriptomics in the whole heart tissue or in circulating blood cells, but little is known on the global gene expression or epigenetic regulation in the different cell types of cardiac tissues. One study examined the gene expression profile of patients with new-onset heart failure due to idiopathic dilated cardiomyopathy, and revealed important sex-specific differences, i.e. 35 overexpressed and 16 down-regulated transcripts in men vs. women.

Long-term effects of hyperlipidemia and obesity on cardiac gene expression have been studied in some experimental models. Expression of numerous genes was altered in rat hearts on cholesterol-enriched chow for 2 months, in high-fat chow fed mice, and in dogs after 9 weeks of high fat diet. In dogs fed a high-fat chow for 9–24 weeks, a time-dependent decrease in expression of several genes implicated in obesity-related cardiac pathologies, such as hypertrophy and fibrosis was found. Expression of several miRs is altered in the ventricles of rats fed a cholesterol-enriched chow. In this study, one of the most robust down-regulation was found with respect to miR-25, which might be responsible for the increased oxidative- and nitrosative stress observed in the heart in this model. However, a more complex bioinformatic analysis of the miRNA omics data is missing from this study. Unfortunately no relevant human data are available on the effect of cardiovascular risk factors on the global gene expression pattern of the normal or ischaemic heart. Although a study showed that the gene expression pattern of the human atrium is significantly altered in obesity, so far data are unavailable from ventricular samples that would be relevant for myocardial ischaemia/reperfusion injury or ischaemic heart failure. For more details on cardiac transcriptomic changes in response to risk factors, see recent reviews.

Data mining analysis has suggested an important role for epigenetic factors in the pathogenesis of type 2 diabetes. However, most epigenetic studies addressing the mechanisms and effects of diabetes were performed in blood cells, endothelial cells, smooth muscle cells, or pancreatic cells, while limited information is currently available in cardiac myocytes. Significant divergence in myocardial gene expression profile has been observed in the hearts of the diabetic Otsuka Long-Evans Tokushima fatty rats when compared to the non-diabetic controls and in Zucker diabetic fatty compared to control lean rats. In type 1 diabetic rats, transcriptomic changes were responsible for altered myocardial metabolic substrate utilization that may account for the exaggerated progression of post-infarction remodelling. Moreover, alterations in miR expression profile in the heart of type 1 diabetic rats coupled with bioinformatic target prediction suggested that the affected miRs make a significant contribution to the myocardial transcriptomic profile involved in hypertrophy and fibrosis.

Currently used drugs to treat comorbidities of myocardial ischaemia may modify the response of the heart to ischaemia/reperfusion injury via several mechanisms including changes in cardiac gene expression pattern. For example, excessive use of nitrate inducing nitrate tolerance also modifies gene expression profile of the heart and leads to disruption of cardioprotective pathways. Statins are potent cholesterol-lowering drugs that are widely used in clinical practice for primary and secondary prevention of coronary heart disease. Although some rodent studies showed alterations in cardiac gene expression profile due to treatment by different statins and despite the abundance of clinical studies investigating the effects of statins, little is known regarding the effects of this class of drugs or other widely used anti-hyperlipidemic drugs on cardiac gene expression.

2.4 Role of mitochondria?

Although most DNA is nuclear, mitochondria also contain a small amount of mitochondrial DNA (mtDNA). Mitochondria have been shown to ‘interfere’ with epigenetic modifications of nuclear DNA by several mechanisms. Moreover, mtDNA mutations have been recently associated with coronary heart disease. The possibility that mtDNA might directly undergo epigenetic modifications (in particular DNA methylation since mammalian mtDNA lacks histones) has been questioned and extensively investigated for several years. Although mtDNA contains a lower number of CpG dinucleotides compared to nuclear DNA and lacks CpG islands, it has been recently proposed that epigenetic modifications of cytosines in mtDNA might be much more frequent than previously believed. Mitochondria-specific oligonucleotide microarrays have been developed to analyse the expression levels
of genes important for mitochondrial function encoded by nuclear or mtDNA in different in vitro and in vivo mouse models and in humans, mostly utilized in hypoxia, toxicogenomics, and aging. Future studies will be needed to determine whether and how changes in cardiac mtDNA methylation or mitochondrial transcriptomics patterns are associated with myocardial ischaemia and its comorbidities.

