ANALYSIS OF MISSENSE MUTATIONS OF HOST INNATE IMMUNE RESPONSE PATHWAY PROTEINS

Table S1: Numeric presentation of the data summary

Variable type	Variable name	Sample_n	Missing_count	Per_of_missing
PANTHER	numeric	5558	1242	0.183
PhD-SNP	numeric	5558	1242	0.183
SNP&GO	numeric	5558	1242	0.183
PROVEAN	numeric	5426	1374	0.202
SIFT	numeric	5426	1374	0.202
Mutation Assessor	numeric	5426	1374	0.202
FATHMM	numeric	5297	1503	0.221
MutationTaster	numeric	5866	934	0.137
PolyPhen-2	numeric	5570	1230	0.181
Condel	numeric	4931	1869	0.275
Envision	numeric	2957	3843	0.565
CADD	numeric	4587	2213	0.325

The numeric form of the mutants in the innate immune response pathway proteins that were successfully classified and those that missed prediction in each of the tools used in the study.

Table S2: Number of predicted damaging and tolerated mutants identified by the individual tools

Tool name	Prediction	Number of predictions	Prediction	Number of predictions
CADD	Damaging	2612	Tolerated	1975
Condel	Damaging	1232	Tolerated	3699
Envision	Damaging	2529	Tolerated	428
FATHMM	Damaging	1397	Tolerated	3896
Mutation Assessor	Damaging	998	Tolerated	4227
MutationTaster	Damaging	2406	Tolerated	3460
PANTHER	Damaging	1463	Tolerated	2922
PhD-SNP	Damaging	1619	Tolerated	3939
PolyPhen2	Damaging	2891	Tolerated	2679
PROVEAN	Damaging	1709	Tolerated	3717
SIFT	Damaging	2462	Tolerated	2964
SNPs&GO	Damaging	813	Tolerated	4745

Table S3a: Pearman correlation matrix (scores of mutants of innate immune response proteins) between the in-silico prediction tools

Tools	PhD-SNP	SNP&GO	PROVEAN	SIFT	Mutation Assessor	FATHMM	MutationTaster	PolyPhen2	Condel	Envision	CADD
PANTHER	0.502	0.732	-0.621	-0.387	0.218	-0.504	0.279	0.190	0.392	-0.493	0.390
PhD-SNP		0.847	-0.637	-0.475	0.222	-0.627	0.371	0.527	0.516	-0.620	0.563
SNP&GO			-0.676	-0.427	0.219	-0.595	0.355	0.563	0.498	-0.611	0.575
PROVEAN				0.546	-0.254	0.640	-0.413	-0.503	-0.406	0.651	-0.485
SIFT					-0.195	0.463	-0.181	-0.479	-0.287	0.470	-0.513
Mutation Assessor						-0.234	0.044	0.152	0.142	-0.198	0.133
FATHMM							-0.330	-0.503	-0.528	0.603	-0.489
MutationTaster								0.219	0.138	-0.422	0.210
PolyPhen2									0.428	-0.592	0.712
Condel										-0.435	0.465
Envision											-0.542

Table S3b: Spearman correlation matrix (scores of mutants of innate immune response proteins) between the in-silico prediction tools

Tools	PhD-SNP	SNPs&GO	PROVEAN	SIFT	Mutation Assessor	FATHMM	MutationTaster	PolyPhen2	Condel	Envision	CADD
PANTHER	0.433	0.649	-0.570	-0.474	0.225	-0.433	0.214	0.431	0.075	-0.433	0.366
PhD-SNP		0.865	-0.613	-0.600	0.222	-0.585	0.325	0.558	0.458	-0.584	0.580
SNP&GO			-0.642	-0.617	0.216	-0.539	0.296	0.606	0.427	-0.571	0.615
PROVEAN				0.708	-0.279	0.616	-0.379	-0.590	-0.306	0.651	-0.521
SIFT					-0.235	0.600	-0.244	-0.652	-0.354	0.651	-0.636
Mutation Assessor						-0.238	0.053	0.174	0.086	-0.193	0.126
FATHMM							-0.295	-0.558	-0.444	0.588	-0.519
MutationTaster								0.214	0.124	-0.376	0.192
PolyPhen2									0.440	-0.655	0.741
Condel										-0.354	0.474
Envision											-0.576

