EDITORIAL COMMENT

Hearing the Noise

The Challenges of Human Genome Variation in Genetic Testing*

Luisa Mestroni, MD,
Matthew R. G. Taylor, MD, PHD

Aurora, Colorado

The completion of the Human Genome Project was a landmark achievement that revealed the reference deoxyribonucleic acid (DNA) sequence for our own genome. Almost immediately, it became clear that there was no single "reference" DNA sequence, as even the approximately one-half dozen human DNA samples used by the Human Genome Project contained tens of thousands of variations (1). As clinical genetic testing becomes more mainstream and various projects under way perform full DNA genome sequencing in hundreds of subjects, the extent of this genetic variation is increasingly being appreciated. It is widely recognized that most of this variation is probably not relevant for determining health or risk for disease and

collectively has been referred to as "genetic noise." As in much of biology, separation of the "signal" from the "noise" can be challenging, and as molecular genetic sequencing expands in use and in the amount of DNA that can be sequenced in a single assay, distinguishing a diagnostic genetic change from background genetic variation will remain a difficult task for researchers and clinicians. Newer DNA sequencing technology can now complete the sequencing of an entire human genome several times over in a matter of days, orders of magnitude faster than the nearly 13 years required for the initial first pass done by the Human Genome Project consortium (2). This technology, which will shortly be widely used in clinical genetic testing, will undoubtedly add to the difficulty of distinguishing signal from noise.

In this issue of the *Journal*, Kapplinger et al. (3) eloquently illustrate the breadth of genetic variation (the "noise") in an important and life-threatening genetic disease, arrhythmogenic right ventricular cardiomyopathy (ARVC). Various investigators had previously studied ARVC families to identify genes and screen patient cohorts to determine the contribution of various ARVC genes to the overall population of ARVC families. Although the screening of control populations for discovered disease-causing mutations is fairly standard in genetic studies, comprehensive DNA sequencing of controls to measure background genetic variation usually is not undertaken. This report and a modest number of others provide compelling data that this evaluation in control populations should become standard for many future genetic studies.

Kapplinger et al. (3) studied 175 subjects from the Netherlands and the United States with confirmed diagnoses of ARVC. In this group of subjects, mutations were found in 58% of cases after sequencing 5 ARVC genes (*PKP2*, *DSP*, *DSG2*, *DSC2*, and *TMEM43*). Although these mutations were not labeled by the investigators as being definitely "pathogenic," the evaluation criteria used would have led to these mutations being classified as "probably" or "presumptively" pathogenic by clinical laboratories and, more importantly, by many clinicians reading the laboratory reports. Interestingly, using the same evaluation criteria for mutations, 16% of 427 healthy controls without ARVC also had mutations, illustrating the level of genetic "noise" and an overall frequency well beyond that predicted assuming the low prevalence of ARVC. In controls, the majority of detected mutations were missense mutations, suggesting that many of this class of mutations were likely benign. When the investigators turned to mutations predicted to cause more substantial consequences to the predicted ARVC protein structures (so-called radical mutations), the prevalence of ARVC mutations dropped to 0.5% in controls yet remained as high as 43% in probands. The criteria for "radical mutations" included in-frame and frame-shift insertions and deletions, splice junction, and nonsense mutations, providing some guidance in how to interpret the likelihood of a given variant's contributing to disease risk. However, as missense mutations may also cause disease, the interpretation of a novel ARVC gene missense mutation likely requires more than just reading the mutation report from the laboratory. The investigators found that in subjects with confirmed ARVC, missense mutations were grouped in "hot spots" in *DSP* and *DSG2* in regions of protein binding domains and that mutations in affected patients occurred in highly conserved residues across species, whereas controls' missense mutations localized in highly variable residues. The complexity involved in the interpretation of these mutations suggests that although ARVC clinical testing is accessible to many patients and cardiologists, significant skill may be required to properly

interpret mutation test results and separate pathogenic genetic mutations from background noise.

A previous study of long-QT syndrome by Kapa et al. (4) investigated the pathogenicity of long-QT genes (*SCN5A*, *KCNQ1*, and *KCNH2*). The investigators screened a large cohort (n = 388) of definitive long-QT syndrome cases and also studied more than 1,300 normal controls. In that case, the "background noise" of mutations was again significant, although at 6% less than in the current study in ARVC. One explanation for the difference in background noise levels is that the total number of patients and controls sequenced for long-QT syndrome was substantially greater than for ARVC; thus, we may expect that as more subjects are sequenced for the ARVC genes, some of the background mutation noise may be reclassified as rare genetic variation, reducing the ARVC mutation noise level somewhat.

An important element missing from the study of Kapplinger et al. (3) was the genetic evaluation of the patients' families. An analysis of cosegregation of a mutation within a family can be critical to help assess the causal role of a putative mutation. Unfortunately, this is difficult in research studies, in which the ascertainment of large families is not always possible. In clinical circumstances, efforts to evaluate, recruit, and test multiple patients in a given family are even less likely to be undertaken. In some circumstances, investigators have taken additional steps to assess mutation pathogenicity using in vitro cellular or in vivo animal assays, but this approach is difficult when large numbers of mutations are identified and is not possible in clinical situations when working with clinical laboratories.

Conclusions. Genetic testing continues to evolve, revealing it to be an imperfect tool that requires careful interpretation before and after testing is done. Criteria for pathogenic mutations are not convincingly settled upon and are liable to undergo some changes as more knowledge is gained. Indeed, the pace of clinical testing seems at times to have moved faster and without circumspect consideration than perhaps research efforts would dictate. Stringent crite-

ria ("radical" mutations, or missense mutations located in highly conserved and functionally important domains) and, whenever possible, cosegregation analysis can be used when applicable to help with genetic test result interpretation. Furthermore, as underlined by Kapplinger et al. (3) and other investigators (5–7), genetic tests must be integrated in the context of an expert clinical evaluation, together with a good family history and accurate clinical information, as with any other diagnostic test. Until the specificity of these types of molecular genetic tests is robust and understood, the clinical application of such tests is probably still better performed at referral centers with expertise in cardiovascular genetics (5–7).

REFERENCES

1. The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. Nature 2010;467:1061–73.
2. Metzker ML. Sequencing technologies—the next generation. Nature Rev Genet 2010;11:31–46.
3. Kapplinger JD, Landstrom AP, Salisbury BA, et al. Distinguishing arrhythmogenic right ventricular cardiomyopathy/dysplasia–associated mutations from background genetic noise. J Am Coll Cardiol 2011;57:2317–27.
4. Kapa S, Tester DJ, Salisbury BA, et al. Genetic testing for long-QT syndrome: distinguishing pathogenic mutations from benign variants. Circulation 2009;120:1752–60.
5. MacRae CA. Closer look at genetic testing in long-QT syndrome: will DNA diagnostics ever be enough? Circulation 2009;120:1745–8.
6. Hershberger RE, Lindenfeld J, Mestroni L, Seidman CE, Taylor MR, Tobwin JR. Genetic evaluation of cardiomyopathy: a HFSA comprehensive heart failure practice guideline. J Card Fail 2009;15:83–97.
7. National Center for Biotechnology Information. GeneTests. Available at: http://www.genetests.org. Accessed December 19, 2010.