3. Prospects and pitfalls of epigenomic and transcriptomic profiling in the ischaemic heart

3.1 Prospects of unbiased omics approaches

Epigenomic and transcriptomic profiling provide quantitative information about epigenetic changes, gene expression and splicing variants, and allow to study variations in the heart throughout disease progression and in response to environmental changes or treatments. As global approaches, they provide large quantities of data that can be used for unbiased assessment of pathophysiological processes without a priori assumption (Figure 2). Newer techniques such as ribosome sequencing (Ribo-seq), which report only those transcripts that are being actively translated, might allow a more nuanced interrogation of the transcriptional landscape. Correlation and causality associations obtained from bioinformatic analyses of these data might represent a great opportunity for researchers to gain important novel insights into the mechanisms underlying myocardial ischaemia and reperfusion. In contrast to other approaches targeting a putative, single molecular target, this strategy might be more helpful to identify multiple key targets determining cardiac dysfunction in response to ischaemia and reperfusion. Once identified crucial networks in preclinical studies and/or in patient’s samples, more focused analyses of specific pathways or targets might offer the potential for rapid diagnostics and prognostics, drug safety testing and patient selection, and stratification for clinical trials. In addition to single-centre experiences, international consortia have been developed to share datasets, resources and protocols, providing a large amount of freely available information that is likely to exponentially increase in the near future (see for example reference90).

Figure 2 Schematic representation of unbiased experimental omics approach (A) or hypothesis-driven traditional approach (B).
Thus, authors of this Position Paper believe that compared to hypothesis-driven investigations, a non-hypothesis-driven research approach through omics methodologies have a strong potential to speed up the discovery process and give broader insight into signalling hubs activated by myocardial ischaemia/reperfusion injury together with its co-morbidities and their routine medications.

3.2 Technical and experimental pitfalls

Despite these promising aspects, there are multiple pitfalls when dealing with epigenomic and transcriptomic analysis in the ischaemic heart. Major technical problems are the heterogeneity within biological specimens, the quality of RNA and DNA (integrity, DNA contamination), and the insufficient availability of clinical samples from patients with myocardial ischaemia and reperfusion of ischaemic heart failure. Indeed, access to human cardiac tissues is often limited and, when possible, it usually involves samples from end-stage heart failure explanted hearts. These hearts have usually undergone multiple treatments and present relevant comorbidities that might directly or indirectly affect gene expression and epigenetic patterns. Cardiac tissue samples often contaminated by blood may also alter their omics readouts. Moreover, fresh healthy hearts are not easily available making the comparison with appropriate control groups often extremely challenging.

Given the relative inaccessibility of cardiac tissue, alternative sources have been considered, including human cell-based samples such as whole blood or white blood cells, or cell-free DNA samples from the plasma. The rationale supporting the use of these samples is that cardiac cells may release epigenetically modified DNA or that blood cells may have undergone changes similar to cardiomyocytes. Although future investigations focused on this topic might be extremely interesting and valuable, at the present time the mechanisms underlying this potential ‘mirrorimg’ are not clearly demonstrated, and therefore studies focusing on cardiac tissue and cells are recommended.

Preclinical studies using animal models have been extremely helpful and valuable as they allow the investigation of the progress of mechanisms of adverse remodelling towards end-stage heart failure, and comparisons of genomic expression across species often identifies patterns of evolutionary conserved transcriptional profiles among mammalian species. However, species-specific expression patterns have also been described and even in the presence of high-degree homology between species, there might be small but significant differences in the activated intracellular pathways. Moreover, the role of genetic background even in the same species is still not clearly established yet. A vast array of heart failure-related genetically modified mouse models has been produced since the techniques became available. Cardiac-specific inducible and conditional gene targeting in mice has greatly enlarged our understanding of the function of individual genes or even gene variants. However, the potential importance of adaptations in the heart to the altered expression of a single gene may modify the global gene expression pattern of the heart.

