# Mutation Assessor

**Approach**: A prediction of the functional impact of protein missense mutations is based on the assessment of evolutionary conservation of amino-acid residues in a protein family multiple sequence alignment [1]. The novelty of the approach is in exploiting the evolutionary conservation in protein subfamilies, which are determined by clustering multiple sequence alignments of homologous sequences on the background of conservation of overall function [2].

**Validation (Figures 1 and 2)**: The scoring function was validated by separation of a large set of disease-associated variants (20,000) from common (benign) polymorphisms (36,500) with the accuracy of 79% and the area under the curve (AUC) 0.86 in the receiver-operation-characteristic (ROC) analysis for a two-class distinction [1]. The predictive power of the score for cancer mutations from COSMIC database (v.49 release of 29/Sep/2010) was assessed by separating assumed to be driver mutations (1800 recurrent mutations - observed in two or more samples and 700 highly recurrent mutations - observed in 5 or more samples) from assumed to be passenger mutations (8200 single mutations - observed only in one sample). The maximal separation accuracies are 69% (AUC 0.75) for recurrent vs. single and 78% (AUC 0.84) for highly recurrent vs. single. Mutations in multiply mutated genes and mutations in known cancer genes tend to have significantly higher functional impact scores than control sets.

**Performance** For every variant with predicted functional impact, the fully automated computational protocol provides with 30+ fields of biological annotations, links to multiple sequence alignments and 3D structures; the service is capable of processing thousands of variants through WEBAPI (built to process output of sequencing machines). Genome or protein coordinates are used as input, two genome builds are supported.

**Availability**: http://mutationassessor.org.

### References

1. Reva, B.A., Antipin, Y.A. and Sander, C. (2010) Predicting the Functional Impact of Protein Mutations: Application to Cancer Genomics (submitted to Nucl. Acids Res.)

2. Reva, BA, Antipin, YA, and Sander, C. (2007) Determinants of protein function revealed by combinatorial entropy optimization. Genome Biol.; 8(11): R232.

Figure 1. Separation of disease-associated and polymorphic variants by functional impact score.

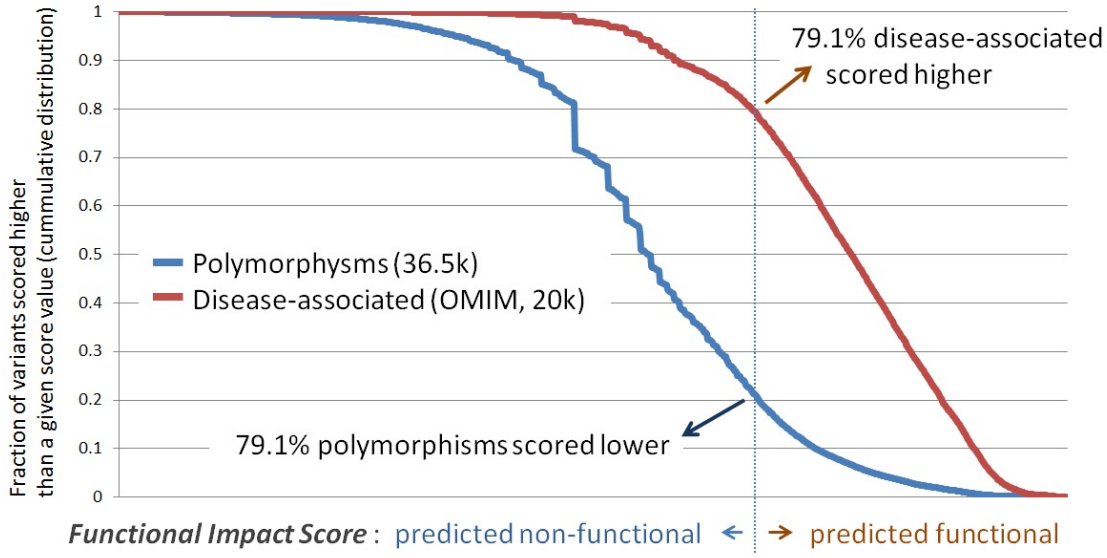


Figure 2. Separation of recurrent and singly observed cancer mutations of COSMIC database by the functional impact score.