Clinically Relevant Functional Annotation of Genotype

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

Editorial

Genetic Testing and Prediction

Genetic testing was originally thought likely to generate deterministic predictions of many clinical outcomes in inherited disorders. Indeed, this premise was one of the original drivers of the precision medicine movement. However, the variable expression of final phenotypes even in the setting of large genetic effects has long reinforced the probabilistic nature of any genotype–phenotype relationship. Recent genome wide association studies have revealed large numbers of loci that contribute to the common heritable variation in many disease traits, and whole-genome sequencing has uncovered a wealth of rare variation, even within a single individual, on a scale previously unimagined.1

Article see p 8

Such a deluge of genomic variation vastly exceeds our ability to model the phenotypic consequences using current approaches whether in vitro or in vivo. Emerging strategies exploiting induced pluripotent stem cells show some promise for focused studies of specific biochemical or cellular phenotypes; however, the limits of modeling integrated disease pathophysiology in an incomplete humoral or cellular context are all too apparent.2 Similarly, although higher throughput in vivo biology is feasible in some model organisms, such experimentation is challenged by concerns about the extent of the generalizability of the experimental findings to humans. Perhaps more importantly, in the era of comprehensive genetic and functional genomic datasets, our current in vivo phenotypic repertoire remains embarrassingly restricted in overall scope, dimensionality, and dynamic range. Indeed, one potential explanation of the so-called missing heritability evident in modern human genetics is the confounding that results from imprecise historical phenotypes with large measurement errors, multiple phenocopies, and variable relevance to the underlying biology.3 Furthermore, the silos in which most modern disease phenotypes have been refined and the aggregating influences of randomized controlled clinical trials have led to dramatic reductions in the effective resolution of many diseases. For example, it is hard to think that some phenotypes still remain in routine epidemiological use when the minute-to-minute variation in an individual subject is an order of magnitude greater than any population effect size ever reported (a high noise-to-signal ratio).

New Functional Assays

One approach that has proven powerful in many genetic studies is to improve the precision of the phenotypic anchor, by focusing on an underlying trait or endophenotype that is more sensitive and specific and is presumed as a result of its relationships in the biological pathway to be more closely reflective of the underlying pathogenesis.4,5 In Circulation: Cardiovascular Genetics of this month, Niemann et al6 have attempted to use a proposed endophenotype to annotate a series of variants in the alpha galactosidase A gene, the causal gene for Fabry disease.7

Fabry disease is an X-linked lysosomal storage disorder that leads to the accumulation of alpha galactosidase substrates including globotriaosylceramides (Gb3) in the lysosomes of multiple tissues.8 The tissues most commonly affected in this condition include the skin, the kidney, the heart, and the nervous system. However, there seems to be substantial variation in the manifestations of disease, even within a single kindred. In light of the availability of disease-modifying therapies, specifically enzyme replacement therapy, there is increasing pressure to be able to identify functionally significant variation in the alpha galactosidase gene rather than wait for documentation of histological involvement of target organs.9 The situation is further complicated by the increasing recognition (although with variable rigor) of clinically significant involvement of female heterozygotes, presumably through partially dominant mechanisms or tissue specific X-inactivation.10 Although the simple plasma or urine levels of Gb3 have been shown not to vary between affected and unaffected carriers of bona fide mutations, a degradation product of Gb3, specifically lyso-Gb3 has been hypothesized as a potential discriminating between those with disease and simple mutation carriers. Niemann et al6 measured plasma lyso-Gb3 levels and compared them with peripheral leukocyte alpha galactosidase activity and a series of clinical organ-specific events previously used to define classical or atypical forms of Fabry disease in a discovery set of 72 patients and a validation set of 52 patients. The authors were able to demonstrate that a simple cutoff level of lyso-Gb3 offered apparently remarkable discrimination between classic and atypical forms of Fabry disease. Skeptics may note that this candidate endophenotype is rather a biomarker equivalent of traditional clinical tools, while the circular logic of a current clinical gold standard suggests that the patients with low lyso-Gb3 Fabry are likely phenocopies. The burden of proof for a disease entity surely must include a discrete outcome, far less a discrete therapeutic response, and it is not yet clear that those individuals labeled

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

From the Cardiovascular Division, Brigham and Women’s Hospital and Harvard Medical School Boston, MA (C.A.M.); and Division of Preventive Medicine, Boston University and The Framingham Heart Study, Boston, MA (R.S.V.).

Correspondence to Calum A. MacRae, MB, ChB, PhD, Thorn 1127, 75 Francis Street, Boston, MA 02115. E-mail cmacrace@partners.org

(Circ Cardiovasc Genet. 2014;7:2-3.)

© 2014 American Heart Association, Inc.

Circ Cardiovasc Genet is available at http://circgenetics.ahajournals.org
DOI: 10.1161/CIRCGENETICS.114.000506

2
as having atypical Fabry disease exhibit outcomes discrete from normal controls. Although these data certainly suggest that lyso-Gb3 may ultimately have some use in the diagnosis of Fabry disease, the absence of any data on unaffected family members or on unselected populations preclude such conclusions.

Essentially every biomarker in history has fallen victim to a dramatic loss of specificity in the context of a reduced pretest probability. One need only reflect on the evolution of cardiac troponin T from a highly specific marker of myocardial necrosis to an archetypal gestalt for general malaise. How then will medical science ever be able to define the biological functions downstream of a given set of genetic variants with the precision necessary for meaningful clinical decision making?

**Framework for Functional Annotation at Scale**

Together these related problems argue for the organized development of a novel repertoire of in vivo phenotypes designed to be representative of the underlying cellular or humoral biology in a fully native context. The ideal phenotypic assays will not be constrained by existing disease silos but rather will offer integrated read-outs of fundamental cellular processes, with a suitable linear dynamic range, ease of multiplexing and cost effectiveness to allow the development of comprehensive panels that could be applied at a population scale. Importantly, both for analytic validity as well as for efficient translation to model organisms or to cellular assays, the optimum phenotypes will include a paired perturbation. The longitudinal collection of basal and perturbed data on multiple pathways in parallel from massive human cohorts and the combination with extant traditional phenotypes and hard outcome endpoints would offer the opportunity to redefine diseases along several quantitative dimensions. Given the scope of the challenges to be overcome in undertaking such an effort, it seems self evident that such data would have to be collected as part of a routine health evaluation across entire populations. Only large datasets of this nature will supply the granularity of phenotype necessary to begin to elucidate the clinical impact of the degree of genetic variation apparent in the average human and to realize the visions of individualized medicine.

Interestingly, such an organized redefinition of the landscape of human phenotyping would be perfectly aligned with many other paradigmatic changes currently underway in healthcare, including the implementation of a common electronic health record, the migration of disease detection and management from the hospital to the home and the community, and the advent of big data strategies across the biomedical arena. The need to map novel phenotypic data types to current clinical traits implies that this effort is likely to begin in existing population research cohorts; however, it will be vital to design strategies for efficient translation and scaling into any such schema if robust functional annotation of complex genomes is ever to be feasible.

**Sources of Funding**

Dr MacRae is supported by the NIH grant HG006500 and the Leducq Foundation.

**Disclosures**

None.

**References**


**Key Words:** Editorials
Clinically Relevant Functional Annotation of Genotype
Calum A. MacRae and Ramachandran S. Vasan

_Circ Cardiovasc Genet_. 2014;7:2-3
doi: 10.1161/CIRCGENETICS.114.000506
_Circulation: Cardiovascular Genetics_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2014 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/7/1/2

**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/