A comprehensive landscape of CYP2D6 variation across 30,000 individuals

Shishi Luo[1], Ruomu Jiang[1], Jasmine Dhaliwal[1], Sharoni Jacobs[1], William Lee[1]

[1] Helix, San Carlos, California

More than one hundred CYP2D6 star alleles (haplotypes of the CYP2D6 gene) are listed in the Pharmacogene Variation Consortium (PharmVar) database, but commercial pharmacogenomic analyses only test for a minor fraction of these, often missing alleles that are common in certain populations. This is due, in part, to structural rearrangements in CYP2D6 (such as the CYP2D6*36 and CYP2D6*68 alleles) that require novel analytical tools and custom assay design to guarantee high confidence allele calls. These arrangements cannot typically be ascertained by off-the-shelf assays. This gap—between alleles that are known to exist and alleles that are reported in tests—can lead to mischaracterization of an individual's CYP2D6 genotype. In the clinical setting, this misclassification may result in prescriptions that are harmful to the patient. In the research setting, misclassification means poor quality input data for association tests. We have developed a clinically validated pipeline that calls 106 CYP2D6 star alleles listed in the PharmVar database, including CYP2D6*36, CYP2D6*59, and CYP2D6*68, which occur with nontrivial frequency but are not typically reported. This pipeline is built on Helix's Exome+ assay and achieves over 99% accuracy without requiring orthogonal assay supplementation. Here we present the results of applying this pipeline to 30,000 exomes, providing an unprecedented characterization of CYP2D6 genotypic diversity in the US population.