JAMA Insights | GENOMICS AND PRECISION HEALTH

Finding the Rare Pathogenic Variants in a Human Genome

James P. Evans, MD, PhD; Bradford C. Powell, MD, PhD; Jonathan S. Berg, MD, PhD

Decreases in the cost of DNA sequencing have enabled substantial progress in fields ranging from archaeology and evolution to basic biomedical science. Concomitantly, there have been calls for routine genome-scale sequencing of healthy individuals in hopes of discovering clinically important information. For example, discovery of a high risk of breast and ovarian cancer due to a *BRCA*1/2 mutation can enable aggressive surveillance or risk-reducing surgery.

How It Works

Massively parallel sequencing of the total complement of an individual's DNA (genome sequencing) has proven to be a powerful diagnostic approach for patients with disorders that have a primarily genetic etiology.[1] Genome-scale sequencing can be performed on DNA from white blood cells or buccal cells from saliva (Audio at time 5:43). In sequence analysis, each individual's genome contains millions of sites where his or her DNA differs from a reference sequence (typically a composite sequence derived by the international Genome Reference Consortium using data from many genomes). Clinical interpretation requires assessing whether any of these variants (eg, a nucleotide change altering an amino acid or a change that results in premature stop of protein translation) are associated with disease. Such determinations rely on control and population data sets, review of the medical literature, and the patient's phenotype (**Figure**).

Editorial page 1842

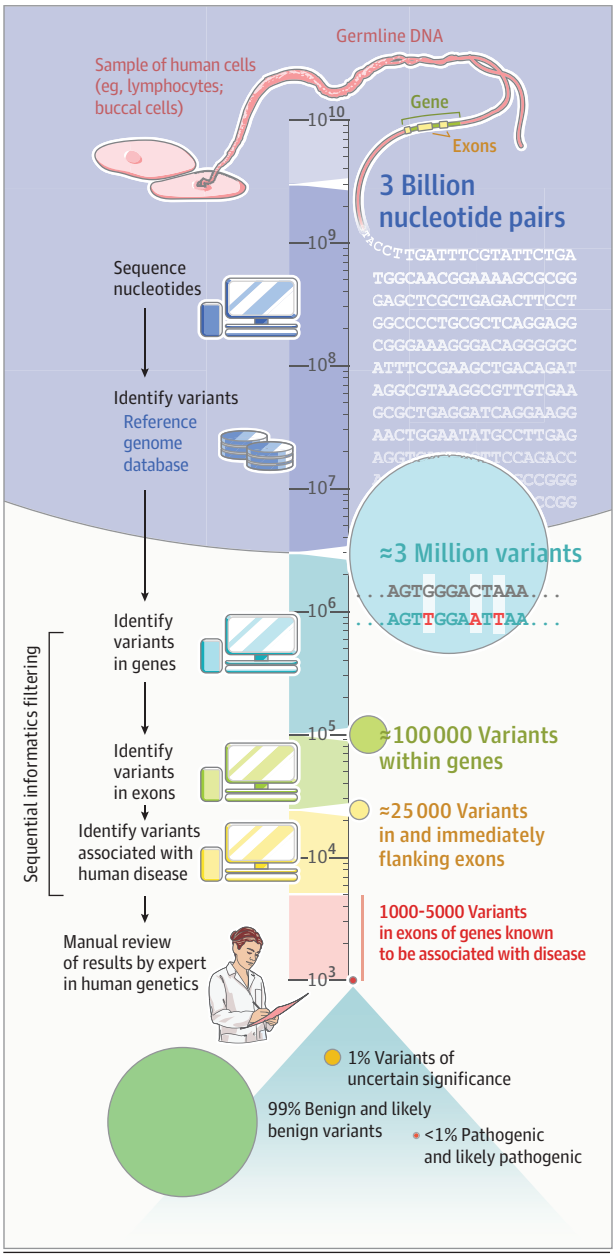
Author Audio Interview

CME Quiz at jamanetwork.com/learning

Despite promising ongoing efforts to catalog and curate genomic variants (eg, via the ClinGen effort[2]), current understanding of most genetic variation in humans is inadequate. Moreover, the interpretation and implications of genomic findings fundamentally differ in healthy individuals vs those with disease. For example, current estimates of genetic risks in a healthy individual carrying a disease-associated variant discovered by sequencing are likely inflated since risk estimates have been derived from studying patients presenting clinically, thus selecting for the most extreme cases. Such limitations render routine performance of genomic sequencing in healthy individuals currently inadvisable.

A genome sequence is not a single test. Rather, it represents millions of tests in which each nucleotide of a genome is queried and variants departing from a reference sequence are reported. Genomic information rarely results in a yes or no answer about disease state, but rather a probabilistic estimate of risk.[3] The chance of an individual manifesting a genetic disease depends on 2 factors: the certainty with which a variant is associated with disease and the penetrance of the condition (the likelihood an individual with a pathogenic variant will develop disease). Lack of knowledge of either factor impairs the predictive value of genetic information, resulting in false positives (variants interpreted as being likely pathogenic but that actually are benign) and overdiagnoses (pathogenic variants that do not lead to disease due to incomplete penetrance). Thus far, genetic risk prediction for common diseases

Figure. Informatic and Human Analysis Required for Finding Rare Pathogenic Variants in a Human Genome



Genetic variants are informatically filtered to remove those with very low likelihood of pathogenicity (eg, variants known to be benign or present at very high frequency in the general population). This informatic processing incorporates annotations of individual variants (eg, population allele frequencies, prior literature reports, computational predictions of functional effect) for use in manual analysis.

is typically no better than family history or other variables,[4] and genomic interpretive abilities are inadequate such that sequence information from the same individuals, analyzed by different experts, results in very different and conflicting medical interpretations.[5]

Important Care Considerations

The most salient harm from genome sequencing of healthy individuals is unnecessary medical intervention due to false-positive findings or overdiagnoses (Audio at time 9:12). Experience with the prostate-specific antigen demonstrated that broad screening with a single test possessing modest predictive value can lead to undesirable outcomes such as overdiagnosis leading to unnecessary procedures and attendant complications. That error should not be repeated by pursuing routine clinical genomic sequencing in healthy individuals until more is known about how to interpret it. Minimizing false-positive findings and overdiagnoses will require gene-specific thresholds for optimal return of results,[6] and the necessary knowledge base will come from sequencing many healthy individuals in a research setting with careful phenotyping and follow-up.