Whatever the experimental model used, since cell-specific gene expression and epigenetic profiles have been consistently reported, it would be ideal to focus and profile specific cell types, rather than the whole tissue, in which the underlying changes might be small and diluted by the presence of non-affected cells. In this regard, it would be crucial to define the target cell type, and describe the methods used to isolate, characterize, culture and eventually treat the cells. Studies have been performed with purified populations of cardiomyocytes obtained using different techniques or isolating cardiomyocyte nuclei from the heart by magnetic or fluorescent-assisted sorting. The isolation procedure of a specific cell-type and cell culturing might also cause changes in RNA expression and epigenetic profiles, thus affecting the analysis. Single-cell techniques to perform transcriptomic analysis are already available and have been used to study gene expression during cardiogenesis in vitro, even if they are not yet in widespread use. Such technologies may also provide insights into the cellular heterogeneity of tissues as recently elegantly shown for cardiac fibroblasts. Single cell approaches also might be especially useful when myocardial ischaemia is treated by cell-based therapies, often after genetic manipulation of the transplanted cells (see for review reference).

From a temporal standpoint, most studies so far have analysed static expression data, such as snapshots in different disease states. This approach might fail to detect disease-causing events that are transient. By using this kind of analysis, it is difficult to determine the changes directly or indirectly initiated by myocardial ischaemia. In addition, static expression data limit the use of reverse engineering approaches to infer a network of gene interactions. In contrast, a dynamic approach describing changes in expression profiles at different time points during the course of a disease would allow network-based analysis and provide novel biological hypotheses. Yet, only very few temporal data sets exist, and the majority of these datasets are limited to blood samples or specimens of patients of different age and conditions. Longitudinal studies describing heart disease in individuals appear rather impossible, however, longitudinal studies using animal models appear promising to infer a network of gene interactions, elucidate underlying mechanisms of heart diseases, and to develop novel therapies.

The number and diversity of transcriptomic and epigenomic profiling techniques is so wide that it can be challenging to select the best one. Although a complete review of all possible methodologies is beyond the scope of this Position Paper, major advantages and disadvantages of the most commonly used techniques for transcriptomics and epigenomic profiling are reported in Table 2. Different results may be obtained by using different technologies, and even with the same technology, variability might start in the wet lab, using different myocardial portions, sample storage conditions, processing methods, and timing. For all these reasons, for high-throughput sequencing could be considered running in duplicates of biological samples. In addition, when RNA is prepared and subsets are enriched or depleted (e.g. poly-A or ribo-minus selection), preparation of libraries, sequencing methods (e.g. strand specific sequencing), as well as the depth of the sequencing and bioinformatics analysis can yield strikingly different results depending on, e.g. the mapping to the reference genome and annotations. Such technical details may become major obstacles when comparing independent studies, which not only use different experimental conditions and platforms, but also varying bioinformatics approaches making potential meta-analyses difficult or useless. Re-analysing raw data may help to overcome at least the variability introduced by the bioinformatics analysis, but may not avoid heterogeneity introduced by the wet lab. Thus, in order to reduce variability, robust and streamlined methods, gold-standard instrumentations, and only widely accepted protocols should be employed to obtain optimal starting samples. Taken together, although generation of omics data could be better harmonized by more standard techniques, novel bioinformatics models may still allow analysis of disease-specific molecular networks from rather heterogeneous data sets as well.

Authors of this Position Paper believe that cardiac tissue samples and isolation of specific cell types should always be preferentially used to derive information on epigenomic and transcriptomic targets in the ischaemic heart whenever possible. Although we recognize the relative inaccessibility of human cardiac tissues in the daily clinical practice, cardiac tissue samples should be preferentially used to derive information on epigenomic and transcriptomic targets related to myocardial ischaemia.
Table 2 Advantages and disadvantages of major genome-wide transcriptomic and epigenomic profiling techniques

