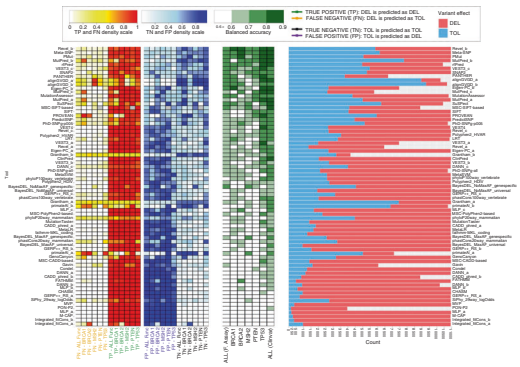
IN THIS ISSUE

Most in silico prediction tools overcall deleterious variants

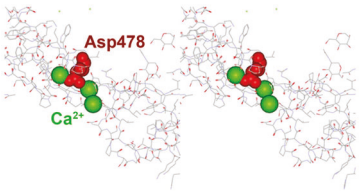
https://doi.org/10.1038/s41436-021-01265-z



Sequence analysis to inform diagnosis and prediction of human Mendelian diseases is now routine, and common frameworks such as those developed by the American College of Medical Genetics and Genomics/Association for Molecular Pathology facilitate variant interpretation. However, variant interpretation for rare missense variants identified in clinical genetic testing frequently relies on predicted changes of protein function using sequence alignment and three-dimensional protein structure modeling. Over the last 20 years, researchers have developed in silico prediction algorithms and tools trained against a truth set—a data set in which the impact of variants on protein function is already quantified. In this issue, Cubuk and colleagues evaluated the predictive performance of more than 40 widely used in silico tools for variants in five cancer susceptibility genes. The researchers generated a combined functional truth set for more than 9,000 missense variants in *BRCA1*, *BRCA2*, *MSH2*, *PTEN*, and *TP53* from clinically validated functional assays. They also assembled a ClinVar truth set for the variants that retained pathogenic/likely pathogenic or benign/likely benign classification. In the combined functional truth set, the prevalence of deleterious variants was 15%. However, the total proportion called as deleterious by in silico tools was greater than 50% for more than half of the 70 tool-threshold combinations examined. The metatools REVEL and Meta-SNP performed the best, with approximately 90% sensitivity and 67% specificity. In contrast, mean positive predictive values for common tools such as SIFT, PolyPhen-2, and Mutation Taster were 26–30%. Most tools were unlikely to miscall truly deleterious variants; however, the high sensitivity came at the expense of very high false-positive call rates. Together the findings indicated wide variation in the predictive performance of many commonly used in silico tools, with REVEL and Meta-SNP displaying the best accuracy. —*V. L. Dengler, News Editor.*

Biallelic *PCDHGC4* variants lead to a novel neurodevelopmental disorder

https://doi.org/10.1038/s41436-021-01260-4



Protocadherins (PCDHs) are a large family of calcium-dependent cell adhesion proteins with prominent expression in the developing vertebrate nervous system. PCDHs, which facilitate discrimination between self and nonself cell surface identities within neuronal circuits, are encoded by two families of genes: nonclustered and clustered (cPCDH). Although rare variants in nonclustered PCDH genes have been found in individuals with neurodevelopmental disorders such as Usher syndrome, a disease-causing variant in cPCDH for a Mendelian disorder in humans had not previously been identified. Iqbal and colleagues discovered biallelic variants in protocadherin-gamma-C4 (*PCDHGC4*) that result in a novel neurodevelopmental disorder. The researchers recruited 19 individuals from nine unrelated families with a neurodevelopmental disorder. Patients displayed developmental delay/intellectual disability, progressive microcephaly, seizures, hypotonia, short stature, and skeletal/joint anomalies. The researchers sequenced probands and proband/parent trios and performed linkage analysis on one family. Given parental consanguinity across all families, the researchers prioritized homozygous, rare exonic, and splice site variants. Sequencing identified three missense variants and five nonsense variants in *PCDHGC4* in affected individuals. All of the variants fully cosegregated with the phenotype and are absent or rare in the general human population, with biallelic variant identification consistent with an autosomal recessive pattern of inheritance. Four homozygous, loss-of-function variants were predicted to result in truncated proteins, and the fifth nonsense variant resulted in aberrant splicing and premature protein truncation. Using the crystal structure of a mouse homologue, the researchers found that all three missense variants are located in or directly adjacent to the calcium-binding motif of *PCDHGC4* and likely impair calcium affinity of the protein. Together, the results provide evidence that biallelic variants in *PCDHGC4* underlie a novel autosomal recessive neurodevelopmental disorder. —*V. L. Dengler, News Editor.*

Genetics in Medicine (2021) https://doi.org/10.1038/s41436-021-01309-4