Supplementary Tables

Supp. Table S1. Performance evaluation of 10 state-of-the-art pathogenicity prediction tools over 5 datasets

		MT2	PP2	MASS	CADD	SIFT	LRT	FatHMM-U	FatHMM-W	Gerp++	phyloP
HumVar	**TP**	18020	17432	16214	-	15877	16520	12920	16392	-	-
	FP	2670	3706	3002	-	3653	3580	3048	2159	-	-
	TN	15392	15592	15982	-	14804	13294	15651	16538	-	-
	FN	1685	3658	4800	-	4335	3460	7251	3763	-	-
	AUC	0.89	0.89	0.88	0.88	0.86	0.83	0.81	0.93	0.75	0.80
	AUC-PR	0.91	0.89	0.89	0.87	0.88	0.86	0.83	0.93	0.74	0.76
	Accuracy	0.88	0.82	0.80	-	0.79	0.81	0.74	0.85	-	-
	F-Score	0.89	0.83	0.81	-	0.80	0.82	0.72	0.85	-	-
	MCC	0.77	0.63	0.61	-	0.59	0.62	0.48	0.70	-	-
	Precision	0.87	0.82	0.84	-	0.81	0.82	0.81	0.88	-	-
	Recall	0.91	0.83	0.77	-	0.79	0.83	0.64	0.81	-	-
	Specificity	0.85	0.81	0.84	-	0.80	0.79	0.84	0.88	-	-
	NPV	0.90	0.81	0.77	-	0.77	0.79	0.68	0.81	-	-
ExoVar	**TP**	4444	4284	4013	-	3884	4124	3140	4070	-	-
	FP	1313	1118	862.00	-	1008	1018	826	497	-	-
	TN	2147	2576	2757	-	2505	2120	2717	2884	-	-
	FN	347	872	1116	-	1039	824	1831	895	-	-
	AUC	0.87	0.84	0.85	0.83	0.82	0.81	0.76	0.92	0.73	0.74
	AUC-PR	0.90	0.88	0.89	0.84	0.87	0.88	0.82	0.94	0.76	0.75
	Accuracy	0.80	0.78	0.77	-	0.76	0.77	0.69	0.83	-	-
	F-Score	0.84	0.81	0.80	-	0.79	0.82	0.70	0.85	-	-
	MCC	0.59	0.53	0.54	-	0.50	0.52	0.39	0.66	-	-
	Precision	0.77	0.79	0.82	-	0.79	0.80	0.79	0.89	-	-
	Recall	0.93	0.83	0.78	-	0.79	0.83	0.63	0.82	-	-
	Specificity	0.62	0.70	0.76	-	0.71	0.68	0.77	0.85	-	-
	NPV	0.86	0.75	0.71	-	0.71	0.72	0.60	0.76	-	-
VariBenchSelected	**TP**	2992	2819	2637	-	2876	2575	2172	3830	-	-
	FP	2766	2264	1777	-	2121	2317	1779	883	-	-
	TN	2823	3692	4073	-	3523	2911	3975	4521	-	-
	FN	1054	1487	1616	-	1343	1578	2089	420	-	-
	AUC	0.65	0.68	0.70	0.66	0.70	0.62	0.64	0.94	0.59	0.58
	AUC-PR	0.62	0.62	0.62	0.56	0.69	0.62	0.58	0.93	0.47	0.46
	Accuracy	0.60	0.63	0.66	-	0.65	0.58	0.61	0.87	-	-
	F-Score	0.61	0.60	0.61	-	0.62	0.57	0.53	0.85	-	-
	MCC	0.25	0.27	0.31	-	0.30	0.18	0.20	0.73	-	-
	Precision	0.52	0.55	0.60	-	0.58	0.53	0.55	0.81	-	-
	Recall	0.74	0.65	0.62	-	0.68	0.62	0.51	0.90	-	-
	Specificity	0.51	0.62	0.70	-	0.62	0.56	0.69	0.84	-	-
	NPV	0.73	0.71	0.72	-	0.72	0.65	0.66	0.91	-	-
predictSNPSelected	**TP**	50	7941	7207	-	7296	7573	5125	7079	-	-
	FP	274	1961	1544	-	1747	2207	1237	638	-	-
	TN	502	4137	4353	-	3914	3001	3815	4256	-	-
	FN	15	2059	2714	-	2287	2007	3219	1263	-	-
	AUC	0.78	0.79	0.80	0.76	0.79	0.71	0.74	0.93	0.67	0.68
	AUC-PR	0.40	0.86	0.87	0.80	0.87	0.83	0.83	0.95	0.74	0.72
	Accuracy	0.66	0.75	0.73	-	0.74	0.72	0.67	0.86	-	-
	F-Score	0.26	0.80	0.77	-	0.78	0.78	0.70	0.88	-	-
	MCC	0.23	0.47	0.45	-	0.45	0.37	0.36	0.70	-	-
	Precision	0.15	0.80	0.82	-	0.81	0.77	0.81	0.92	-	-
	Recall	0.77	0.79	0.73	-	0.76	0.79	0.61	0.85	-	-
	Specificity	0.65	0.68	0.74	-	0.69	0.58	0.76	0.87	-	-
	NPV	0.97	0.67	0.62	-	0.63	0.60	0.54	0.77	-	-
SwissVarSelected	**TP**	3391	3086	2457	-	2592	2985	2250	2039	-	-

	MT2	PP2	MASS	CADD	SIFT	LRT	FatHMM-U	FatHMM-W	Gerp++	phyloP
FP	3180	2623	2299	-	2515	2675	2089	1541	-	-
TN	4114	5580	5214	-	4828	3958	6080	6614	-	-
FN	829	1440	1943	-	1617	1184	2250	2381	-	-
AUC	0.73	0.71	0.68	0.73	0.68	0.68	0.67	0.71	0.65	0.68
AUC-PR	0.66	0.60	0.55	0.58	0.61	0.62	0.54	0.58	0.47	0.50
Accuracy	0.65	0.68	0.64	-	0.64	0.64	0.66	0.69	-	-
F-Score	0.63	0.60	0.54	-	0.56	0.61	0.51	0.51	-	-
MCC	0.36	0.35	0.25	-	0.26	0.30	0.25	0.29	-	-
Precision	0.52	0.54	0.52	-	0.51	0.53	0.52	0.57	-	-
Recall	0.80	0.68	0.56	-	0.62	0.72	0.50	0.46	-	-
Specificity	0.56	0.68	0.69	-	0.66	0.60	0.74	0.81	-	-
NPV	0.83	0.79	0.73	-	0.75	0.77	0.73	0.74	-	-

For columns with a “-” it was not possible to compute the metric since the tools return a score but not a prediction label.

Supp. Table S2. Performance evaluation of FatHMM-W, a logistic regression over the features used to weight FatHMM-W, and a protein majority vote (MV), over 5 datasets

		FatHMM-W	Logistic Regression over the Features ln(Wn) & ln(Wd)	Protein Majority Vote (MV)
HumVar	**AUC**	0.93	0.92 (0.00)	0.96 (0.00)
	AUC-PR	0.93	0.92 (0.00)	0.96 (0.00)
	Accuracy	0.85	0.84 (0.01)	0.88 (0.00)
	F1-Score	0.85	0.85 (0.01)	0.89 (0.00)
	MCC	0.70	0.69 (0.01)	0.76 (0.01)
	Precision	0.88	0.84 (0.01)	0.84 (0.01)
	Recall	0.81	0.86 (0.01)	0.95 (0.00)
	NPV	0.81	0.84 (0.01)	0.93 (0.00)
	Specificity	0.88	0.83 (0.01)	0.80 (0.01)
ExoVar	**AUC**	0.92	0.90 (0.01)	0.94 (0.01)
	AUC-PR	0.94	0.92 (0.01)	0.96 (0.01)
	Accuracy	0.83	0.84 (0.01)	0.81 (0.01)
	F1-Score	0.85	0.87 (0.01)	0.86 (0.01)
	MCC	0.66	0.66 (0.02)	0.62 (0.02)
	Precision	0.89	0.84 (0.01)	0.78 (0.01)
	Recall	0.82	0.89 (0.01)	0.94 (0.01)
	NPV	0.76	0.83 (0.02)	0.89 (0.02)
	Specificity	0.85	0.76 (0.03)	0.64 (0.02)
VariBenchSelected	**AUC**	0.94	0.93 (0.01)	0.98 (0.00)
	AUC-PR	0.93	0.92 (0.01)	0.98 (0.00)
	Accuracy	0.87	0.87 (0.01)	0.82 (0.01)
	F1-Score	0.85	0.85 (0.01)	0.82 (0.01)
	MCC	0.73	0.74 (0.02)	0.68 (0.02)
	Precision	0.81	0.84 (0.02)	0.71 (0.02)
	Recall	0.90	0.87 (0.01)	0.97 (0.00)
	NPV	0.91	0.89 (0.01)	0.97 (0.00)
	Specificity	0.84	0.87 (0.02)	0.71 (0.02)
predictSNPSelected	**AUC**	0.93	0.92 (0.01)	0.97 (0.00)
	AUC-PR	0.95	0.95 (0.01)	0.98 (0.00)
	Accuracy	0.86	0.85 (0.01)	0.88 (0.01)
	F1-Score	0.88	0.89 (0.01)	0.91 (0.00)
	MCC	0.70	0.68 (0.02)	0.74 (0.01)
	Precision	0.92	0.86 (0.01)	0.85 (0.01)
	Recall	0.85	0.91 (0.01)	0.97 (0.01)
	NPV	0.77	0.84 (0.01)	0.93 (0.01)
	Specificity	0.87	0.74 (0.02)	0.73 (0.01)
SwissVarSelected	**AUC**	0.71	0.70 (0.01)	0.84 (0.01)
	AUC-PR	0.58	0.57 (0.02)	0.79 (0.01)
	Accuracy	0.69	0.69 (0.01)	0.75 (0.01)
	F1-Score	0.51	0.45 (0.02)	0.67 (0.01)
	MCC	0.29	0.27 (0.03)	0.47 (0.02)
	Precision	0.57	0.61 (0.03)	0.63 (0.01)
	Recall	0.46	0.35 (0.02)	0.72 (0.01)
	NPV	0.74	0.71 (0.01)	0.83 (0.01)
	Specificity	0.81	0.88 (0.01)	0.76 (0.01)