<table>
<thead>
<tr>
<th>Method</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TRANSCRIPTOMIC PROFILING</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>qRT–PCR microarrays</td>
<td>• Widely available</td>
<td>• High costs for genome-wide analysis</td>
</tr>
<tr>
<td></td>
<td>• Low costs for selection of genes</td>
<td>• Normalization sensitive to method/choice of reference genes</td>
</tr>
<tr>
<td></td>
<td>• High precision and sensitivity</td>
<td>• Not suitable for discovery of novel gene transcripts</td>
</tr>
<tr>
<td></td>
<td>• Increasingly multiplexed</td>
<td></td>
</tr>
<tr>
<td>cDNA microarrays</td>
<td>• Relatively low costs</td>
<td>• Prior probe selection</td>
</tr>
<tr>
<td></td>
<td>• Established method</td>
<td>• Probe redundancy and annotation</td>
</tr>
<tr>
<td></td>
<td>• Useful for large transcriptomes</td>
<td>• Sensitivity due to hybridization (background hybridization, probe saturation)</td>
</tr>
<tr>
<td></td>
<td>• Easy</td>
<td>• Complex computational analysis</td>
</tr>
<tr>
<td></td>
<td>• Fast</td>
<td>• Difficult absolute quantification</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not suitable for discovery of novel gene transcripts</td>
</tr>
<tr>
<td>SAGE</td>
<td>• Lack of hybridization problems</td>
<td>• Required validation by qRT–PCR</td>
</tr>
<tr>
<td></td>
<td>• Suitable for discovery of novel gene transcripts</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• No need to know sequence in advance</td>
<td></td>
</tr>
<tr>
<td>MPSS</td>
<td>• No need to know sequence in advance</td>
<td>• Tag specificity (multiple transcripts might have the same tag)</td>
</tr>
<tr>
<td></td>
<td>• Suitable for discovery of novel gene transcripts</td>
<td>• High costs and long times</td>
</tr>
<tr>
<td></td>
<td>• Digital data easy to store in databases and compare</td>
<td>• Required recognition site for restriction enzyme</td>
</tr>
<tr>
<td></td>
<td>• Larger library size and longer signatures compared to SAGE</td>
<td>• Restriction enzyme function</td>
</tr>
<tr>
<td></td>
<td>• Detection of lower expression levels transcripts</td>
<td>• Tag annotation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Still relatively new</td>
</tr>
<tr>
<td>NanoString</td>
<td>• Versatile</td>
<td>• Sequence selection</td>
</tr>
<tr>
<td></td>
<td>• Useful for clinical applications</td>
<td>• Not suitable for discovery of novel gene transcripts</td>
</tr>
<tr>
<td></td>
<td>• Library and processing not needed</td>
<td>• Hybridization problems</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Medium-throughput</td>
</tr>
<tr>
<td>RNA-Seq</td>
<td>• Analysis of any RNA type</td>
<td>• Not yet suitable for low amount of samples</td>
</tr>
<tr>
<td></td>
<td>• No need to know sequence in advance</td>
<td>• Complex computational analysis</td>
</tr>
<tr>
<td></td>
<td>• Reliable on low abundance transcripts</td>
<td>• RNA fragmentation</td>
</tr>
<tr>
<td></td>
<td>• Suitable for discovery of novel gene transcripts</td>
<td>• Storage of a large amount of data</td>
</tr>
<tr>
<td></td>
<td>• Identification of genetic variants and isoforms</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Broader sensitivity to highly expressed transcripts</td>
<td></td>
</tr>
<tr>
<td><strong>EPIGENOMIC PROFILING</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA methylation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHARM</td>
<td>• Low costs</td>
<td>• Moderate resolution</td>
</tr>
<tr>
<td></td>
<td>• Irrespective of proximity to genes or CpG islands</td>
<td>• Limited to regions next to restriction enzyme sites</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Does not detect 5hmC</td>
</tr>
<tr>
<td>MBDCapSeq</td>
<td>• Relatively low costs</td>
<td>• Relatively low resolution</td>
</tr>
<tr>
<td></td>
<td>• MBD proteins can discriminate between 5mC and 5hmC</td>
<td>• No absolute quantification of methylation levels</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Dependent on MBD binding sensitivity and specificity</td>
</tr>
<tr>
<td>MeDIPSeq</td>
<td>• Relatively low costs</td>
<td>• Does not identify single 5mC sites</td>
</tr>
<tr>
<td></td>
<td>• Detection in regions with higher and lower CpG density</td>
<td>• Sensitive to CpG density and copy numbers</td>
</tr>
<tr>
<td></td>
<td>• Antibodies can also identify 5hmC (hMeDIP-Seq)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Feasible with small amounts of DNA</td>
<td>• No absolute quantification of methylation levels</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Does not detect single 5mC sites</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Dependent on antibody sensitivity and specificity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Resolution of 100–300bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Sensitive to CpG density and copy numbers</td>
</tr>
</tbody>
</table>

