Top Advances in Functional Genomics and Translational Biology for 2009

Patrick T. Ellinor, MD, PhD; Jennifer E. Van Eyk, PhD

During the past years vast strides have been made in population genetics and proteomics. Genome-wide association studies (GWAS) have become large, collaborative, multicenter efforts often involving in excess of 50,000 subjects with dramatically increased power compared with previous efforts. Rapid evolution in sequencing technology has made comprehensive sequencing of the exome feasible and enabled a unique approach toward the identification of rare disease-causing variants. This year, several key proteomic articles have been published on the discovery and verification of new protein biomarkers that mark a new stage in this field, specifically in our efforts to move these emerging markers toward clinical application.

The American Heart Association Functional Genomics and Translational Biology Interdisciplinary Council provides a forum for a multidisciplinary group of volunteers committed to making a substantial contribution to reduce the burden of heart disease and stroke through genetics, genomics, proteomics, and metabolomics research and translational science. On behalf of the Council on Functional Genomics and Translational Biology, we have selected 4 manuscripts published in 2009 that reflect these advances.

Genome Wide Association Study of Premature Myocardial Infarction Identifies 9 Genetic Loci


Principal Findings
Most patients who have a myocardial infarction (MI) are elderly; however, a minority of patients with a MI present at an early age have a higher heritability of MI and are thus more likely to have a genetic basis for their condition. Investigators from the Myocardial Infarction Genetics Consortium performed a multistage case-control GWAS in subjects with early-onset MI. In an initial stage, a GWAS for single-nucleotide polymorphisms (SNPs) and copy number variants was undertaken using 2967 cases and 3075 controls. The top hits from the initial analysis and from the previously published studies were analyzed in 3 successive replication stages consisting of an additional 12,800 cases and controls. A total of 9 loci were significantly associated with premature MI, in which 6 were previously reported and 3 were novel. None of the common or rare copy number variants were associated with premature MI.

Implications
This manuscript highlights some of the great strengths and inherent weaknesses of GWAS. As for the strengths, the investigators are to be commended for the large, multistage replication strategy used because they have convincingly identified many common genetic variants associated with premature MI. Second, as an unbiased and genome-wide analysis, GWAS offers the potential to discover new pathways for disease pathogenesis. The limitations of this manuscript are similar to other GWAS studies. Considerable future work will be necessary to elucidate the precise mechanisms by which genetic variants identified in this GWAS, and in other association studies, lead to disease. Furthermore, the majority of GWAS have been performed in individuals of European ancestry, and future work will be necessary to determine whether the results can be generalized to other races and ethnicities. Finally, despite the tremendous efforts of the investigators, the overall effect sizes remain small, with odds ratios ranging from 1.12 to 1.29. Even when combined, the top variants at each of the 9 loci account for <3% of variance in MI risk. Thus, despite the heritability of premature MI, a large part of the genetic basis of the disease remains unexplained.

Rapid Sequencing of the Human Exome


Principal Findings
As illustrated in the preceding manuscript, despite the ever increasing scale of genome-wide studies, the common variants identified have typically only accounted for a small percentage of the heritability of a condition. Therefore, other disease mechanisms including the role of rare genetic variants have continued to be explored. Although the costs for
sequencing the entire genome are falling quickly, genome-wide sequencing is likely to continue to remain too expensive for individual investigators for at least a few more years. Based on these constraints, a recent focus on sequencing all of the protein coding regions, or exomes, that comprise \( \approx 1\% \) of the human genome, has emerged.\(^2\)

Ng et al sequenced the exomes of 12 individuals. Of these, 8 were referent subjects, and 4 had Freeman-Sheldon syndrome, a rare, autosomal dominant craniosynostotic dysplasia due to mutations in \( MYH3 \). Starting with genomic DNA, they used 2 microarrays to enrich the exome in each individual, and then shotgun sequenced the enriched library. They obtained an average of 6.4 gigabases of sequence or \( \approx 50\)-fold coverage of the exomes for each individual. Use of referent subjects from the HapMap and Human Genome Structural Variation projects enabled a detailed comparison of the call rates of coding SNPs (cSNPs) between the current exome sequencing and traditional sequencing or genotyping techniques. Overall, exomic sequencing was highly sensitive at cSNP detection with concordance rates in excess of 99%. There were an average of 17,272 cSNPs identified in each individual, and as expected more cSNPs were identified in Africans than in non-Africans. Polymorphisms that disrupt splice-site junctions or alter protein coding are typically assumed to be pathological, and a striking number of such variants were identified; among all of the subjects studied, there were a total of 225 nonsynonymous SNPs, 102 splice-site disruptions, and 664 small insertions or deletions. Finally, in a proof of concept study, the investigators used exomic sequencing and filtering of common cSNPs to identify multiple mutations in \( MYH3 \) among the 4 subjects with Freeman-Sheldon syndrome.