Supp. Table S3. Number of pathogenic and neutral variants from *VariBenchSelected* in identical or similar proteins in *HumVar* and *ExoVar*

Protein Similarity (%)	#Pathogenic Variants in Similar Proteins	Pathogenic Variants in Similar Proteins (%)	#Neutral Variants in Similar Proteins	Neutral Variants in Similar Proteins (%)
100	3929	91.1%	4026	67.6%
90	4241	98.4%	4207	70.6%
80	4248	98.6%	4255	71.4%
70	4255	98.7%	4333	72.7%
60	4261	98.8%	4482	75.5%
50	4276	99.2%	4704	79.0%
40	4283	99.4%	4976	83.5%
30	4289	99.5%	5306	89.1%

Supp. Table S4. Number of pathogenic and neutral variants from *predictSNPSelected* in identical or similar proteins in *HumVar* and *ExoVar*

Protein Similarity (%)	#Pathogenic Variants in Similar Proteins	Pathogenic Variants in Similar Proteins (%)	#Neutral Variants in Similar Proteins	Neutral Variants in Similar Proteins (%)
100	8584	82.1%	3813	57.0%
90	10221	97.8%	4708	70.4%
80	10265	98.2%	4771	71.3%
70	10275	98.3%	4855	72.6%
60	10280	98.4%	5031	75.2%
50	10299	98.5%	5267	78.7%
40	10332	98.9%	5597	83.6%
30	10363	99.1%	5944	88.8%

Supp. Table S5. Number of pathogenic and neutral variants from *SwissVarSelected* in identical or similar proteins in *HumVar* and *ExoVar*

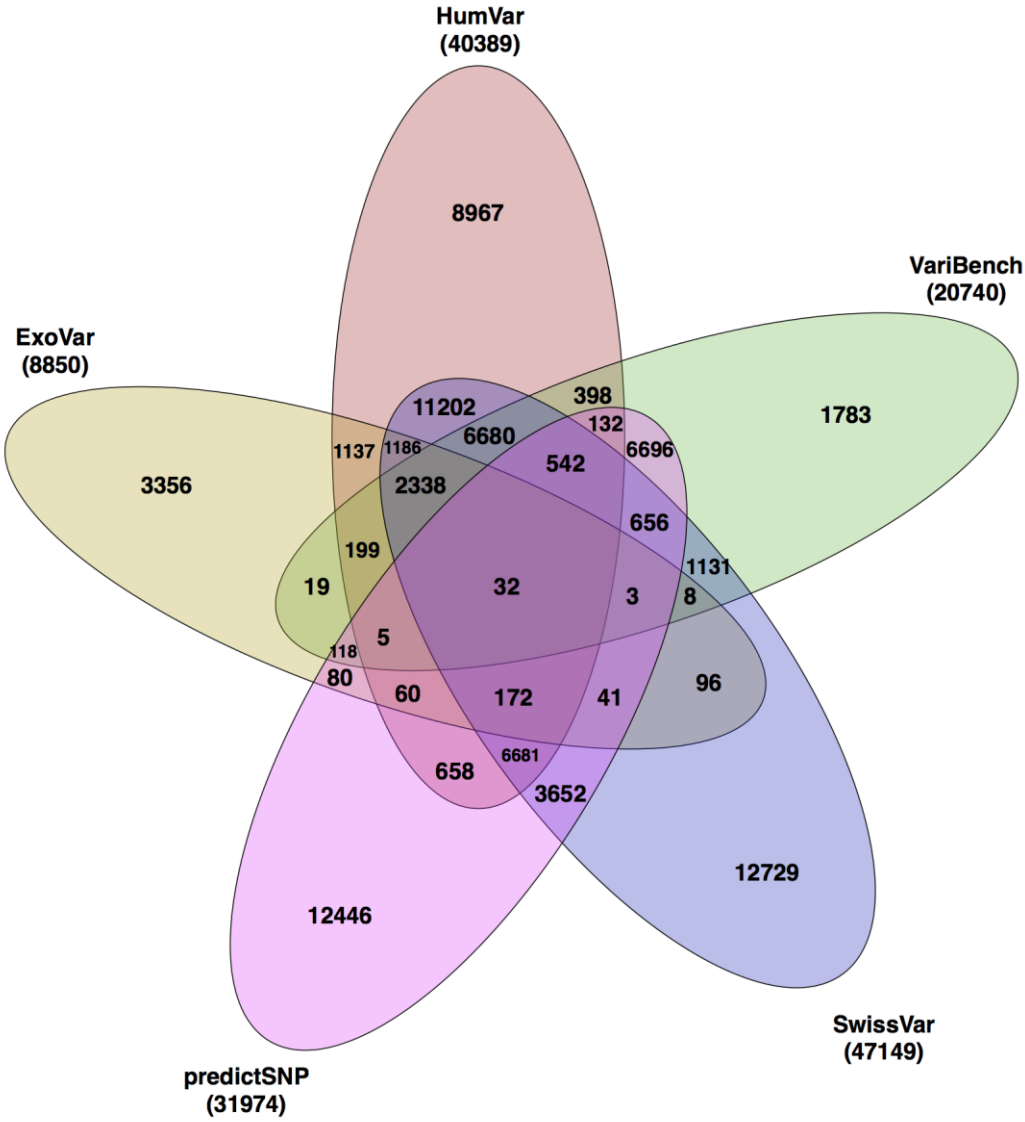
Protein Similarity (%)	#Pathogenic Variants in Similar Proteins	Pathogenic Variants in Similar Proteins (%)	#Neutral Variants in Similar Proteins	Neutral Variants in Similar Proteins (%)
100	2745	60.6%	4562	55.6%
90	3363	74.3%	5753	70.1%
80	3428	75.7%	6155	75.0%
70	3516	77.7%	6341	77.3%
60	3644	80.5%	6546	79.8%
50	3791	83.8%	6794	82.8%
40	3955	87.4%	7138	87.0%
30	4088	90.3%	7496	91.4%

