The study of the interaction between genes and environment dates back through the work of R.A. Fisher and Lancelot Hogben in the early 20th century, to that of Charles Darwin and Alfred Russel Wallace in the mid-19th. Darwin and Wallace’s focus was, of course, evolution by natural selection rather than human disease. However, their fundamental insight was that different members of the same species, faced with a change in environment, respond in different ways and that these differences in response areheritable. In the context of human disease, we think of gene–environment interactions in terms of people with different genotypes at a particular locus responding to an environmental stimulus, such as exposure to tobacco smoke, in different ways.

More broadly, there is abundant evidence that practically all human disease can be seen as being due to the effect of multiple genetic variants, and in the interaction of those genetic variants with each other, and with the environment. For example, although we think of injuries as purely environmental in nature, in fact trauma is also strongly genetically determined. The Y chromosome is a major genetic risk factor for trauma at all ages after 12 months, and beyond that, genetic variants contributing to personality traits, such as impulsivity, also serve as risk factors. Similarly, infectious disease represents an interaction between environmental exposure (contact with a pathogen) and genetic predisposition. Host genetic factors are important determinants of whether exposure to a pathogen passes unnoticed by the individual or causes severe morbidity or even mortality.

Conversely, it has become clear that apparently single gene disorders are nothing of the sort. Readers of Circulation: Cardiovascular Genetics will be familiar with the extreme variability of Mendelian disorders, such as hypertrophic cardiomyopathy, even between individuals with the same causative variant and even within the same family. This is, at least in part, because of the ameliorating or exacerbating effects of variants in genes other than the causative gene responsible for the condition (known as modifier genes), as well as environmental factors. Environmental influences on monogenic disease are mostly poorly understood, but can be striking. The impact of dietary modification on phenylketonuria and the effects of exposure to Burkholderia cepacia, complex organisms in cystic fibrosis are just 2 examples among many.

So, we know that genetic variation is important in causing human disease. We know that environmental factors are also important. And we know that the two interact. This means that a full understanding of the basis of any disorder will never be possible, until we understand the genetic (and epigenetic) and the environmental factors, which give rise to that disorder, as well as the way that they interact with one another. Currently, however, we are a long way from this goal, even in a field as important and well-studied as cardiovascular disease. A great deal is known about cardiac environmental risk factors. Considerable effort has been expended, particularly in the past decade, on identifying genetic loci which modify disease risk. But little is known about the intersection between the two.

Part of the reason for this is the formidable difficulty of studying such interactions. A major tool for the study of the genetic basis of common disease is the genome-wide association study (GWAS). A GWAS involves searching for associations between single-nucleotide polymorphisms and a phenotype of interest. A large number of single-nucleotide polymorphisms, typically hundreds of thousands, distributed across the genome, are tested in a cohort of thousands of individuals. The National Human Genome Research Institute - European Bioinformatics Institute Catalog of published GWAS lists ≈2500 such studies published since the first (in 2005), identifying ≈25000 single-nucleotide polymorphism trait associations. This approach has the benefit of being hypothesis-free, and unexpected associations between genetic variants and disease states discovered by GWAS have led to new biological insights.

However, reproducibility of GWAS results has been problematic, and it is demanding and expensive to conduct such studies; large numbers of well-phenotyped patients need to be recruited, consented, and genotyped. To adapt the technique to the study of gene–environment interactions adds a further layer of difficulty. The variants identified in GWAS studies are only sometimes located within genes or close enough to a particular gene that they might be surmised to be relevant to its function. This means that it is often difficult to determine the underlying biological mechanism behind the effect of a particular variant, even if the effect is relatively large and has been replicated in more than one study. Moreover, to date, most of the single-nucleotide polymorphisms which have been associated with a human trait to date, have only a modest impact on the phenotype being studied, whether that is a discrete outcome such as stroke or a continuous measure such as blood pressure.
pressure. Studying the interaction between a variant of modest effect and an environmental exposure adds a layer of difficulty and increases the required sample sizes.

The CHARGE Consortium (Cohorts for Heart and Aging Research in Genomic Epidemiology) is a major international collaboration, the aim of which is to facilitate GWAS meta-analyses and the replication of GWAS results. In this issue of Circulation: Cardiovascular Genetics, Rao et al report the formation, structure, and administration of, and methodology used by, the Gene-Lifestyle Interactions Working Group. This Working Group works with the resources of CHARGE to study the interactions of genetic variation and environmental factors. The phenotypes being studied in this first phase of the project are blood pressure and lipids. The environmental factors are smoking, alcohol consumption, education (as a surrogate for socioeconomic status), physical activity, a set of psychosocial attributes, and sleep duration.

The scale of this undertaking is enormous. In a field which was once highly competitive, no fewer than 124 cohorts, including 610,475 subjects from 5 ancestry groups, have been drawn together into a coordinated whole. The approach involves agreement on a standard study design, which is then implemented in each separate cohort. The resulting data are uploaded in a standard format to a central server, and meta-analysis is conducted across the cohorts. The investigators leverage the power of their huge sample size by conducting analyses using different models and with different structures. For example, the now standard approach of doing genome-wide analyses in a discovery cohort, followed by targeted analysis in a replication cohort is used, with 149,684 individuals in stage 1 and 490,791 individuals in stage 2. However, combined analyses using all 124 cohorts are also performed; using both approaches allows the maximum possible discovery to be made with the resources available.

Why have the investigators gone to the considerable lengths that they have? Why does it matter to the field that they have done so, and what kind of results can we expect and hope for from this study? The GWAS Catalog lists just 22 studies of gene–environment interaction, all published after 2010 and only 4 of which are directly relevant to cardiovascular disease. Already, 4 projects from the Working Group have completed all analyses and are being taken toward publication; effectively, they are already set to double the world literature in this field. Five other projects are well underway and more are being planned. We can expect to see a steady flow of papers, reporting new associations and, importantly, bridging the gene/environment gap, in a systematic way that has not been possible in the past. It is likely that novel biology will be revealed. Moreover, there is evident potential to expand the scope of the phenotypes and interactions that the group studies.

What about impacts for patient care? It seems unlikely we will ever be saying to patients “your particular set of genetic variants means that it’s fine for you to smoke”…or to refrain from exercise, or subsist on fast food. However, it is not unrealistic to expect that better understanding of the genetic basis of cardiovascular phenotypes, especially the role that genotypes play in modifying response to environmental exposures, will allow improved stratification of patients, both by risk and by likely response to interventions. It is no stretch to say that this article represents an important milestone on the path toward a far more complete understanding of the origins of cardiovascular disease and a better understanding of how to manage it.

Disclosures
None.

References
Genes, Environment, and the Heart: Putting the Pieces Together
Edwin P. Kirk

Circ Cardiovasc Genet. 2017;10:
doi: 10.1161/CIRCGENETICS.117.001764
Circulation: Cardiovascular Genetics is published by the American Heart Association, 7272 Greenville Avenue,
Dallas, TX 75231
Copyright © 2017 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/10/3/e001764

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/