GOLDEN HELIX
Accelerating the Quest for Significance™

Five Functional Prediction Algorithms for NGS Data

Algorithm	Pub Year	Citations	Host Institution	Category	Distinguishing Characteristic	URL
SIFT	2003	>1200	JCVI (UW)	Untrained	Popular, broadly applicable and intuitive method to identify functional mutations.	*sift.jcvi.org*
PolyPhen-2	2010	>1000	Harvard/BWH	Trained	Provides two scores (HumDiv and HumVar) for applications to complex and Mendelian disease, respectively.	*genetics.bwh.harvard.edu/pph2*
MutationAssessor	2011	57	MSKCC	Untrained	Considers amino acid conservation in protein subfamilies to refine important functional regions. Interactive user interface.	*mutationassessor.org*
MutationTaster	2010	199	Charite-Berlin	Trained	Native support for DNA (rather than amino acid) variant analysis. Allows online submission of VCF files.	*mutationtaster.org*
FATHMM	2013	NA	University of Bristol	Trained (weighted)	Uses HMM method (rather than BLAST) to create MSA. Weighted extensions for human disease and cancer analysis.	*fathmm.biocompute.org.uk*

Cautions on Usage

- The common belief is that variants called damaging by multiple algorithms are most likely to have true disease causing potential. (Although this is not always true.)
- Published comparisons aren't exhaustive, and usually focus on prediction performance for detecting a particular category of mutations.
- Each prediction tool has its own strengths and weaknesses and may carry certain biases based on the authors' own research interests.
- All of the algorithms generally perform well for distinguishing between known damaging variants and known neutral variants.
- False positive rate can be high when the methods are applied to a broad range of variants of unknown significance.

GOLDEN HELIX
Accelerating the Quest for Significance™