Supp. Table S6. Overview of all combined predictors

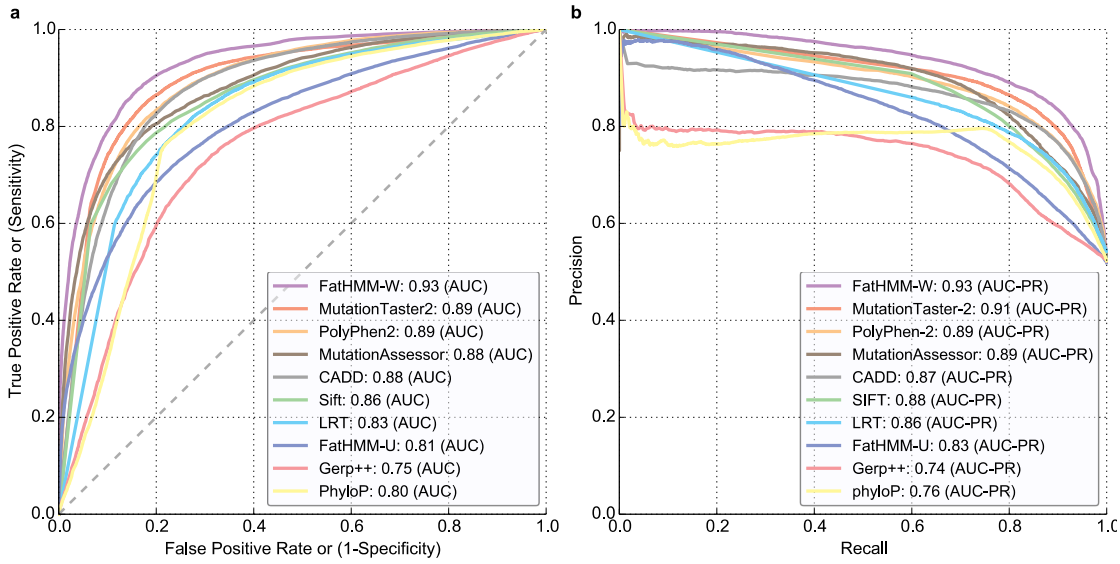
Combined Predictors (Abbreviation)	Combined Tools	Version	Tool/Website
Condel	PP2, SIFT, MASS	Old website (will be replaced by new website end of April 2014)	http://bg.upf.edu/condel
Logit	PP2, SIFT, MASS	KGGSeq 0.4	http://statgenpro.psychiatry.hku.hk/limx/kggseq/
Condel+	PP2, SIFT, MASS, FatHMM-W	New Website (April 2014)	http://bg.upf.edu/fannsdb/
Logit+	PP2, SIFT, MASS, FatHMM-W	KGGSeq 0.4	http://statgenpro.psychiatry.hku.hk/limx/kggseq/

A list of all investigated combined predictors, the tools that were used to create them, and links to the corresponding webservers.

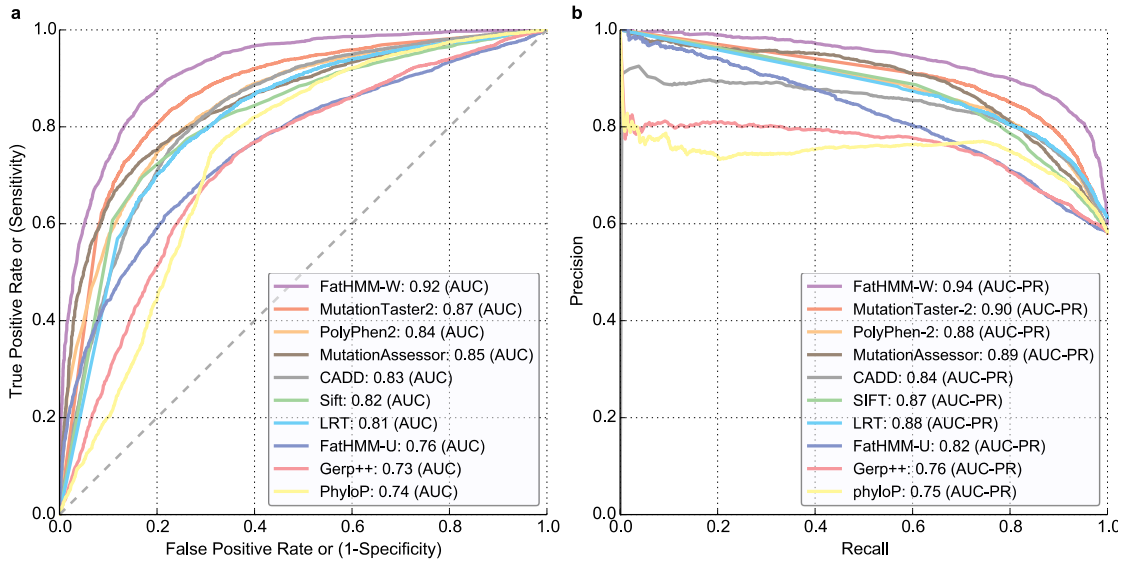
Supplementary Figures



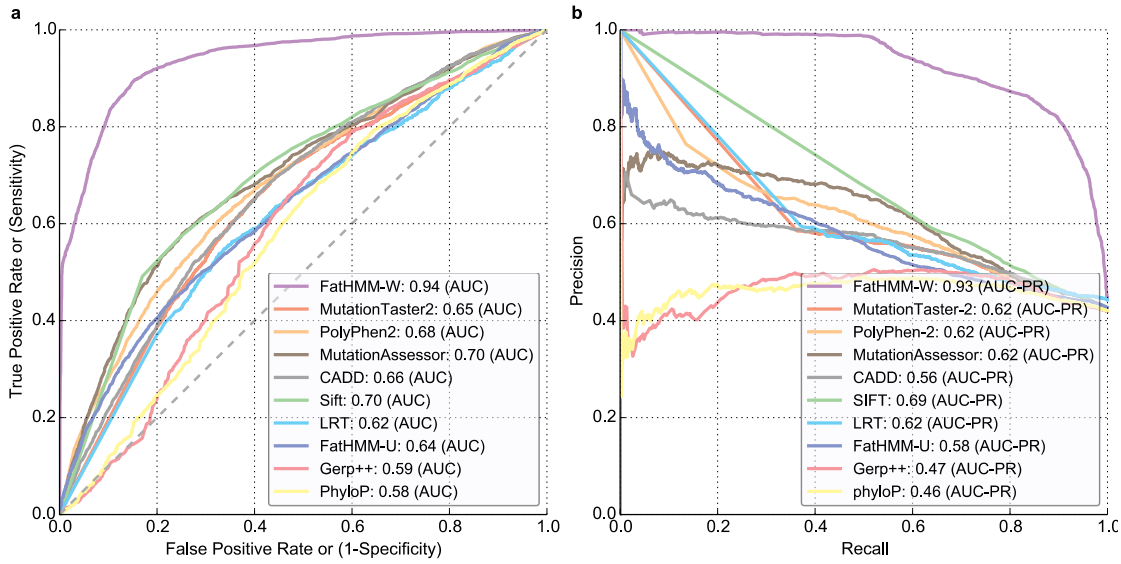
Supp. Figure S1. Venn diagram showing the overlap between five datasets used in this study. *VariBenchSelected* (10266 variants) is the part of *VariBench* not overlapping with *HumVar* nor *ExoVar*. *predictSNPSelected* (16098 variants) is the part of *predictSNP* not overlapping with *HumVar*, *ExoVar* nor *VariBench. SwissVarSelected* (12729 variants) is the part of *SwissVar* that does not overlap with *HumVar*, *ExoVar*, *VariBench, nor predictSNP*.



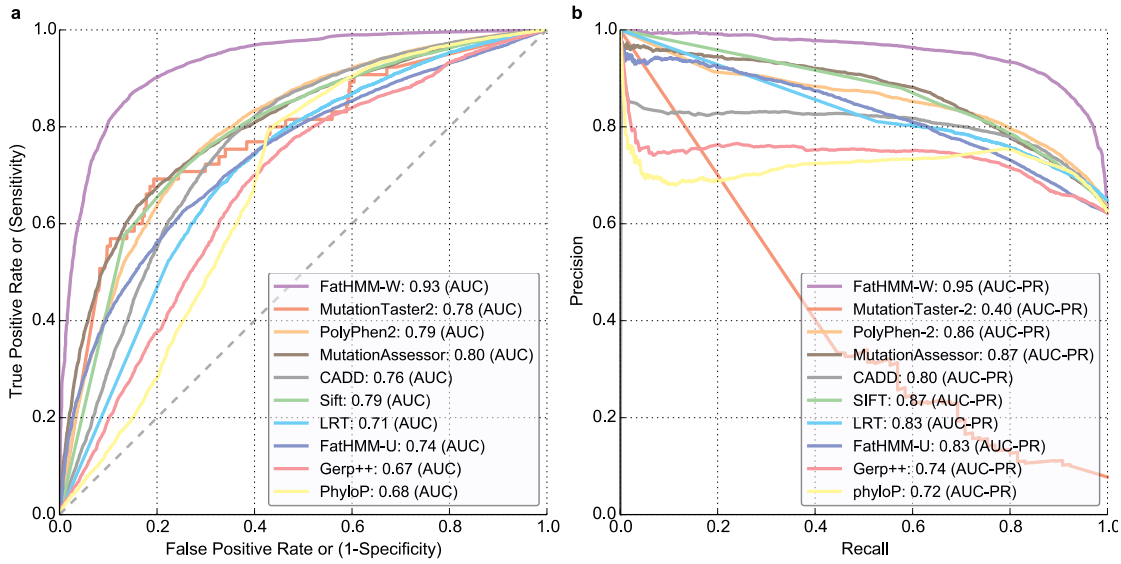
Supp. Figure S2. Predictive performance of 10 state-of-the-art pathogenicity prediction tools on the *HumVar* dataset. (a) ROC curves and corresponding AUC values. (b) Precision-Recall curves (ROC-PR) and corresponding AUC-PR values.