Continued
target discovery is the crucial initial step in biomarker and drug development. Undoubtedly, a critical aspect in the search of novel targets in the post-genomic era is probably no longer data generation, but rather data analysis and interpretation. The search, identification, validation, and interpretation of novel disease-associated targets will require innovations in high-throughput technologies, biostatistics and bioinformatics, and will necessitate interdisciplinary expertise and teamwork of clinicians, biochemists, biologists, and bioinformaticians. In this context, final data interpretation usually develops in a multidisciplinary, shared environment. Accordingly, bioinformatics has undoubtedly gained wider prominence over the last decade in myocardial ischaemia/reperfusion and other research areas. Bioinformatics areas of interest include the development of new algorithms and statistics to assess relationships among members of large data sets, the ability to analyse and interpret different types of sequences, domains, and structures, and the capacity to develop or implement tools enabling efficient access and management of different types of information. A large number of methods, including pattern recognition algorithms, network analysis, and other artificial intelligence-based algorithms have been applied to bioinformatics research and will not be discussed here in detail.12,101–103 In general, all approaches are characterized by specific advantages or limitations for target identification and, in order to apply a ‘tailored’ bioinformatics approach to the research plan, it is usually necessary to accurately predetermine and focus on the elements to be investigated, on the interactions between the different elements, on the mathematical models to describe these interactions and on statistics to be used to analyse the data.10 Thus, combining biological concepts with computer tools or statistical methods it is possible to discover, select and prioritize targets.

### 3.3 Target identification: importance of bioinformatics

Target discovery is the crucial initial step in biomarker and drug development. Undoubtedly, a critical aspect in the search of novel targets in the post-genomic era is probably no longer data generation, but rather data analysis and interpretation. The search, identification, validation, and interpretation of novel disease-associated targets will require innovations in high-throughput technologies, biostatistics and bioinformatics, and will necessitate interdisciplinary expertise and teamwork of clinicians, biochemists, biologists, and bioinformaticians. In this context, final data interpretation usually develops in a multidisciplinary, shared environment. Accordingly, bioinformatics has undoubtedly gained wider prominence over the last decade in myocardial ischaemia/reperfusion and other research areas. Bioinformatics areas of interest include the development of new algorithms and statistics to assess relationships among members of large data sets, the ability to analyse and interpret different types of sequences, domains, and structures, and the capacity to develop or implement tools enabling efficient access and management of different types of information. A large number of methods, including pattern recognition algorithms, network analysis, and other artificial intelligence-based algorithms have been applied to bioinformatics research and will not be discussed here in detail.12,101–103 In general, all approaches are characterized by specific advantages or limitations for target identification and, in order to apply a ‘tailored’ bioinformatics approach to the research plan, it is usually necessary to accurately predetermine and focus on the elements to be investigated, on the interactions between the different elements, on the mathematical models to describe these interactions and on statistics to be used to analyse the data.10 Thus, combining biological concepts with computer tools or statistical methods it is possible to discover, select and prioritize targets. Importantly, ‘targets’ should be broadly considered molecular entities (such as genes, proteins, RNA, and their modifications) as well as biological phenomena (such as molecular functions and signalling pathways).

The complex features induced by myocardial ischaemia are difficult to ‘capture’ using single omic data modality, addressing only one mechanism of disease. There is now significant interest in the use of multiple high-throughput omic data for the identification of novel molecular subtypes of a disease. Integration of expression data with epigenetic datasets is a promising strategy to improve accuracy of the inferred networks, enable the

