

Genomewide Association Studies and Common Disease — Realizing Clinical Utility

Lars Fugger, M.D., Ph.D., Gil McVean, Ph.D., and John I. Bell, D.M.

Ever since Archibald Garrod showed, nearly 100 years ago, that human disease has a heritable component, scientists and clinicians have sought to translate fundamental genetic findings into a better understanding of disease mechanisms and effective treatments. In the past 10 years, since the human genome was sequenced and genomewide association studies have permitted the examination of variation in the genomes of thousands of people, we have learned much about which genes and biologic pathways are associated with both mendelian and complex diseases.

Studies of diseases such as type 2 diabetes, vascular diseases, autoimmune diseases, and cancers are now routinely conducted in thousands of people. The entire genomes of many thousands of patients have been sequenced, and vast resources are being poured into finding rare and common genetic variations that influence disease risk. Yet despite this investment, and with a few notable exceptions, the number of tangible benefits to patients remains small. In the face of skepticism about the value of genetic work for finding effective therapies, what can be done to accelerate progress in translating disease genetics into treatment?

The debate about the value of genetic work for addressing complex disease has two parts. The first is irrelevant; the second is critical. The first question concerns the usefulness for diagnosis and prognosis of genetic variants that are identified through association studies. The field of disease genetics grew out of the

study of mendelian traits, in which there is typically a one-to-one relationship between a mutation and a phenotype. The diagnostic and prognostic value of identifying such mutations is unequivocal. In contrast, the study of the genetics of complex diseases has largely, at least in recent years, entailed efforts to find variants whose prevalence differs only slightly between persons with a given disease and those without it. On its own, such variation has little or no predictive utility, and even if a moderate fraction of a trait's genetic variance (or heritability) can be accounted for by adding up the effects over many loci, the diagnostic utility is limited by the substantial environmental component in the causation of most complex traits.

This apparent limitation of complex-disease genetics is irrelevant to their translational value because genetic association studies were never designed to result in genetic tests. Rather, genes and biologic pathways implicated in disease through association may represent targets for interventions that are likely to have much greater effects than any naturally occurring variation. Indeed, efficient therapies may well be directed against gene products that exhibit no naturally occurring variation.

In a classic example, the genetic associations between some HLA loci and certain autoimmune disorders are some of the strongest associations that have been identified in complex disease, but they have very limited predictive value. Rather, they have implicated T-cell activation and adap-

tive immunity in causing many autoimmune diseases. Interventions that modulate aspects of T-cell function, including cytokine inhibitors, T-cell trafficking inhibitors, and therapies that attack the T cell itself, have proved useful in treating many of these disorders.¹ These therapies do not target the polymorphic locus of association, but rather are founded on knowledge of associated pathways.

Of course, the association of HLA with autoimmune disease has been recognized for many years. How can we begin to benefit from all the loci identified by genomewide association studies in recent years? Here's where the second part of the debate comes into play. The problem is that for the vast majority of variants associated with complex disease, we have no understanding of the biologic processes and pathways involved in disease causation. That's because we know very little about the function of the majority of genes in the genome, and even when we know something about function, it's often not possible (because of the structure of allelic association within the region) to identify the specific change that influences risk. Consequently, we have no way of knowing whether the causal variant increases or decreases gene expression, alters protein interactions or localization, increases or decreases activity, changes protein or RNA stability, interferes with local gene regulation, and so on. At this point, many observers would argue that fine mapping to a single causal variant is not essential. What matters is the func-

tion of the genomes that carry the variant. Nevertheless, without a single variant to examine (or at least a small number of variants in a highly localized genetic region), the range of assays that could be considered is effectively unlimited.

