EDITORIALS

Individual Genomes on the Horizon

Richard P. Lifton, M.D., Ph.D.

Physicians have long recognized that pinpointing specific causes of disease in individual patients enables therapies that are the most likely to confer benefit with the fewest adverse effects. We also recognize the potential for disease prevention through identification of specific risk factors and mitigation of their effects. For a century, we have known that many of these risk factors are genetic. In the past 20 years, the genomic revolution has translated this knowledge into a new understanding of disease: mutations that cause more than 2000 mendelian diseases have been identified, which has led to the rewriting of textbooks of pathophysiology of every organ system and the identification of rational targets for therapeutic intervention. Genes also play a major role in risk for virtually every common disease, affording the possibility of identifying persons who have a specific inherited predisposition.

The field has been driven by saltatory leaps in technology. The development of complete genetic maps of the human genome fueled the mapping and identification of genes underlying mendelian traits in nuclear families. Subsequently, the ability to inexpensively genotype hundreds of thousands of common sequence variants across the genome enabled the discovery of common variants contributing to common diseases in large cohorts of case patients and controls.

Building on the complete sequence of the human genome, spectacular reductions in the cost of DNA sequencing now point to a coming era of genomics based on identification of rare variants that confer disease risk in individual patients. When the sequencing of the first human genome was initiated, the cost to produce 1 million bases of sequence was $100,000. The development of new technologies that permit simultaneous sequencing of hundreds of millions of DNA templates has recently driven the cost to sequence 1 million bases to under $1.

This advance creates myriad opportunities for the use of DNA sequencing in gene discovery. For example, the discovery of the comprehensive set of somatic mutations in cancer[1] and suspected de novo mutations underlying diseases ranging from congenital malformations to autism become tractable goals. Similarly, common variants have explained only a small fraction of the inherited risk for most common diseases, findings that suggest a role for rare variants with relatively large effect,[2,3] which can be discovered by sequencing large cohorts. Finally, thousands of known and suspected mendelian traits that have thus far eluded understanding will most likely be solvable with the use of high-throughput sequencing.

Genome sequencing will also have a role in translating these discoveries into clinical diagnosis. Traditionally, the genetic diagnosis of a mendelian disorder relied on the establishment of a clinical diagnosis followed by the sequencing of previously implicated genes. Practical limitations of this approach include frequent diagnostic uncertainties, which thwart efforts to define a short list of genes for sequencing. Similar limitations arise for diseases in which mutations in many genes can cause the same disease. Sequencing these genes one by one is cumbersome and limits the number that can be efficiently examined. Supplanting this approach with routine sequencing of all the genes is consequently attractive and, more importantly, scalable. Although daunting challenges, such as distinguishing clinically significant mutations from nonconsequential variation, remain, the cost to

sequence all the genes in the genome with the use of new technology is already approaching the fee charged to sequence single genes in some diagnostic laboratories.

In this issue of the *Journal*, Lupski and colleagues report on their study that shows the power of this new technology.[4] They used whole-genome sequencing to make a specific diagnosis in a family in which four siblings were affected by Charcot–Marie–Tooth disease, a peripheral polyneuropathy. Mutations in 31 known genes and additional unidentified loci can produce Charcot–Marie–Tooth disease. The investigators produced nearly 90 billion base pairs of genomic sequence in one affected subject (sufficient to ensure that both alleles at nearly every base pair have been sampled repeatedly) and identified variations from the reference sequence. As expected, they found a large number of common and novel variants. When they examined genes known to be mutated in patients with Charcot–Marie–Tooth disease, they found two compelling mutations in *SH3TC2* (the SH3 domain and tetratricopeptide repeats 2 gene), which causes autosomal recessive Charcot–Marie–Tooth disease. They also found complete cosegregation of these mutations with disease status in the family, providing convincing evidence that these *SH3TC2* mutations are the cause of Charcot–Marie–Tooth disease in this family.

The sequence production for this project cost less than $50,000. More traditional approaches could have obtained the same answers; nonetheless, the study provides a striking proof of principle. Moreover, there is every reason to believe that the cost of sequencing will continue to plummet. Owing to innovation and intense competition, the cost of sequence production 2 years from now will almost certainly be at most one tenth of the current cost of using current technologies. Moreover, there are widespread efforts to advance new technologies to achieve further drastic drops in cost.[5]

In addition, large cost reductions can be achieved by shrinking the target for sequencing. Protein-encoding exons of the roughly 23,000 genes in humans constitute approximately 1% of the genome but harbor about 90% of all mutations with large effects. Efficient methods for whole-exome sequencing (that is, sequencing of all the exons in a genome) have recently been reported,[6,7] and their usefulness for both clinical diagnosis[7] and disease-gene identification[8] has been shown. Current costs for whole-exome sequencing are only about $4,000, and as long as expense remains a factor, a 90 to 95% reduction of that cost will be significant. Notably, this approach could have led to the same conclusion far less expensively in the current study.

It is increasingly clear that the cost is fast approaching a threshold at which DNA sequencing will become a routine part of the diagnostic armamentarium. This raises many critical questions. Who will benefit from comprehensive sequencing? When in a person's life should sequencing be done? How should we deal with the many variants of uncertain clinical significance? How should we interpret changes found outside of genes? How should we effectively communicate the results to patients in ways that will improve health without inducing neurosis? These questions have far-reaching implications for the education of health care professionals and patients as well as for health and social policy. Lupski and colleagues provide a glimpse of the future for which we need to prepare.

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

From the Departments of Genetics and Internal Medicine, Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, CT.

This article (10.1056/NEJMe1001090) was published on March 10, 2010, at NEJM.org.

1. Parsons DW, Jones S, Zhang X, et al. An integrated genomic analysis of human glioblastoma multiforme. Science 2008;321:1807-12.
2. Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. Science 2004;305:869-72.
3. Ji W, Foo J-N, O'Roak BJ, et al. Rare independent mutations in renal salt handling genes contribute to blood pressure variation. Nat Genet 2008;40:592-9.
4. Lupski JR, Reid JG, Gonzaga-Jauregui C, et al. Whole-genome sequencing in a patient with Charcot–Marie–Tooth neuropathy. N Engl J Med 2010;362:1181-91.
5. Drmanac R, Sparks AB, Callow MJ, et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. Science 2010;327:78-81.
6. Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes. Nature 2009;461:272-6.
7. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A 2009;106:19096-101.
8. Ng SB, Buckingham KJ, Lee C, et al. Exome sequencing identifies the cause of a mendelian disorder. Nat Genet 2010;42:30-5.