

GENETICS

Deep mutational scans with targeted editing

CRISPR-based gene editing can be scaled up to study the function of sequence variation in the context of native chromatin.

Genetic association can only go so far in determining the significance of genomic variants that sequencing studies uncover by the millions. “There’s an increasing appreciation that we can’t get all the way with genetics and we’re going to have to do functional analysis,” says Jay Shendure, a geneticist at the University of Washington. “And one-by-one type analysis doesn’t really cut it,” he says.

A promising approach is deep mutational scanning, which maps biological function by measuring the effects of systematic sequence variation. Once considered a ‘brute force’ technique, it can now be carried out efficiently with high-throughput sequencing—but with a nagging caveat.

“There’s always been this question about the biological relevance” of introducing sequences outside of chromatin, says Shendure. Synthetic sequences delivered by plasmid or inserted into arbitrary genomic sites lack the larger native regulatory context, making it difficult to match endogenous expression levels and potentially missing the effects on splicing and regulatory cues when testing protein-coding regions.

The Shendure lab is using another new technology, RNA-guided nucleases, to fix the problem. Clustered, regularly interspaced, short palindromic repeats (CRISPR)-Cas9 is a bacterial defensive system that guides the Cas9 nuclease via short RNA sequences to digest invading viral DNA. The guide RNA can also be engineered to cut mammalian genomic sequence, which prompts the cell to attempt repair; in the presence of donor template, homology-dependent repair replaces native sequence with the edited version.

The researchers transfected large numbers of mammalian cells with Cas9 and guide RNA, at the same time introducing a complex library of donor templates to direct many possible changes at a single site. Because the efficiency of getting both the cut and homology-directed repair is very low, first author Gregory Findlay introduced a short ‘handle’ sequence, enabling enrichment of successful edits during PCR amplification. The edits were also designed to destroy the short protospacer-adjacent motif, which

Cas9 needs to recognize its cut site, thus preventing repeat cuts and reducing the chance of introducing nonhomologous changes.

The approach allowed the researchers to introduce every possible sequence combination into a 6-base portion of a *BRCA1* exon. Using RNA sequence as a readout, the team measured the effects of each change on transcript splicing. They also tested every single-base variant in the 78-base-long exon. The number of edited cells is still a bottleneck, and Shendure stresses that replicates are important to confidently interpret the results.

As with all deep mutational scans, a critical conceptual design issue is to connect sequence changes to a functional assay. “And the key is to identify generic approaches, ideally,” says Shendure. The scientists screened a large number of protein-coding changes in the essential gene *DBR1* and were able to relate sequences to changes in cellular fitness, an approach that should work for any essential gene. The work was carried out in haploid cells but should also work in diploid cells that lack one functional copy of the gene.

The method will not entirely replace other mutational scanning approaches because of the difficulty of connecting functional assays with the sequence changes, the efficiency bottleneck and the restricted size of editing regions, which may be constrained by the dynamics of the homologous repair process. Shendure and Findlay are working on scanning entire genes more efficiently. Currently this requires tiling across exons, which is more labor intensive than simply testing an open reading frame library in other systems.

Shendure points out that maintaining context goes beyond isolating individual changes to a genome—ideally, these types of large-scale functional assays will be possible in complex tissues or organisms. But the approach moves genetics a step forward on the important question of function, by providing precomputed functional values that could be “translated into probabilities and odds ratios for interpreting variants of uncertain significance,” he says.

Tal Nawy

RESEARCH PAPERS

Findlay, G.M. *et al.* Saturation editing of genomic regions by multiplex homology-directed repair. *Nature* **513**, 120–123 (2014).