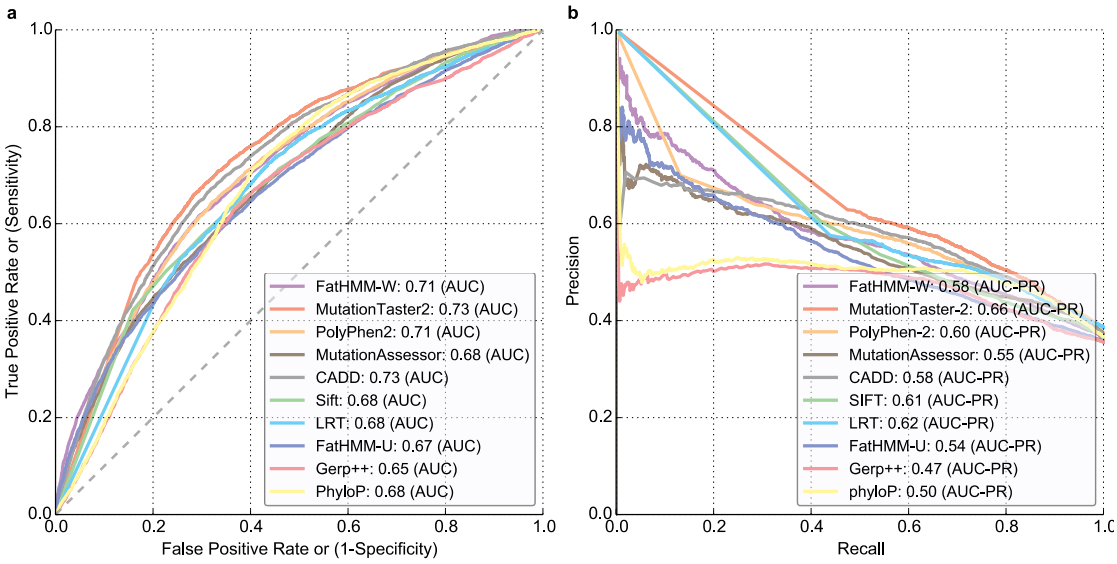
Supp. Figure S3. Predictive performance of 10 state-of-the-art pathogenicity prediction tools on the *ExoVar* dataset. (a) ROC curves and corresponding AUC values. (b) Precision-Recall curves (ROC-PR) and corresponding AUC-PR values.



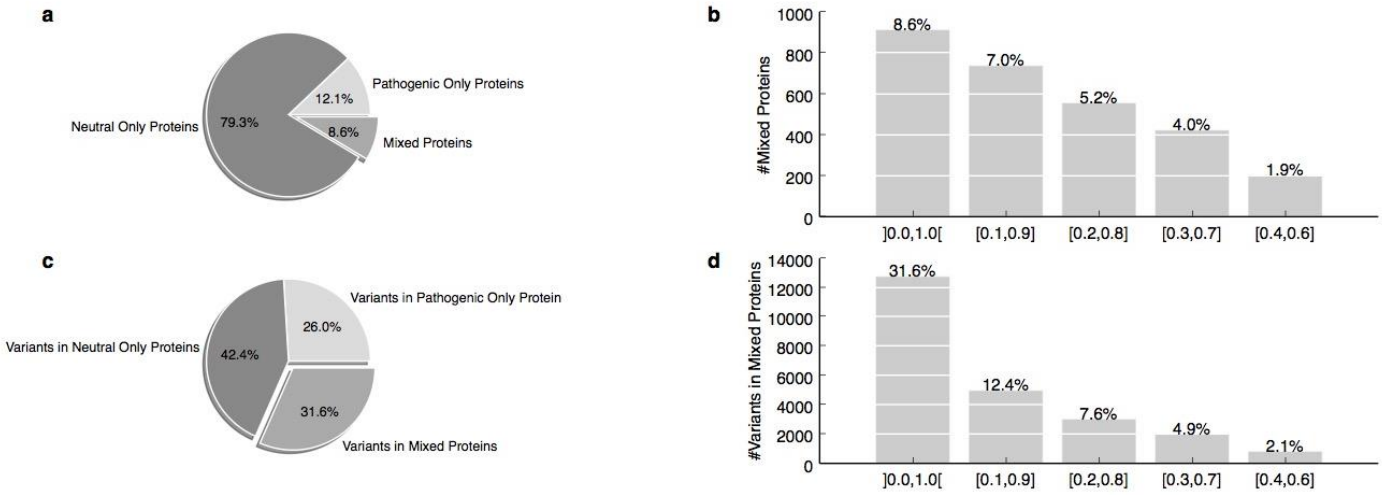
Supp. Figure S4. Predictive performance of 10 state-of-the-art pathogenicity prediction tools on the *VariBenchSelected* dataset. (a) ROC curves and corresponding AUC values. (b) Precision-Recall curves (ROC-PR) and corresponding AUC-PR values.



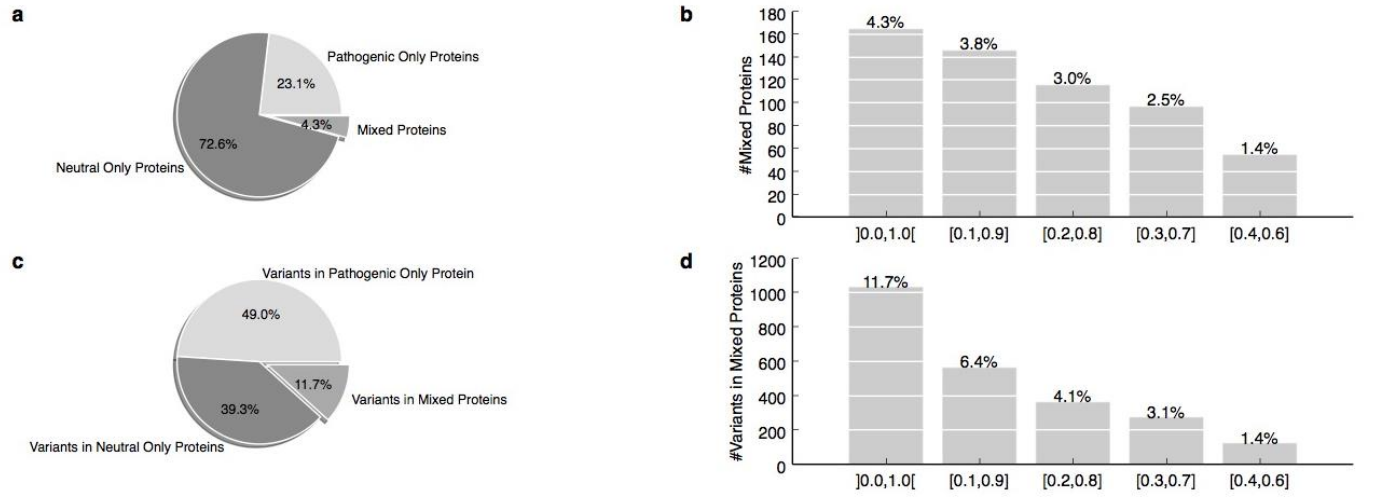
Supp. Figure S5. Predictive performance of 10 state-of-the-art pathogenicity prediction tools on the *predictSNPSelected* dataset. (a) ROC curves and corresponding AUC values. (b) Precision-Recall curves (ROC-PR) and corresponding AUC-PR values.



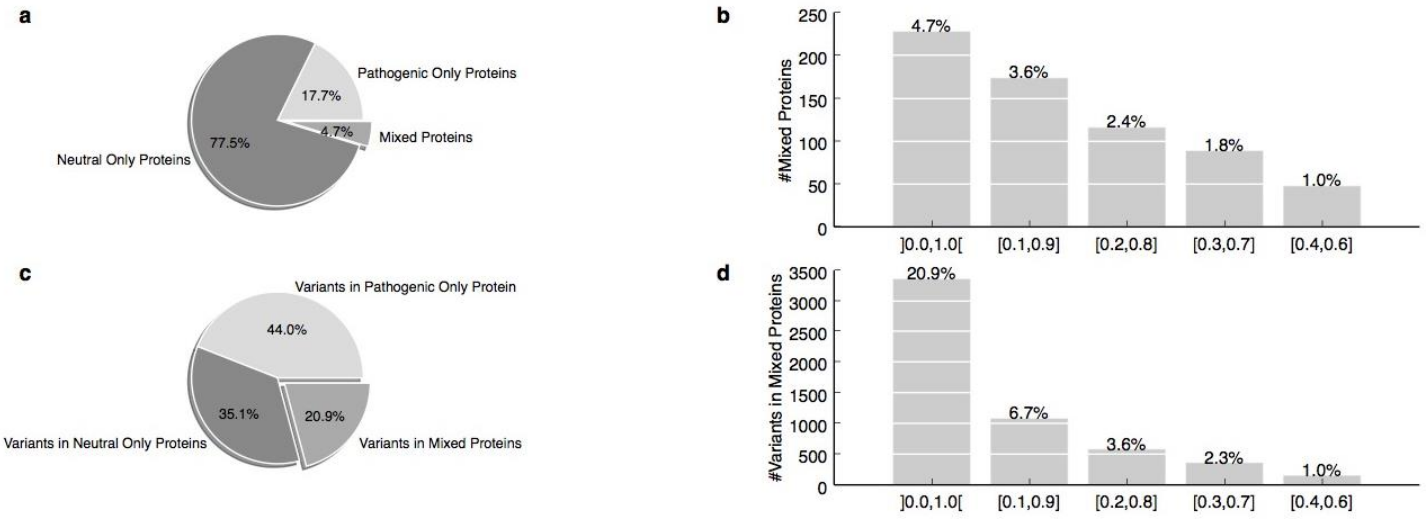
Supp. Figure S6. Predictive performance of 10 state-of-the-art pathogenicity prediction tools on the *SwissVarSelected* dataset. (a) ROC curves and corresponding AUC values. (b) Precision-Recall curves (ROC-PR) and corresponding AUC-PR values.