### Table 2 Continued

<table>
<thead>
<tr>
<th>Method</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| Methylation microarrays | • Relatively low costs  
• High sensitivity     | • Medium coverage (not every CpG site)  
• To date only for human samples                                      |
| WGBS                  | • Methylation state of almost every CpG site  
• Can distinguish between 5mC and 5hmC if bisulfite sequencing is performed after chemical oxidation of 5hmC to 5FU (oxBS) | • DNA degradation after bisulfite treatment  
• Does not discriminate between 5mC and 5hmC  
• Limited to regions next to restriction enzyme sites  
• Lower coverage at distant and intergenic regulatory sites |
| RRBS                  | • Relatively low costs  
• High sensitivity and coverage | • High costs  
• DNA degradation after bisulfite treatment  
• Does not discriminate between 5mC and 5hmC  
• Limited to regions next to restriction enzyme sites  
• Lower coverage at distant and intergenic regulatory sites |
| TabSeq                | • Can distinguish between 5mC and 5hmC  
• High sensitivity and coverage | • High costs  
• DNA degradation after bisulfite treatment  
• High sequencing depth required to detect 5hmC  
• Conversion dependent on efficiency of Tet enzymes |
| Histone modifications |                                                                           | • Labor and skill intensive  
• Can be limited by operator experience  
• Array-specific resolution  
• Requires higher amounts of ChIP DNA  
• Lower detection limit  
• High signal saturation |
| ChIPchip              | • Relatively low costs | • Labor and skill intensive  
• Can be limited by operator experience  
• High costs |
| ChIPSeq               | • Single nucleotide resolution  
• Lower amounts of ChIP DNA | |
discovery of new insights and create a global perspective from which novel diagnostic, prognostic or therapeutic targets in the ischaemic heart could be discerned. Despite the abundance of genome-wide transcriptomic and epigenomic datasets, there is a surprising shortage of statistical tools aimed at their specific integrative analysis. Such integrative tools might be crucial for the discovery of novel subnetworks (disease module) and/or relevant molecular pathways. Combination of these methods usually entails specific difficulties, and therefore establishing the concepts and fundamentals of data integration is clearly needed. Traditional database techniques are not focused on the problems that arise in the search and finding of data in large, not pre-organized repositories, for linking what is found to other data, or for reusing and repurposing the data without significant effort. Moreover, different omics data have different distributions and variation patterns that configure data heterogeneity. In addition to the specific variation patterns in each layer of information, inter-layer regulatory co-variations can be identified, increasing complexity. While multiple programs and web services are available for the analysis of single omics data, the majority of these tools has a command-line interface, performs only one specific task and typically requires file conversions. To better explore and integrate multiple omics data modalities (for example epigenetic, transcriptional, proteomic, and metabolomic modifications), novel and more sophisticated bioinformatics methods, that can seamlessly account for the networked nature of the subcellular environment, and that are simple to use by the general research community would be needed.

Authors of this Position Paper believe that in order to fully realize the scientific value of the vast amount of omics data, it is essential to fill the gap between data generation and investigators’ ability to retrieve and interpret the data. Most network and clustering methods proposed to identify molecular targets were developed for single omic data types and may not be appropriate when more than one omic data type is collected in the same study subjects. To better explore and integrate multiple omics data modalities, novel and sophisticated bioinformatics methods would be needed, allowing automated, systematic, and unbiased identification of targets. This approach will need more financing and more collaboration between cardiovascular and bioinformatic research groups, ideally hiring biomedical scientists with strong bioinformatic expertise, a species hardly found due to gaps in biomedical education.

### 3.4 Data availability, storage, and sharing

Collaborative research is usually the rule in this field, and it requires implementation of an integrative framework with standardization of data formats and possible access to the data from multiple web-based user positions. Distinct data warehousing approaches have been used, and which is more useful is yet to be determined. Shared warehousing possibilities and software components might be extremely helpful for successful collaborative work. Moreover, since several heterogeneous databases are currently available and more are under steady development, a standardized interface might represent an important target of research for developers. In general, more advanced and powerful computing will be needed in the future for large-scale omics analysis in terms of hardware and development of novel software. Additional issues regarding security and appropriate control of access to data, privacy protection, protection by backup, ethical issues, real-time availability and eventually data mobility will also likely be a matter of research and discussion in this field. These next-generation post-genomic challenges will need to be addressed in the future by computer developers, researchers, and the general population of patients who might benefit from the potential applications of these studies.