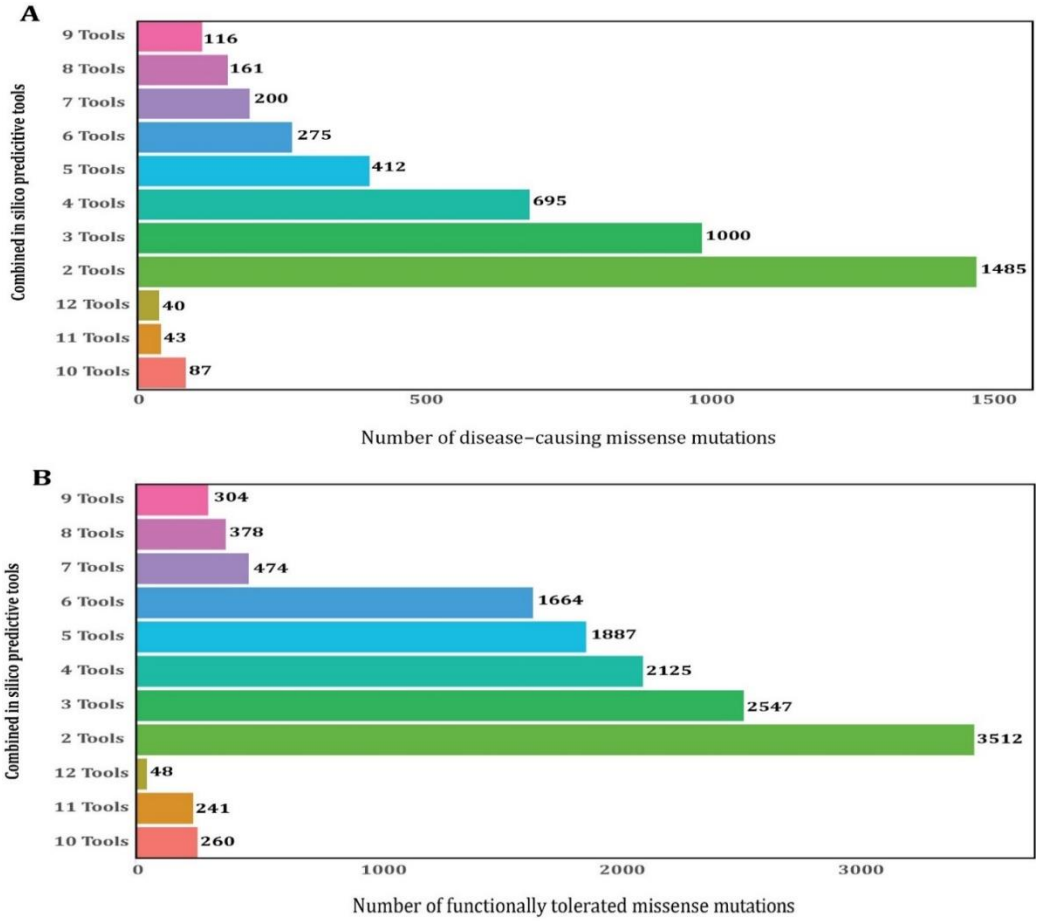

Figure S1a:

Screening the mutants in the innate immune pathway proteins to identify those highly pathogenic and with high potential of compromising protein's function. This was accomplished by combining output from multiple of the tools. Among the 6,802 mutants analyzed, combination of the outputs from all the 12 tools identified only 40 of the mutants as highly pathogenic

Figure S1b: Screening the mutants in the innate immune response pathway proteins to identify those least harmful (most tolerated) with protein's function. This was accomplished by combining output from multiple of the tools. Among the 6,802 mutants analyzed, a combination of outputs from all the 12 tools identified only 48 of the mutants as most tolerated.