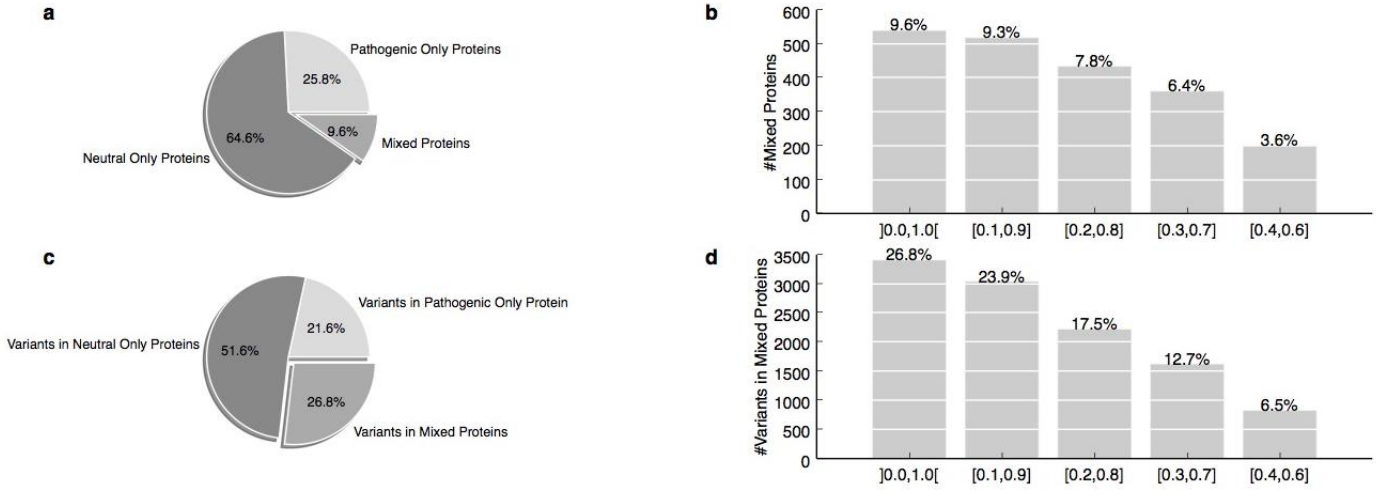
Supp. Figure S7. Composition of the *HumVar* dataset. (a) Relative proportions of proteins containing only neutral variants ("neutral-only"), proteins containing only pathogenic variants ("pathogenic-only"), and proteins containing both types ("mixed") in the ***HumVar*** data set. 8.6% of the proteins are mixed. (b) Fractions of proteins of ***HumVar*** containing various ratios of pathogenic-to-neutral variants, binned into narrower and narrower bins closer and closer to perfectly balanced proteins. The open interval]0.0, 1.0[corresponds to all mixed proteins; [x,y] indicates that the corresponding bar reports the fraction of all proteins in the data that have a ratio of pathogenic-to-neutral variants greater than or equal to x and lower than or equal to y. (c) Relative proportions, in the ***HumVar*** data set, of the number of variants belonging to these three categories of proteins. (d) Fractions of variants, in the ***HumVar*** data set, belonging to those same categories of mixed proteins. 2.1% of all variants belong to almost perfectly balanced proteins.



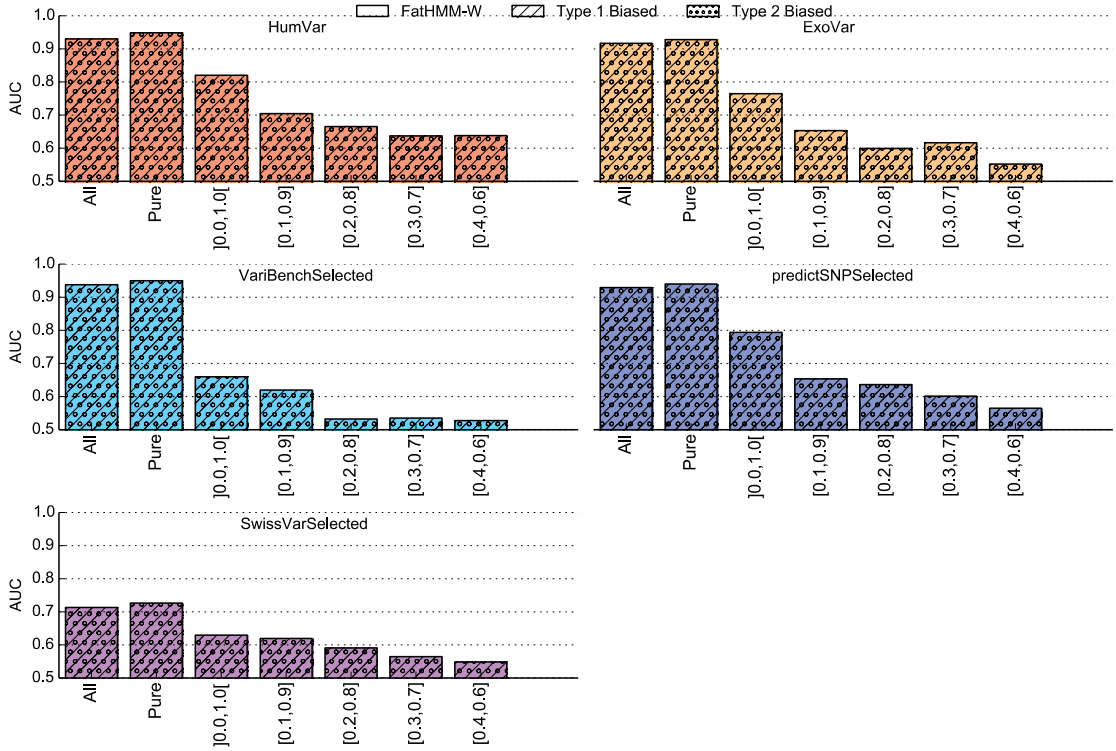
Supp. Figure S8. Composition of the *ExoVar* dataset. (a) Relative proportions of proteins containing only neutral variants ("neutral-only"), proteins containing only pathogenic variants ("pathogenic-only"), and proteins containing both types ("mixed") in the ***ExoVar*** data set. 4.3% of the proteins are mixed. (b) Fractions of proteins of ***ExoVar*** containing various ratios of pathogenic-to-neutral variants, binned into narrower and narrower bins closer and closer to perfectly balanced proteins. The open interval]0.0, 1.0[corresponds to all mixed proteins; [x,y] indicates that the corresponding bar reports the fraction of all proteins in the data that have a ratio of pathogenic-to-neutral variants greater than or equal to x and lower than or equal to y. (c) Relative proportions, in the ***ExoVar*** data set, of the number of variants belonging to these three categories of proteins. (d) Fractions of variants, in the ***ExoVar*** data set, belonging to those same categories of mixed proteins. 1.4% of all variants belong to almost perfectly balanced proteins.



