Toward a Holistic View of Transcriptional Regulation

Keith A. MacCannell, PhD; Ralph V. Shohet, MD

The first response to the flood of transcriptomic information obtained from microarray technology was reductionist, identifying individual transcripts that are regulated during an experimental perturbation. This approach continues to identify many previously unsuspected actors in physiological and pathophysiological processes. Coordinate regulation, illustrated by previously ubiquitous heat maps and self-organizing maps, provided some further summary evaluation and suggested associations between genes that had not been previously recognized. Additional connections between regulated transcripts were obtained by mapping them onto known functional pathways, which began to provide mechanistic understanding directly from transcriptional regulation. Over the past few years, a further evaluation of the network of interacting gene regulation has emerged that can, in at least some circumstances, provide a further level of refinement of our understanding of the molecular changes that contribute to disease.

One application of network analysis is the statistical evaluation of large amounts of genomic data to derive new insights into previously unsuspected aspects of molecular coordination. The study of Dewey et al in this issue of Circulation: Cardiovascular Genetics performs such an analysis on cardiac transcriptional data from a large number of mouse experiments. These investigators use both meta-analysis, the statistical combination of multiple studies, with network analysis, which in this case is applied to the relations of regulated transcripts.

Combining multiple studies is now possible because of the requirement that published microarray data be submitted in a single format into a publicly available database. Dewey et al refined a set of data from the Gene Expression Omnibus relating to “pathological” cardiac hypertrophy from all data that mentions the heart. The investigators elected to evaluate only murine data and to exclude models of physiological hypertrophy or transgenic manipulation. These decisions resulted in a dataset of 478 microarray experiments selected out of 2546. The number of genes evaluated was also limited to 12,620, presumably those present in all of the array platforms. The loss of transpecies data is unfortunate because findings that are preserved across species may be more likely to be true. This decision was based on preliminary analysis that suggested that array experiments “clustered by species.” However, it seems likely that authentic, functional networks of gene expression will be preserved between human and mouse, and those that are functional in humans are of the greatest interest. The absence of interspecies confirmation may increase the number of false-positive results, and, perhaps, reveal coordinated expression networks, unique to the rodent lineage, that are not shared with primates. Large quantities of data exist on human hypertrophy and heart failure and may, in the future, yield insights into our shared responses to hypertrophy. Nonetheless, the murine data evaluated in this study constitute more than 5 million determinations of mRNA abundance in the cardiac models examined. It is an impressive effort to have curated and normalized this huge amount of data into a tractable form for the subsequent analysis.

The investigators then applied a variety of sophisticated statistical analyses to create a map of common coordinated gene networks that are regulated and differentially expressed during pathological hypertrophy, heart failure, and fetal development. The mathematics used for network analysis is still in development and may seem occult to the cardiovascular scientists who will need to interpret it. There are many mathematical models that can be applied, and it is likely that different approaches will work optimally for different types of data and network structures. For example, the investigators have used a more rigorous permutation method in their analysis rather than assuming the independence of the statistical relation. The sophistication of the work analysis presented generates hypotheses, not demonstrations. Although genes that are coregulated are often affected by, or responding to, a common stress, without direct experimental manipulation of these pathways it is impossible to judge the validity of the statistical relation. The sophisticated statistical mathematics of this analysis is primarily an effort to identify potential coordination while limiting false-positive implications of coincidental coregulation. However, there are many other confounding factors that will inevitably produce inaccuracies in the identified networks.

The opinions expressed in this article are not necessarily those of the editors or of the American Heart Association.

Correspondence to Ralph V. Shohet, MD, Center for Cardiovascular Research, John A. Burns School of Medicine, University of Hawaii, Honolulu, HI.


Circ Cardiovasc Genet is available at http://circgenetics.ahajournals.org

DOI: 10.1161/CIRCGENETICS.111.959452
They include indirect regulation, in which the crucial function of the product of an unregulated gene may not be recognized because it is modified by regulation of an inhibitor or activator. The authors also recognize the potential importance of microRNA-mediated regulation, which is not included in the data they are analyzing. Finally, the important, and probably often predominant, influence of posttranscriptional and posttranslational events is also only indirectly revealed in transcriptional response, if at all.

Importantly, this study does produce predictions that are subject to experimental validation. For example, targets of the transcription factor, ZIC2, not previously suspected of a role in cardiac stress, were overrepresented in the fetal module recapitulated in hypertrophy and heart failure. Transcriptional evaluation of a mouse model, or even a cell culture system, in which ZIC2 activity is knocked down or overexpressed, should demonstrate a specific response in the gene network it coordinates, if the underlying hypothesis is correct.

What’s next? There have already been several efforts to integrate different types of data across networks. In an Einsteinian effort to achieve a grand synthesis, large scale information from genomic, proteomic, and even metabolomic network analysis is being integrated in networks of networks. Initial efforts have focused on yeast but are certain to approach more complicated organisms soon. There is also an effort to examine network evaluations across species to improve the reliability and generalizability of these results.

Ultimately, network analysis will prove useful to the extent that it provides predictions that are unexpected and that lead to definitive biological insights confirmed by experiment. The work in yeast is enticing, and one should expect that other systems will prove tractable. The study of Dewey et al is an impressive first step in applying these methods to the heart.

Disclosures
None.

References
Toward a Holistic View of Transcriptional Regulation
Keith A. MacCannell and Ralph V. Shohet

doi: 10.1161/CIRCGENETICS.111.959452
Circulation: Cardiovascular Genetics is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2011 American Heart Association, Inc. All rights reserved.
Print ISSN: 1942-325X. Online ISSN: 1942-3268

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/4/1/2

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Genetics can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Genetics is online at:
http://circgenetics.ahajournals.org//subscriptions/