Description of Additional Supplementary Files

Supplementary Data 1

MisFit estimated gene-level missense selection for 830 genes with known disease mechanisms. ‘MisFit_sgene_mis’ and ‘logit_s_map_mis’ columns represent MisFit predicted gene-level missense selection in original and logit scale via maximize a posteriori (MAP). ‘Type’ column indicates known disease mechanism of the gene: autosomal recessive (AR), gain of function (GOF), haploinsufficient (HI), dominant negative (DN).

Supplementary Data 2

Genetic variants data used in analysis. ‘Sample’ is the number of sequenced samples. ‘Synonymous’, ‘missense’ and ‘protein_truncating’ columns represent number of unique variants of the type in the dataset.

Supplementary Data 3

Deep mutational scanning experiments used in analysis and Spearman correlation coefficient with multiple computational methods.

Supplementary Data 4

Deep mutational scanning experiments that have bimodal distribution of functional scores used in analysis, and AUROC of computational methods in classification.

Supplementary Data 5

MisFit estimated gene-level missense selection for all genes trained in model. ‘MisFit_sgene_mis’ and ‘logit_s_map_mis’ columns represent MisFit predicted gene-level missense selection in original and logit scale via maximize a posteriori (MAP).