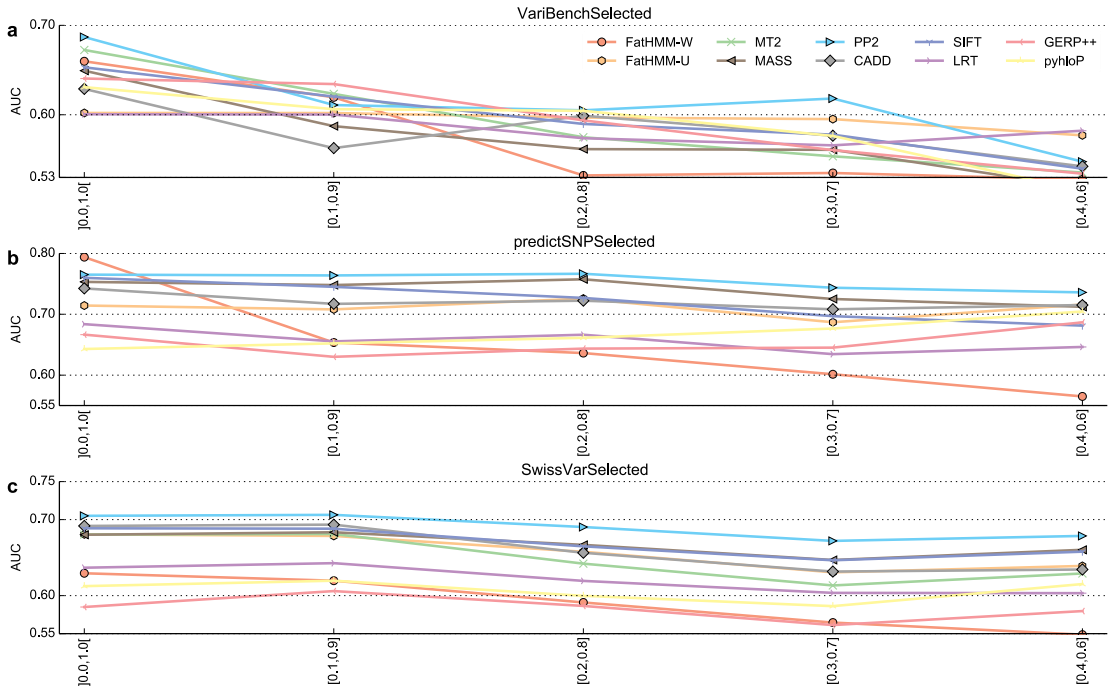
Supp. Figure S9. Composition of the dataset *predictSNPSelected*. (a) Relative proportions of proteins containing only neutral variants ("neutral-only"), proteins containing only pathogenic variants ("pathogenic-only"), and proteins containing both types ("mixed") in the ***predictSNPSelected*** data set. 4.7% of the proteins are mixed. (b) Fractions of proteins of ***predictSNPSelected*** containing various ratios of pathogenic-to-neutral variants, binned into narrower and narrower bins closer and closer to perfectly balanced proteins. The open interval]0.0, 1.0[corresponds to all mixed proteins; [x,y] indicates that the corresponding bar reports the fraction of all proteins in the data that have a ratio of pathogenic-to-neutral variants greater than or equal to x and lower than or equal to y. (c) Relative proportions, in the ***predictSNPSelected*** data set, of the number of variants belonging to these three categories of proteins. (d) Fractions of variants, in the ***predictSNPSelected*** data set, belonging to those same categories of mixed proteins. 1.0% of all variants belong to almost perfectly balanced proteins.



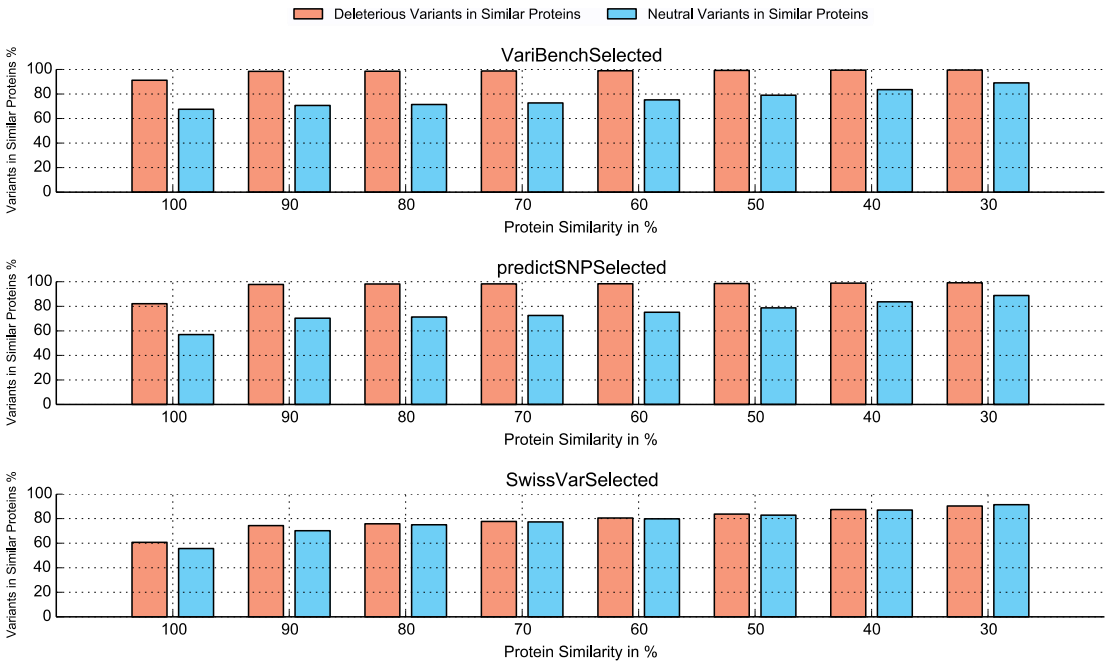
Supp. Figure S10. Composition of the dataset *SwissVarSelected*. (a) Relative proportions of proteins containing only neutral variants ("neutral-only"), proteins containing only pathogenic variants ("pathogenic-only"), and proteins containing both types ("mixed") in the ***SwissVarSelected*** data set. 9.6% of the proteins are mixed. (b) Fractions of proteins of ***SwissVarSelected*** containing various ratios of pathogenic-to-neutral variants, binned into narrower and narrower bins closer and closer to perfectly balanced proteins. The open interval]0.0, 1.0[corresponds to all mixed proteins; [x,y] indicates that the corresponding bar reports the fraction of all proteins in the data that have a ratio of pathogenic-to-neutral variants greater than or equal to x and lower than or equal to y. (c) Relative proportions, in the ***SwissVarSelected*** data set, of the number of variants belonging to these three categories of proteins. (d) Fractions of variants, in the ***SwissVarSelected*** data set, belonging to those same categories of mixed proteins. 6.5% of all variants belong to almost perfectly balanced proteins.



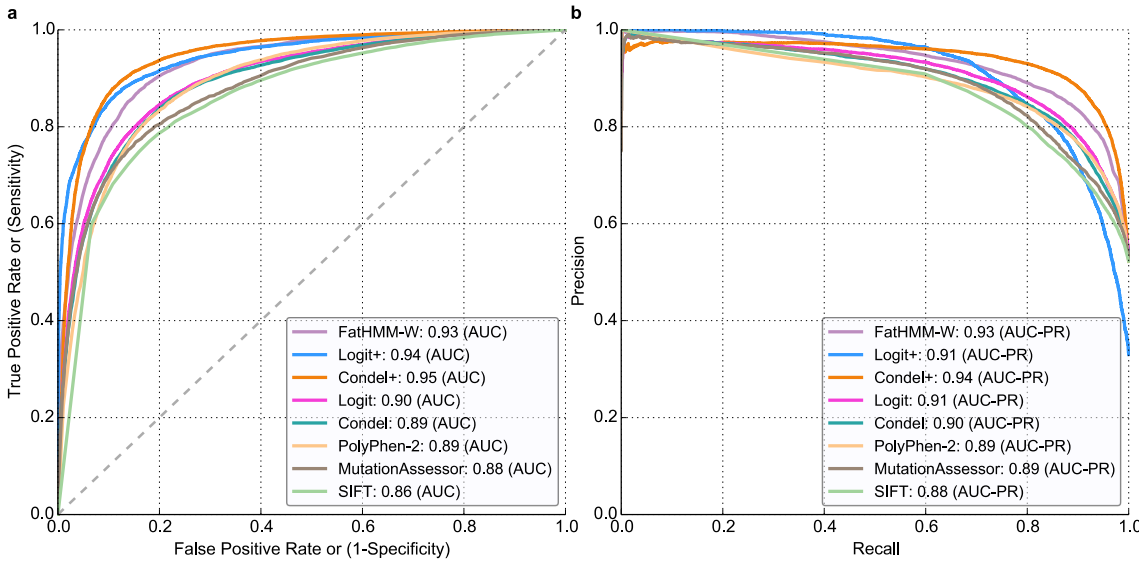
Supp. Figure S11. Evaluation of FatHMM-W according to pathogenic-to-neutral ratios. AUC of FatHMM-W over the five benchmark datasets. For each dataset, “All” denotes the whole dataset; “Pure” denotes the subset that is composed of variants appearing in pure proteins only; the open interval]0.0, 1.0[denotes the subset composed of all mixed proteins; [x,y] indicates the subset composed of all mixed proteins with a ratio of pathogenic-to-neutral variants greater than or equal to x and lower than or equal to y. Hatched bars indicate potentially biased results, due to the overlap (or suspected overlap) between the evaluation data and the data used for training FatHMM-W. Dotted bars indicate type 2 circularity bias. FatHMM-W performs poorly on mixed proteins, even on the data on which it has been trained.