Various datasets derived from different detection systems are in need of harmonization as most profiling studies are performed using dissimilar platforms, requiring re-analysis of existing raw datasets, and after harmonization of datasets. Since availability of data is crucial for validation of the bioinformatics approach, authors of this Position Paper recommend that publishers and researchers commit to make freely available all relevant information, raw data sets, and their analysis. Details of protocols are essential allowing investigators to compare or pool data sets. As soon as technologies become state-of-the-art, the development of standard operating procedures (SOPs) might be helpful to provide guides for the research community. Moreover, acquired data should be stored and managed in open-source, web-based platforms for data-intensive biomedical research.

### 3.5 Target validation

After *in silico* identification of key gene/RNA/protein targets by different bioinformatic tools analysing epigenomic and transcriptomic data, experimental validation of these predicted molecular targets at the protein level is essential. Using *in vitro* cellular systems (genome editing, transfections, etc.) for target validation has become a relatively easy approach. However, target validation in *in vivo* in the target disease models preferably in the presence of relevant co-morbidities, obviously cannot be spared. An obvious pitfall of such *in vivo* validation is that the translational value of routinely used in *in vivo* disease models of myocardial ischaemia/reperfusion and heart failure is not sufficient (e.g. lack of comorbidities and co-medications), and the cost of such target validation is relatively high. Nevertheless, an increasing number of credible data on cellular signalling in health and disease in multiplex preclinical disease models and from human samples will lead to creation of an accurate disease module, a subnetwork of the full human interactome, integrating knowledge from DNA, epigenome, transcriptome, proteome, and metabolome. Since the network analysis approach of Big Data is somewhat tolerant to input of data with heterogenous quality, this will allow us to integrate data in the ischaemic heart in the context of the ‘dissosphere’, allowing us to explore its relationship to common cardiac comorbidities and other diseases.

Authors of this Position Paper believe that predicted targets identified by bioinformatics approaches from omics studies should be always validated in order to avoid generation of purely descriptive data and

<table>
<thead>
<tr>
<th>Table 3 Major recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Unbiased bioinformatic analysis of full epigenomic and transcriptomic profiles of the ischaemic heart (preferably including proteomic and metabolomic data) may lead to identification of novel molecular targets;</td>
</tr>
<tr>
<td>- Cardiac tissue samples should be used for omics assays and cell-specific analysis should be attempted;</td>
</tr>
<tr>
<td>- User-friendly bioinformatic tools should be developed for target prediction from large omics data;</td>
</tr>
<tr>
<td>- Omics data should be stored in open-access databases to enable their analysis by the global scientific community;</td>
</tr>
<tr>
<td>- Experimental validation of predicted targets should be performed in relevant models of the ischaemic heart.</td>
</tr>
</tbody>
</table>
putative predictions. This Working Group recommends that a heart-specific targeted approach in relevant disease models should always be used for validation whenever possible.

4. Conclusions and recommendations

Omnics technologies have the potential to revolutionize our understanding of complex diseases at the molecular network level. In order to optimize the search for novel diagnostic or therapeutic targets for the ischaemic heart in the post-genomic era, this Working Group provides the following recommendations (summarized in Table 3):

1. Compared to hypothesis-driven approaches based on a priori target selection, global, unbiased assessment of epigenomic and transcriptomic profiles (preferably together with proteomic and metabolomic profiles) combined with bioinformatics and biostatistics evaluations have the potential to identify multiple novel molecular targets in the ischaemic heart.

2. Whenever possible, cardiac samples should be used as sources of biological material and cell-specific analysis should be attempted.

3. To better explore and integrate multiple omics data modalities, complex but simple to use bioinformatics tools should be developed, allowing automated, systematic, and unbiased identification of targets.

4. Publishers and researchers should always commit to make freely available all relevant data sets and their analyses. Acquired data should be stored and managed in open-source, web-based platforms for data-intensive biomedical research.

5. After the unbiased, non-hypothesis-driven target identification, experimental validation should be performed in relevant experimental models of myocardial ischaemia and reperfusion and ischaemic heart failure to avoid generation of purely descriptive data with putative predictions.

