

Standards for clinical use of genetic variants

The price of DNA sequencing has never been lower. However, there is little consensus as to when the identification of a genetic variant is clinically useful. Clinical geneticists carrying out systematic community reviews of evidence for the pathogenicity of variants collected in locus-specific and disease-specific databases are beginning to bridge the gap between research evidence and rules used to make clinical decisions.

Clinical classification of hereditary sequence variants identified in disease-associated genes directly affects clinical management of affected individuals and their relatives. Addressing this need in the Analysis on page 107, the International Society for Gastrointestinal Hereditary Tumours (InSiGHT) has applied expert curation to thousands of different mutations found in families affected by colon cancer all over the world, adding up the evidence for variants conferring inherited risk of cancer, identifying types of mutations that obviously impair the genome's ability to repair its DNA sequence and separating these variants from harmless common variants that confer no risk of cancer above that present in individuals without family-specific mutations. Individuals carrying rare 'spell-checking' mutations in mismatch-repair genes have a substantially increased risk of colorectal and endometrial cancer as well as a range of other cancers. However, once such a mutation causing Lynch syndrome has been detected in a family, intensive management of carrier individuals reduces their mortality. Expert review of the evidence for colon cancer risk conferred by constitutional variants in *MLH1*, *MSH2*, *MSH6* and *PMS2* has now resulted in the reclassification of two-thirds of the variants reported in existing databases and has led to clinical recommendations for the interpretation of 1,370 variants that do not result in obvious protein truncation. This kind of individual scrutiny of the evidence within an explicit framework for likelihood-based decision-making is by its nature painstaking, as mutations in different genes have different penetrance and the number of individuals in each family identified with the disease and with mutations varies for each report. Accordingly, precise attribution has been systematically credited to the participating laboratories reporting individual variants, published or unpublished, as an incentive to continue the collection, curation and stewardship of this clinically valuable data set.

In the absence of unequivocal support from segregation analysis, many variants still cannot currently be classified, for example, large duplications, most synonymous and some missense mutations, intronic variants and most variants in promoter and enhancer elements. Interpretation of variants in terms of their disruption of protein-coding sequence may also be misleading, as these variants may also be associated with transcription factor binding sites overlapping the coding regions of exons (*Science* **342**, 1367–1372, 2013) or microRNA target sites, or they may have effects on RNA splicing or stability.

In an earlier study this year (*Nat. Genet.* **45**, 1160–1167, 2013), *CFTR* variants were systematically assessed for the phenotypes of cystic fibrosis and sweat chloride concentration, and a set of missense variants was examined in transgenic cells in tissue culture for effect on protein localization and chloride channel function. In addition, the genetic architecture and biology of cystic fibrosis enabled the reassignment of a number of candidate missense variants as nonpathogenic when found in *trans* with the common *CFTR* allele encoding p.Phe508del in fathers, as a man without a functional *CFTR* allele would be sterile, lacking vas deferens development. Although this finding depends upon the particular gene mutations and their biological consequences, the concept of a genetic test for normal alleles is a powerful one.

These studies are models for turning genome sequence into medically actionable information via review by an expert community. The statistical model for assessment is extendable to other genetic disorders with and without explicit modeling of the protein functions encoded by the affected genes. The difficulty encountered in these studies adapting variant calls for clinical decision-making suggests that, in addition to research aimed at the discovery of new disease genes, we also need more discussion of systematic ways to assess the variety of variants that can now be found by genomic sequencing. ■