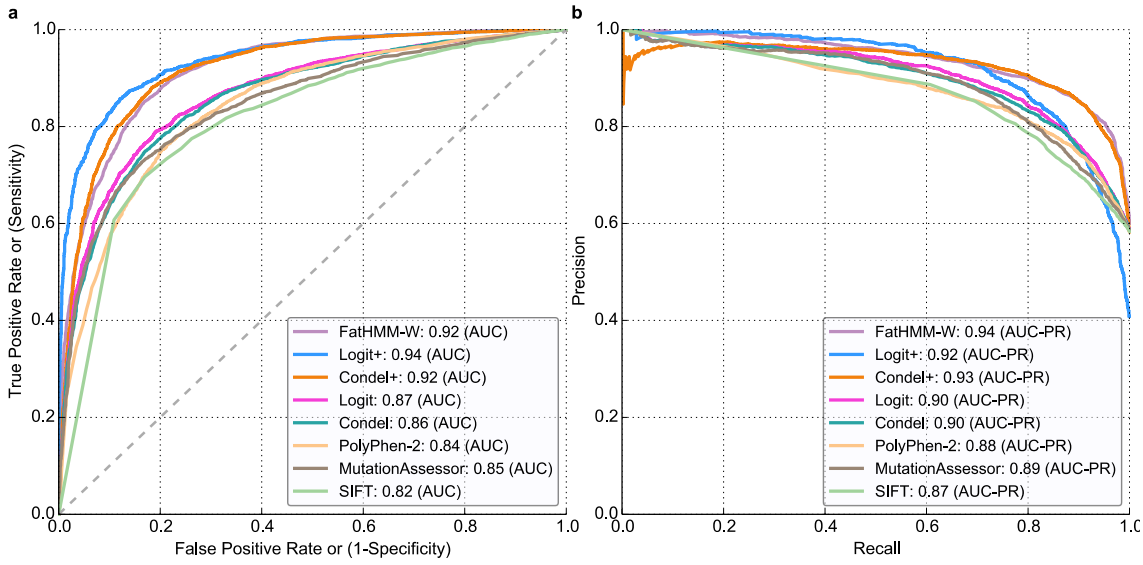
Supp. Figure S12. Performance of ten pathogenicity prediction tools according to protein pathogenic-to-neutral variant ratio in mixed proteins only. Evaluation of tool performance on subsets of *VariBenchSelected*, *predictSNPSelected* and *SwissVarSelected* defined according to the relative proportions of pathogenic and neutral variants in the proteins they contain. [x, y] indicate variants belonging to mixed proteins, containing a ratio of pathogenic-to-neutral variants between x and y.]0.0, 1.0[therefore indicates all mixed proteins (the ratios of 0.0 and 1.0 being excluded by the reversed brackets). While FatHMM-W performs well or excellently on variants belonging to pure proteins, it performs poorly on those belonging to mixed proteins.



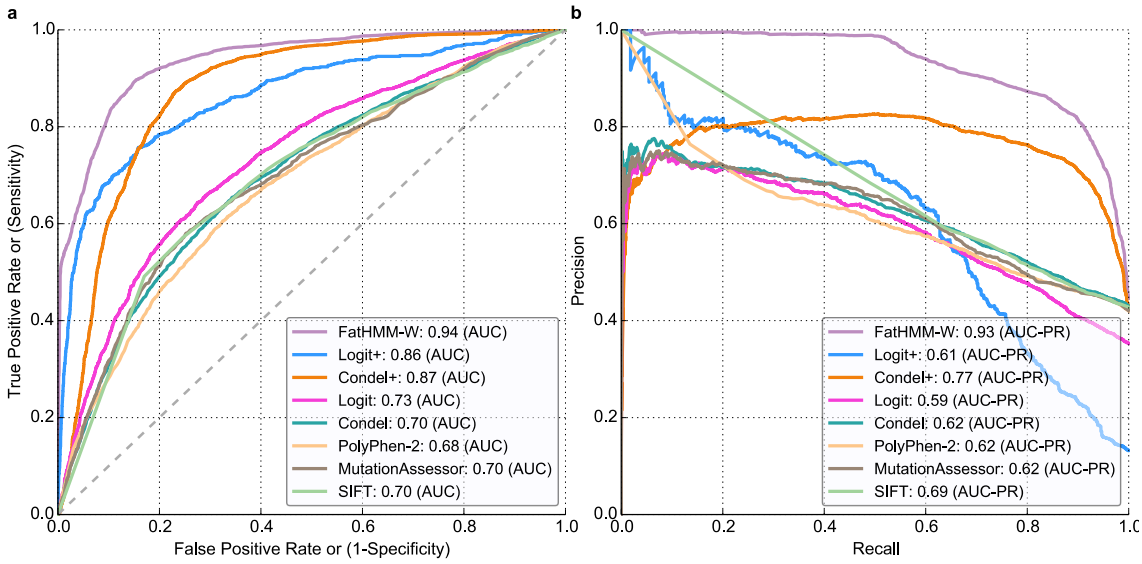
Supp. Figure S13. Variants from the *Selected* datasets that are in identical or similar proteins in the proxy training datasets *HumVar/ExoVar*. Percentage of pathogenic and neutral variants that can be found in identical or similar proteins in *HumVar/ExoVar*. The x-axes show different similarities between the proteins in the *Selected* dataset and the proteins in *HumVar/ExoVar*. The y-axis is the percentage of variants that can be found in identical or similar proteins.



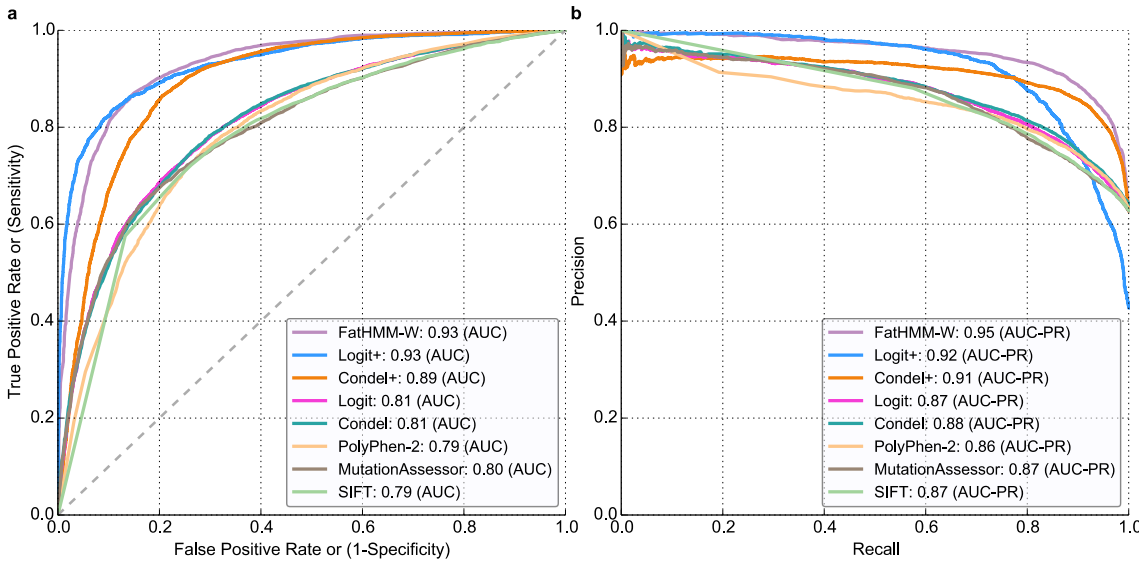
Supp. Figure S14. Performance of meta-predictors on *HumVar*. (a) ROC and (b) ROC-PR curves of FatHMM-W, MASS, PP2 and SIFT compared to those of the two meta-predictors Logit+ and Condel+ which combine them. Note that this evaluation is biased by both type 1 and type 2 circularity.



