P. Morgan/NPG

DISEASE GENETICS

# Network effects of disease mutations

Of the >100,000 genetic variants that have been statistically associated with human disease, a major challenge is to identify and characterize the subset that have functional consequences. Such information helps to pinpoint variants that causally contribute to diseases (thus refining disease risk estimates for individuals harbouring these variants) and can reveal the underlying molecular mechanisms. A new study uses high-throughput functional assays for an extensive characterization of the effects of disease-associated variants on protein stability and intermolecular interactions.

Sahni *et al.* focused on disease-associated missense genetic variants, which result in amino acid substitutions in the encoded proteins. Based on mutation data from the Human Gene Mutation Database (HGMD), they generated human gene libraries consisting of 2,890 missense gene variants (1–4 variants per gene) and the corresponding 1,140 wild-type gene counterparts.

To determine the extent to which these mutations caused protein unfolding (as a measure of protein destabilization), the authors used a high-throughput luciferase-based screen to test 2,332 mutant and 992 wild-type proteins for binding to 7 diverse chaperone proteins. Almost one-third (~28%) of mutations resulted in increased binding to chaperones. For these variants, further evidence for unfolding and decreased stability was provided by experimental verification of reduced expression levels and lower protein solubility, as well as by computational analyses of predicted perturbations to the protein structure.

As most of the disease-associated mutations did not substantially affect protein stability, the authors tested for effects on intermolecular interactions. 2,449 protein variants and 1,072 wild-type counterparts were first tested in a yeast two-hybrid system for binding to 7,200 wild-type proteins, followed by a focused pairwise screen of the identified interactions using 460 mutant and 220 wild-type proteins. They categorized the mutations according to the extent to which they abolished protein–protein interactions (PPIs), classifying them as 'quasi-null' when they abolished all interactions, 'edgetic' when they lost a subset of specific interactions, and 'quasi-wild-type' when none of the tested interactions were apparently affected. Of these categories, the quasi-null variants were enriched for those variants that had scored in the chaperone-interaction screen, indicating that overall destabilization of a protein was underlying the perturbation of all its interactions.

Most of the disease-associated mutations had functional consequences on PPIs and were either edgetic or quasi-null. This was in contrast to the equivalent examination of proteins harbouring normal human variation, as these variant proteins were predominantly quasi-wild-type. Thus, interaction profiling could be a useful tool to distinguish benign from disease-causal variants. Indeed, the authors showed that their functional analyses performed favourably compared with purely computational predictors of variant deleteriousness, and the number of PPIs disrupted by the mutations correlated positively with disease severity.

Furthermore, the PPI data provided insights into the mechanisms by which different mutations in the same gene can have distinct clinical consequences. For example, different mutations in tropomyosin 3 (*TPM3*) resulted in either disruption of interactions with a subset of muscle-expressed proteins or a complete loss of interactions, which might explain the distinct forms of myopathy that are associated with these mutations.

As some of the disease-associated mutations in transcription factors were quasi-wild-type with respect to PPIs, the authors looked for effects on DNA binding. Using yeast one-hybrid assays to examine 70 mutant and 28 wild-type transcription factors for binding to 152 developmental enhancers, they found that >80% of mutations altered DNA-binding profiles. Rather than just conferring a general binding disruption, the mutations generally caused complex edgetic effects by impairing binding to a subset of sites while increasing binding to other sites.

Overall, this study reports a valuable strategy and resource for investigating the functional consequences of disease-associated genetic variants. Further advances could be to incorporate additional interaction types (such as protein–RNA interactions) and to extend analyses to additional types of variants in more genomic regions.

*Darren J. Burgess*

**ORIGINAL RESEARCH PAPER** Sahni, N. *et al.* Widespread macromolecular interaction perturbations in human genetic disorders. *Cell* **161**, 647–660 (2015)