ANALYSIS OF MISSENSE MUTATIONS OF SARS-COV-2 VRNA-HOST INTERACTION PROTEINS

Table S1: Numeric presentation of the data summary

Variable type	Variable name	Sample_n	Missing count	Per_of_missing
PANTHER	numeric	3964	638	0.139
PhD-SNP	numeric	3964	638	0.139
SNP&GO	numeric	3964	638	0.139
PROVEAN	numeric	3584	1018	0.221
SIFT	numeric	3584	1018	0.221
Mutation Assessor	numeric	3584	1018	0.221
FATHMM	numeric	3584	1018	0.221
MutationTaster	numeric	3839	763	0.166
PolyPhen-2	numeric	3283	1319	0.287
Condel	numeric	3292	1310	0.285
Envision	numeric	2775	1827	0.397
CADD	numeric	3875	727	0.158

The numeric form of the mutants in the SARS-CoV-2 vRNA-host interaction proteins that were successfully classified and those that missed prediction in each of the tools used in the study.

Table S2: Number of predicted damaging and tolerated mutants identified by the individual in silico tools

Tool name	Prediction	Number of predictions	Prediction	Number of predictions
CADD	Damaging	2971	Tolerated	904
Condel	Damaging	947	Tolerated	2345
Envision	Damaging	2545	Tolerated	230
FATHMM	Damaging	1047	Tolerated	2514
Mut Assessor	Damaging	727	Tolerated	2607
MutationTaster	Damaging	3101	Tolerated	668
PANTHER	Damaging	1072	Tolerated	2184
PhD-SNP	Damaging	1283	Tolerated	2681
PolyPhen2	Damaging	1726	Tolerated	1557
PROVEAN	Damaging	1363	Tolerated	2221
SIFT	Damaging	1679	Tolerated	1905
SNPs&GO	Damaging	554	Tolerated	3410

Table S3a: Pearson correlation matrix (scores of mutants of SARS-CoV-2 vRNA-host interaction proteins) between the in-silico tools

	PhD-SNP	SNP&GO	PROVEAN	SIFT	Mutation Assessor	FATHMM	MutationTaster	PolyPhen2	Condel	Envision	CADD
PANTHER	0.492	0.727	-0.549	-0.344	0.454	-0.419	0.321	0.425	0.392	-0.466	0.400
PhD-SNP		0.869	-0.628	-0.521	0.660	-0.550	0.366	0.475	0.558	-0.552	0.523
SNPs&GO			-0.660	-0.431	0.625	-0.537	0.381	0.462	0.565	-0.547	0.480
PROVEAN				0.514	-0.667	0.566	-0.455	-0.518	-0.568	0.563	-0.543
SIFT					-0.557	0.441	-0.213	-0.483	-0.347	0.449	-0.601
Mutation Assessor						-0.586	0.218	0.541	0.663	-0.559	0.537
FATHMM							-0.388	-0.447	-0.498	0.504	-0.449
MutationTaster								0.225	0.214	-0.436	0.264
PolyPhen2									0.348	-0.576	0.602
Condel										-0.491	0.386
Envision											-0.519

Table S3b: Spearman correlation matrix (scores of mutants of SARS-CoV-2 vRNA-host interaction proteins) between the in-silico tools

	PhD-SNP	SNP&GO	PROVEAN	SIFT	Mutation Assessor	FATHMM	MutationTaster	PolyPhen2	Condel	Envision	CADD
PANTHER	0.435	0.673	-0.507	-0.429	0.414	-0.386	0.225	0.450	0.272	-0.424	0.433
PhD-SNP		0.891	-0.630	-0.646	0.640	-0.556	0.297	0.514	0.461	-0.528	0.567
SNPs&GO			-0.649	-0.611	0.623	-0.527	0.303	0.522	0.432	-0.522	0.570
PROVEAN				0.667	-0.678	0.583	-0.373	-0.593	-0.504	0.561	-0.638
SIFT					-0.622	0.593	-0.276	-0.602	-0.409	0.589	-0.688
Mutation Assessor						-0.596	0.168	0.584	0.551	-0.531	0.560
FATHMM							-0.322	-0.526	-0.454	0.517	-0.558
MutationTaster								0.216	0.166	-0.347	0.255
PolyPhen2									0.353	-0.650	0.682
Condel										-0.398	0.398
Envision											-0.581