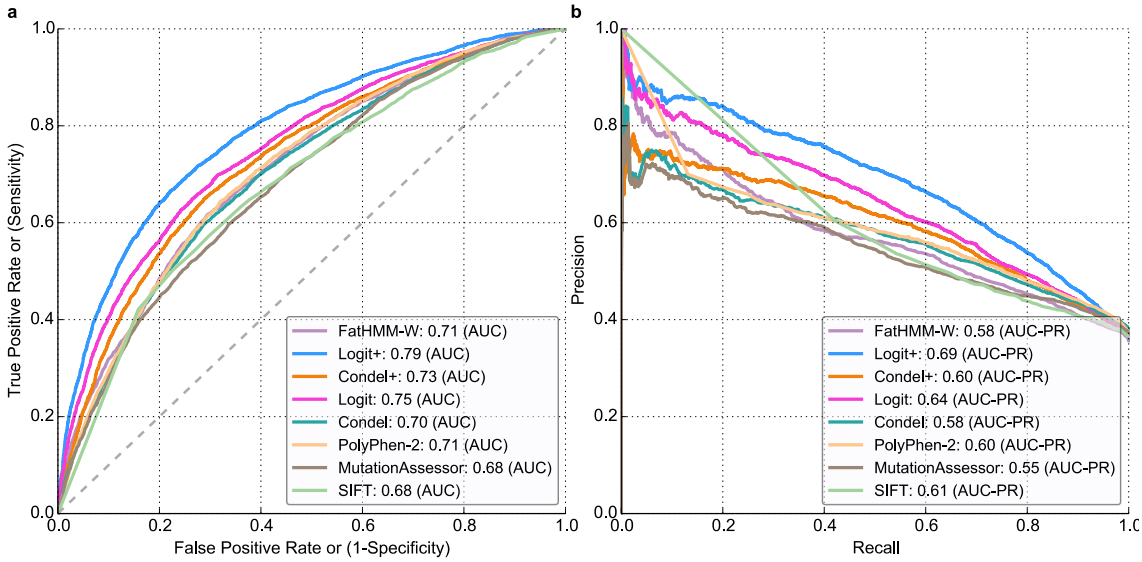
Supp. Figure S15. Performance of meta-predictos on *ExoVar*. (a) ROC and (b) ROC-PR curves of FatHMM-W, MASS, PP2 and SIFT compared to those of the two meta-predictors Logit+ and Condel+ which combine them. Note that this evaluation is biased by both type 1 and type 2 circularity.



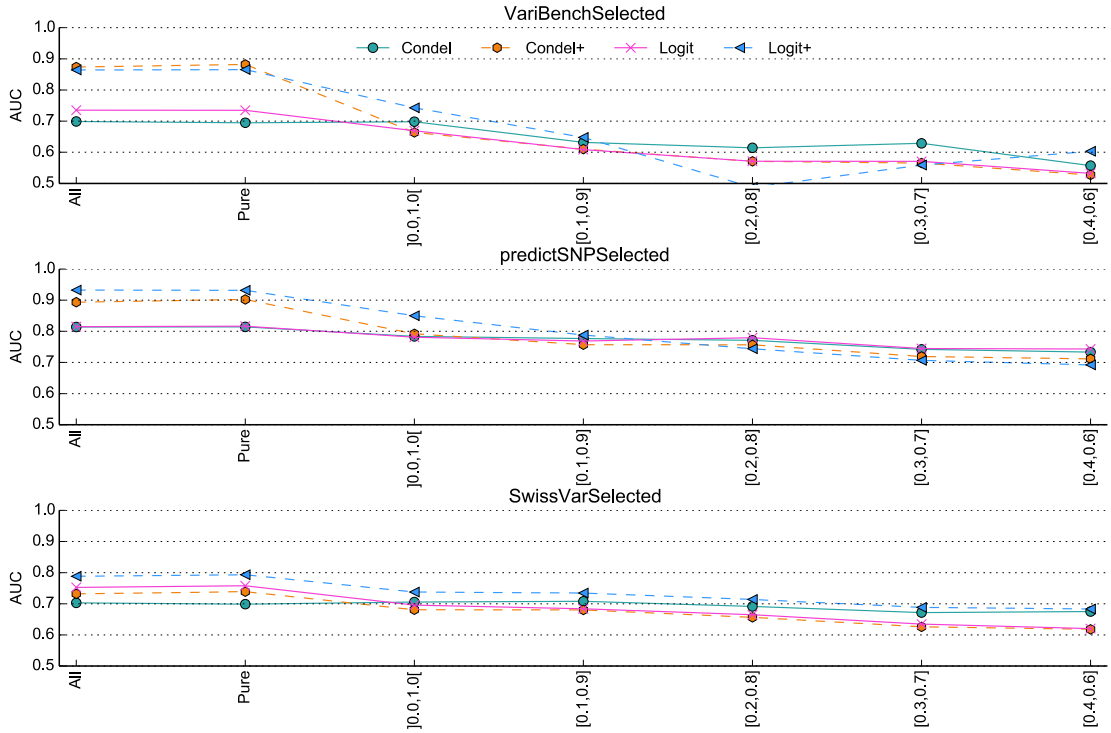
Supp. Figure S16. Performance of meta-predictors on *VariBenchSelected*. (a) ROC and (b) ROC-PR curves of FatHMM-W, MASS, PP2 and SIFT compared to those of the two meta-predictors Logit+ and Condel+ which combine them. FatHMM-W clearly outperforms both Logit+ and Condel+. Note that this evaluation is biased by both type 1 and type 2 circularity.



Supp. Figure S17. Performance of meta-predictors on *predictSNPSelected*. (a) ROC and (b) ROC-PR curves of FatHMM-W, MASS, PP2 and SIFT compared to those of the two meta-predictors Logit+ and Condel+ which combine them. FatHMM-W is on par with Logit+ and outperforms and Condel+. Note that this evaluation is biased by type 2 circularity.



Supp. Figure S18. Performance of meta-predictors on *SwissVarSelected*. (a) ROC and (b) ROC-PR curves of FatHMM-W, MASS, PP2 and SIFT compared to those of the two meta-predictors Logit+ and Condel+ which combine them.



Supp. Figure S19. Performance of 4 meta-predictors according to protein pathogenic-to-neutral variant ratio. Evaluation of Condel, Condel+, Logit and Logit+ on subsets of ***VariBenchSelected***, ***predictSNPSelected*** **and** ***SwissVarSelected*** defined according to the relative proportions of pathogenic and neutral variants in the proteins they contain. “Pure” indicates variants belonging to proteins containing only one class of variant. [x, y] indicate variants belonging to mixed proteins, containing a ratio of pathogenic-to-neutral variants between x and y.]0.0, 1.0[therefore indicates all mixed proteins (the ratios of 0.0 and 1.0 being excluded by the reversed brackets). While Logit+ and Condel+, which include FatHMM-W, perform well or excellently on variants belonging to pure proteins, they perform poorly on those belonging to mixed proteins. Logit performs better on pure proteins than Condel, whereas on the mixed proteins Condel is better or on par with Logit.

Supplementary Text S1

Brief summary about FatHMM

In its unweighted version, FatHMM-U [1] scores each variant by the log odds ratio of wild-type (P_w) to mutation amino acid (P_m), where the probabilities of observing each version of the amino acid are determined by an HMM-based multiple-sequence alignment against UniRef90 [2] sequences. The score for FatHMM-U is computed as follows:

$$\text{FatHMM-U} = \ln \frac{\frac{P_m}{1 - P_m}}{\frac{P_w}{1 - P_w}} = \ln \frac{P_m(1 - P_w)}{P_w(1 - P_m)} \qquad (1)$$

Essentially, FatHMM assumes that the more conserved the position at which the mutation occurred, the more likely it is to be pathogenic.

The weighted version (FatHMM-W) also takes into account how tolerant to mutations the sequence is. The tolerance to mutation of a sequence is evaluated using its relative proportions of known neutral (W_n) versus known pathogenic (W_d) variants. For this purpose, the FatHMM-U score (Equation 1) is weighted by the relative frequency of benign variants in (UniPort [3]) and pathogenic variants from (HGMD [4]). The updated score for the weighted version of FatHMM-W is defined as:

$$\text{FatHMM-W} = \ln \frac{(1.0 - P_w)(W_n + 1.0)}{(1.0 - P_m)(W_d + 1.0)} \qquad (2)$$

References

1. Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GLA, et al. (2013) Predicting the Functional, Molecular, and Phenotypic Consequences of Amino Acid Substitutions using Hidden Markov Models. Hum Mutat 34: 57–65. doi:10.1002/humu.22225.
2. Suzek BE, Huang H, McGarvey P, Mazumder R, Wu CH (2007) UniRef: comprehensive and non-redundant UniProt reference clusters. Bioinformatics 23: 1282–1288. doi:10.1093/bioinformatics/btm098.
3. Magrane M, Consortium U (2011) UniProt Knowledgebase: a hub of integrated protein data. Database 2011: bar009–bar009. doi:10.1093/database/bar009.
4. Stenson PD, Mort M, Ball EV, Shaw K, Phillips AD, et al. (2014) The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. Hum Genet 133: 1–9.

Supplementary Data S1

Supp. Data can be downloaded at the VariBench website:

http://structure.bmc.lu.se/VariBench/GrimmDatasets.php

The code as well as the data is available at:

http://www.bsse.ethz.ch/mlcb/research/bioinformatics-and-computational-biology/pathogenicity-prediction.html

This archive contains the variants and labels for each dataset as well as all retrieved tool scores and predicted labels from each tool. This data can be found in the folder “ToolScores”. Additionally, we provide all Python scripts to reproduce all main Figures, Supp. Figures and Tables from this study. The command “python start.py”, in the root directory of the ZIP file, automatically executes all scripts and stores all Figures and Tables into the folder “Output”.