Although clinical genome-scale sequencing in healthy individuals is premature, the concept is worth pursuing. However, it will be important to think smaller. Instead of sequencing the whole genome of healthy individuals, necessarily yielding vast amounts of poorly understood data susceptible to misinterpretation, a better strategy may be to interrogate the small number of well-understood human genes that, when mutated, result in a high risk of readily preventable disease (Audio at time 11:20).

Approximately 1% to 2% of the US population harbors a genetic variant in such well-studied genes with the potential to benefit from established preventive modalities if the elevated risk were known. For example, more than half a million individuals in the United States harbor a pathogenic variant in a Lynch syndrome gene[7] conferring a high risk of colorectal cancer—an outcome readily preventable through annual colonoscopy. There exist a few dozen such genes that could be candidates for focused screening, including those responsible for Lynch syndrome, hereditary breast and ovarian cancer (*BRCA1*, *BRCA2*), and familial hypercholesterolemia. Finding those individuals at high risk for serious genetic disorders in time to implement prevention has been highly successful in newborn screening. Yet, while only approximately 0.1% of newborns screen positive for conditions that will benefit from early intervention to prevent or substantially ameliorate poor medical outcomes. An order of magnitude more adults are at risk for readily preventable disorders now discoverable through carefully targeted sequencing.[8]

Value of Whole-Genome Sequencing

Advocacy of broadly applying genome sequencing in healthy individuals is typically justified by its relatively low cost. However, while DNA sequencing is much less expensive than a decade ago, clinical exome sequencing still costs thousands of dollars. Even if sequencing becomes truly inexpensive, substantial medical costs result not from the simple act of performing a medical test, but from the interpretation of those tests and from the downstream actions that their results trigger.

Evidence Supporting Genome Sequencing in Healthy Individuals

Thus far, no systematic evidence supports genome sequencing in healthy individuals. Indeed, studies showing highly inconsistent interpretation argue against its broad application in healthy individuals.[5] A proper evidence base must be established before clinical implementation occurs, determining the true penetrance of genomic variants in diverse, unselected populations and ultimately demonstrating improved outcomes to those screened (Audio at time 13:30).

Bottom Line

Technology allows sequencing of the entire human genome, but for healthy people, it currently has no clear clinical benefit and is inconsistent with a fundamental clinical axiom—to refrain from seeking uninterpretable and misleading information in patients or healthy individuals. Future research should investigate the utility of more targeted genome sequencing for screening healthy people.

Audio, CME, and Additional Resources

Listen to the accompanying audio program for more information about genomic sequencing. In some instances, the answers to CME questions are in the audio and not the text of the article (Audio at time 2:53). Take the quiz at http://jamanetwork.com/learning/article-quiz/10.1001/jama.2017.0432. Learn more at jamagenetics.com.

ARTICLE INFORMATION

Author Affiliations: Department of Genetics, University of North Carolina at Chapel Hill (Evans, Powell, Berg); Department of Medicine, University of North Carolina at Chapel Hill (Evans).

Corresponding Author: James P. Evans, MD, PhD, Departments of Genetics and Medicine, CB #7264, Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7264 (jpevans@med.unc.edu).

Section Editor: W. Gregory Feero, MD, PhD, Associate Editor, *JAMA*.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Berg reports being a funded investigator on the National Human Genome Research Institute–funded GeneScreen project, which is investigating the potential use of genomic screening of a small panel of genes in generally healthy adults. No other disclosures were reported.

REFERENCES

1. Strande NT, Berg JS. Defining the clinical value of a genomic diagnosis in the era of next-generation sequencing. *Annu Rev Genomics Hum Genet*. 2016;17:303-332.

2. Rehm HL, Berg JS, Brooks LD, et al; ClinGen. ClinGen—the clinical genome resource. *N Engl J Med*. 2015;372(23):2235-2242.

3. Goldstein DB, Allen A, Keebler J, et al. Sequencing studies in human genetics: design and interpretation. *Nat Rev Genet*. 2013;14(7):460-470.

4. Khoury MJ, Janssens AC, Ransohoff DF. How can polygenic inheritance be used in population screening for common diseases? *Genet Med*. 2013;15(6):437-443.

5. Corpas M, Valdivia-Granda W, Torres N, et al. Crowdsourced direct-to-consumer genomic analysis of a family quartet. *BMC Genomics*. 2015;16:910-926.

6. Adams MC, Evans JP, Henderson GE, Berg JS. The promise and peril of genomic screening in the general population. *Genet Med*. 2016;18(6):593-599.

7. Chen S, Wang W, Lee S, et al; Colon Cancer Family Registry. Prediction of germline mutations and cancer risk in the Lynch syndrome. *JAMA*. 2006;296(12):1479-1487.

8. Evans JP, Berg JS, Olshan AF, Magnuson T, Rimer BK. We screen newborns, don't we? *Genet Med*. 2013;15(5):332-334.