A recent example highlights the issue. Variation near the gene *TNFRSF1A* is associated with a modestly increased risk of multiple sclerosis (odds ratio, 1.12). Unlike most of the genome, the variant driving the association can be mapped exactly to a mutation that abolishes an essential splice site.² Functional work showed that this mutation results in exon skipping and the production of a novel, soluble form of the tumor necrosis factor (TNF) receptor, which binds to active TNF.² This finding is clearly not a full understanding of the disease pathway, but it mirrors observations from clinical trials of TNF antagonists, which can enhance the severity of, or even induce, multiple sclerosis. Without functional data, there was no way of knowing whether the associated variant was indeed causal. This study also showed that even if a genetic variant confers only modest risk of a disease and leads to a subtle alteration in biologic function, a drug mimicking its activity can have a much more powerful effect.

What are the implications of a study that has led to successful mapping of a disease-risk locus to a causative molecular phenotype? First, detail matters. It matters

that the genetic evidence generates a strong hypothesis about the molecular phenotype associated with risk and that the molecular phenotype can be assayed in cells or models that are relevant to the disease. This kind of dogged dissection of individual associations may take a long time to lead to new therapies, but it is essential.

The second implication is, perhaps, more heartening. Slightly perversely, one might say that the human body is full of chemicals that provide natural models for existing or potential drugs. Genetic variation within these pathways can provide clues as to the likely efficacy of such interventions. For example, recent work on rare genetic variants that influence high-density lipoprotein (HDL) cholesterol levels strongly suggests that HDL cholesterol is unlikely to be a useful therapeutic target in myocardial infarction.³ More generally, **multiple genomic regions have been identified through genome-wide-association-studies overlap with known drug targets.**⁴ For example, genetic variation in the interleukin 23–interleukin 17 axis is associated with susceptibility to psoriasis, which suggests that targeting this pathway might have therapeutic benefit. **Monoclonal antibodies binding or neutralizing these interleukins have shown efficacy against psoriasis, and several other compounds targeting them are in clinical development.**⁵

The ideal drug is one that mimics and enhances the effects of a genetic variant that has been identified as protective against a

given disease and that is not a risk factor for other diseases. Of course, such variant–drug combinations are not easy to come by, but careful dissection of existing association data could well identify important leads. Furthermore, such potential is precisely why it continues to be valuable to identify new genetic risk factors for complex disease through large-scale genetic studies. But to maximize the health benefits from genetics and genomics, it is imperative that a detailed translational phase be undertaken in parallel. Otherwise, the rich yield of disease loci identified by genome-wide association studies will be for naught.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

From the Weatherall Institute of Medicine (L.F.), the Nuffield Department of Clinical Neurosciences (L.F.), the Wellcome Trust Centre for Human Genetics (G.M.), and the Office of the Regius Professor of Medicine (J.I.B.), University of Oxford, Oxford, United Kingdom.

1. Rosenblum MD, Gratz IK, Paw JS, Abbas AK. Treating human autoimmunity: current practice and future prospects. *Sci Transl Med* 2012;4:125sr1.
2. Gregory AP, Dendrou CA, Attfield KE, et al. TNF receptor 1 genetic risk mirrors outcome of anti-TNF therapy in multiple sclerosis. *Nature* 2012;488:508-11.
3. Voight BF, Peloso GM, Orho-Melander M, et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet* 2012;380:572-80. [Erratum, *Lancet* 2012;380:564.]
4. Travers ME, McCarthy MI. Type 2 diabetes and obesity: genomics and the clinic. *Hum Genet* 2011;130:41-58.
5. Gudjonsson JE, Johnston A, Ellis CN. Novel systemic drugs under investigation for the treatment of psoriasis. *J Am Acad Dermatol* 2012;67:139-47.

DOI: 10.1056/NEJMp1212285

Copyright © 2012 Massachusetts Medical Society.

The Question

Chris Adrian, M.D.

“When people at parties ask what you do, just say you’re a pediatrician or make

something up. Telling them what you really do will just make them uncomfortable.” So

my attendings had advised me early in my pediatric oncology fellowship. Now I was at a party