Conflict of interest. P.F. is a founder and CEO of Pharmahungary Group, a group of R&D companies. A.L.B. is the co-founder of DZZOM that commercializes network-based bioinformatics tools. R.S. received grants from Sanofi and Astra Zeneca, lecture honoraria from Amgen, Recordati, Sanofi. S.D. is advisor for Miragen. F.P. received a grant from Bayer, and lecture honoraria from Novartis and MSD. L.D.W. is co-founder and stockholder of Mirabilis Therapeutics BV. T.T. is founder of Cardior Pharmaceuticals GmbH and has filed and licensed patents in the field of noncoding RNAs. None declared for other authors.

Funding

This work was supported by the Italian Ministry of Health (GR-2009-1596220), by the Italian Ministry of University (RBFR124FEN), and by the Federico II-STARR-grant (junior principal investigator) to C.P.; National Institute of Health (NIH). USA 1R01HL118455-01A1 and Centers of Excellence of Genomic Science (CEGS) 1505HG004233 to A.L.B.; European Foundation for the Study of Diabetes (EFSD) New Horizons Collaborative Research Initiative from European Association for the Study of Diabetes (EASD) and European Cooperation in Science and Technology (COST EUROS) to R.S. and P.F.; European Research Council Advanced Grant (CardioEpiGen, # 294609), Associazione Italiana per la Ricerca sul Cancro, the Italian Ministry of Health and the National Research Council, Italian Ministry of Research and Education, and the CARIPLO Foundation to G.C.; National Research, Development, and Innovation Office, (OTKA K. 109737, OTKA ANN 107803, NVKP 16-1-2016-0017) to P.F.; European Research Council (ERC) grant 311549 and a VIF award 918-156-47 from The Netherlands Organization for Scientific Research (NWO) to L.D.W.; ERC grant (Angiolic) and Leducq foundation grant (MIRVAD) to S.D.; the Netherlands Organization for Health Research and Development (ZonMW Veni 91612147) and Netherlands Heart Foundation (Dekker 2013T056) grants to L.V.L., and the Deutsche Forschungsgemeinschaft (DFG Research Unit FOR 2149) to F.B.E. Duke-National University Singapore Medical School, National Medical Research Council National Heart Centre Singapore Collaborative Centre Grant, British Heart Foundation (FS/10/039/28270), the Rosetrees Trust, and the National Institute for Health Research University London Hospitals Biomedical Research Centre, UK to D.J.H., ERC Grant (Longheart) and Leducq Fondation grant (MIRVAD) to T.T.

References


Epigenomics and transcriptomics in the ischaemic heart

735


25. Greco CM, Condorelli G. Epigenetic modifications and noncoding RNAs in cardiac

Trends Cardiovasc Med

19. Tunon J, Barbas C, Blanco-Colio L, Burillo E, Lorenzo O, Martin-Ventura JL, Mas S,

20. Haldar SM, McKinsey TA. BET-ting on chromatin-based therapeutics for heart fail-

22. Haberland M, Montgomery RL, Olson EN. The many roles of histone deacetylases

J Am Coll Cardiol

2015;70:751–762.

Circ Res


J Mol Cell Cardiol

2015;0:88–94.

Nature Genet


FEBS Lett


J Mol Cell Cardiol

2015;132:852–872.

Acta Physiol

2016;228:813–832.

Expert Opin Biol Ther


J Mol Cell Cardiol


Cardiovasc Med


Rev Endocr Metab Disord


Nat Cell Biol


Nat Commun


Nat Commun

2005;130:1326–1332.

Nat Commun


Nature


Nature

2010;467:2177–2187.

Am J Physiol Cell Physiol

2002;282:C63–C73.

Nat Med

2005;11:75–78.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg

2005;130:1326–1332.

J Thorac Cardiovasc Surg


J Mol Cell Cardiol

2011;124:2871–2880.

J Mol Cell Cardiol

2011;124:2871–2880.

J Mol Cell Cardiol

2011;124:2871–2880.

J Mol Cell Cardiol

2011;124:2871–2880.

J Mol Cell Cardiol

2011;124:2871–2880.

J Mol Cell Cardiol

2011;124:2871–2880.

J Mol Cell Cardiol

2011;124:2871–2880.

J Mol Cell Cardiol

2011;124:2871–2880.