**Implications**

In the near future, exomic sequencing will offer the potential to screen the entire coding region of the genome. Particularly promising are the applications of this technique for the identification of causative genes in diseases that seem to have a genetic basis, but in which the families are too small for traditional linkage analysis.\(^3\) Limitations at the moment include cost, the vast amount of genotypic data generated, the need for expertise in data processing, and the large sample sizes necessary to detect rare variants for more common diseases.

**Mass Spectrometry for Multiplexed Biomarker Detection**


**Principal Findings**

Biomarkers have a vital role in promoting innovation in diagnostics and treatments. The application of proteomic technology in the area of circulating biomarkers has been technically challenging with respect to discovery of new markers and their verification. Kuhn and coworkers\(^5\) use a mass spectrometry method called selective reaction monitoring (or multiple reaction monitoring) to obtain absolute quantification of 2 known biomarkers in serum. This method is based on quantification of signature peptides that are unique to a particular protein and can even be used to track selective peptides with disease-induced modifications (eg, phosphorylation or proteolytic fragments), single/multiple amino acid changes due to polymorphism or unique to specific protein isoforms. The use of selective reaction monitoring was coupled with immuno-affinity enrichment of the representative low abundant proteins including cardiac troponin I and interleukin 33. The latter was selected to represent a candidate biomarker for which an ELISA does not exist. Selective reaction monitoring completely removes the need for antibodies or as in this case, the need for 2 high affinity specific antibodies while allowing one to obtain quantitative data that is comparable with the data (with similar reproducibility and coefficient of variation) obtained by classic methods such as ELISA or radioimmunoassay.\(^6\) Thus, the selective reaction monitoring technique described in this manuscript provides a novel approach for the targeted discovery and verification of biomarkers in cardiovascular disease.

**Transforming Growth Factor Beta in Marfan Syndrome**

7. Matt P, Schoenhoff F, Habashi J, Holm T, Van Erp C, Locht D, Carlson OD, Griswold BF. Developing multiplexed assays for troponin I and interleukin 33. The latter was selected to

**Principal Findings**

Transforming growth factor beta (TGF-\(\beta\)) is a ubiquitous cytokine present in virtually all mammalian cells that controls cell differentiation and proliferation.\(^7\) Circulating TGF-\(\beta\) levels (bound and free [active]) were elevated in a mouse model, \( Fbn1^{C1039G/\#} \) that mimics many clinical features of individuals with Marfan syndrome, including the progressive dilation of the aorta. TGF-\(\beta\) levels were also found to be elevated in a large cohort of patients with Marfan syndrome obtained through the multicenter GenTAC (National Registry of Genetically Triggered Thoracic Aortic Aneurysms and Cardiovascular Conditions) consortium. Importantly, TGF-\(\beta\) levels are sensitive to therapeutic intervention and are lower in both the \( Fbn1^{C1039G/\#} \) mouse model and in patients treated with the AT1-antagonist losartan, which has been shown to reduce aortic root growth and dimensions in the mouse model of Marfan syndrome.
Implications
This manuscript provides an elegant example of the ability to directly translate the findings from an animal model of a disease rapidly into patients. As pointed out in the accompanying editorial, measuring circulating levels of TGF-β may serve as a prognostic and therapeutic marker that can ultimately be integrated into the clinical management of patients with Marfan syndrome.

Sources of Funding
This work was supported National Institute of Health grants HL092577 and DA027021 (to P.T.E.), NO1-HV28180 and 1U54RR023561 (to J.V.E).

Disclosures
None.

References

Key Words: aorta • diagnosis • genetics • signal transduction
Top Advances in Functional Genomics and Translational Biology for 2009
Patrick T. Ellinor and Jennifer E. Van Eyk

doi: 10.1161/CIRCGENETICS.110.937367
Circulation: Cardiovascular Genetics is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2010 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/3/1/106

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/