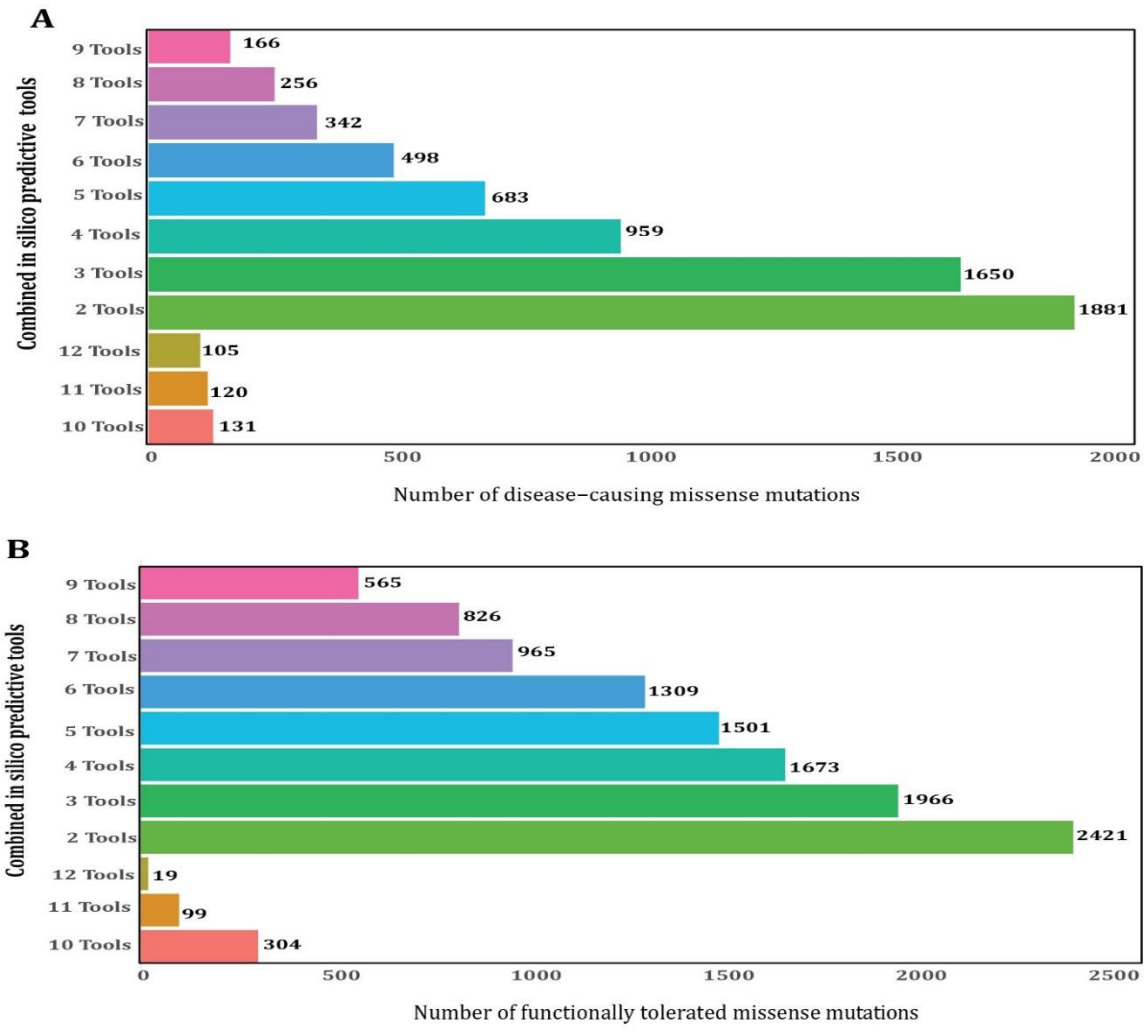

Figure S1A: Screening the mutants in the individual proteins to identify those highly pathogenic and with high potential of compromising protein's function. This was accomplished by combining output of multiple of the tools. Among the 4,600 mutants analyzed, combination of the outputs from all the 12-tool identified only 105 of the mutants as highly pathogenic.

Figure S1B: Screening the mutants in the individual proteins to identify those with least harmful impact (most tolerated) on protein's function. This was accomplished by combining output of multiple of the tools.