Comparative Analysis of Tools for Predicting the Functional Impact of mtDNA Variants

Madeline P. Griffin and Catherine E. Welsh

Rhodes College — 1848 —

INTRODUCTION

The development of a number of functional impact predictors and annotation tools has been implemented to aid in DNA variant analysis. Many conventional prediction tools, such as SIFT [3] and PolyPhen2 [1], are built for nuclear DNA, and fail to produce accurate predictions for mitochondrial mutations. Newer tools such as APOGEE [2] have addressed the need to annotate mtDNA exonic variants with high-confidence, but are limited when it comes to variants in mitochondrial tRNA and rRNA, which are both common causes of mitochondrial disease. A few papers, such as MitoTIP [4], address the need for tools that predict the pathogenicity of tRNA variants, while no known tools exist for annotating rRNA variants. Our purpose for this experiment is to provide an extensive review of available variant annotation tools that can be used to accurately predict the functional consequence of mtDNA variants.

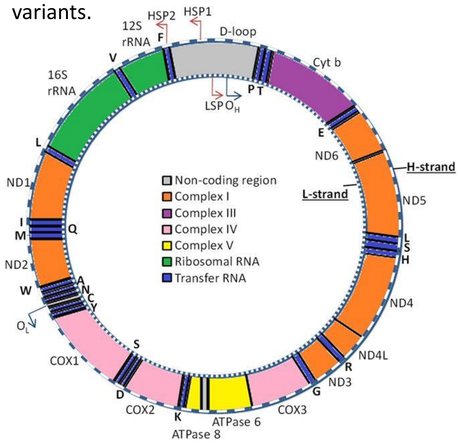


Figure 1 – Mitochondrial DNA –The circular 16,569 bp human mitochondrial genome encodes 13 genes, 22 tRNA and 2 rRNA. The variants we tested are found in Complexes I, III, IV, and V, which together form the coding region of mitochondrial DNA.

MATERIALS & METHODS

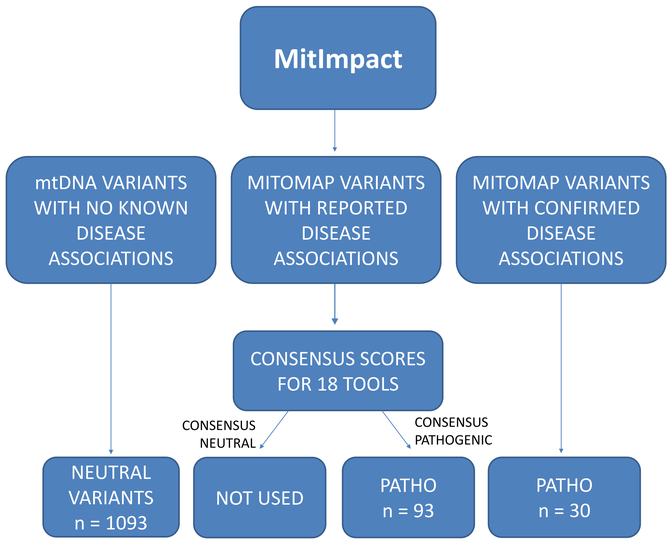
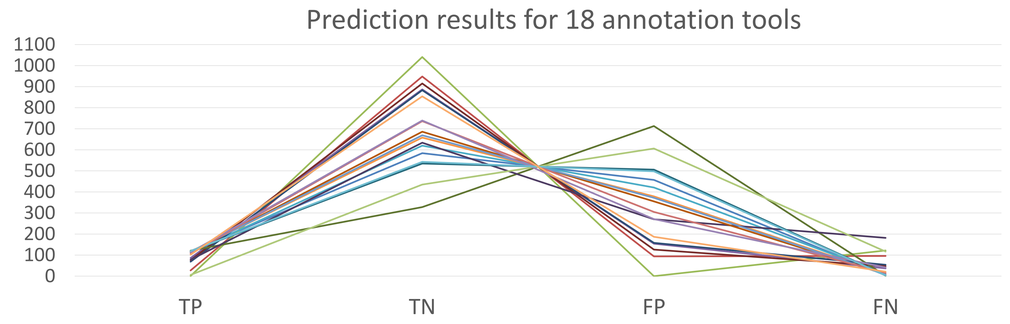
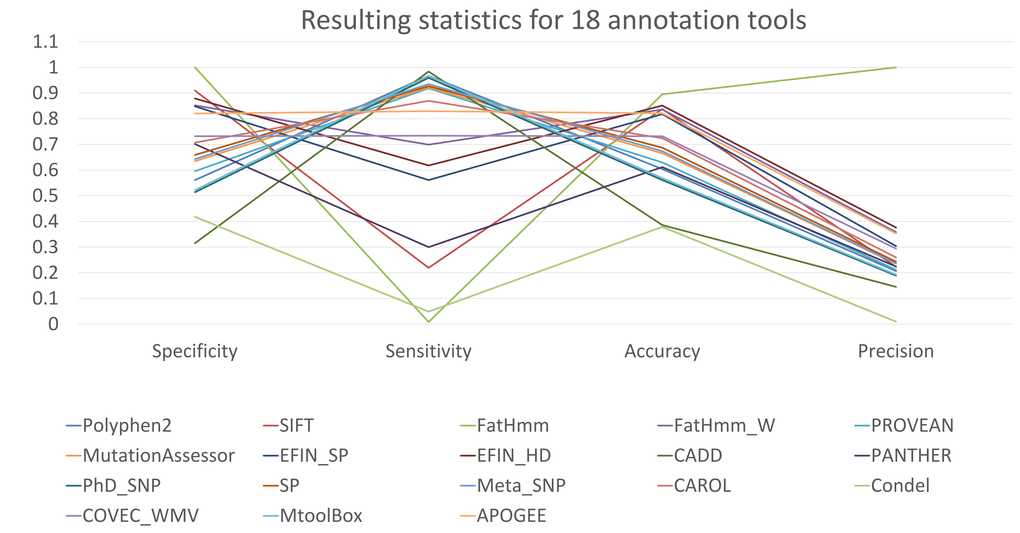


Figure 2 – Curated Test Set of mtDNA variants – Mitochondrial variants with reported, confirmed, and no known disease associations were acquired from MitImpact 2.9[5]. The pathogenic consensus scores from the 18 annotation tools were crosschecked with the reported disease associated variants and combined with the confirmed disease associations and neutral variants with no disease associations. We analyzed the performance of each tool based off of the resulting true or false predictions using a test set of size 1216 mtDNA variants.

Tools	TP	TN	FP	FN	Specificity TN/(FP+TN)	Sensitivity TP/(TP+FN)	Accuracy (TP+TN)/(P+N)	Precision TP/(TP+FP)	FDR FP/(TP+FP)	MCC	MCR
PolyPhen2	119	584	457	4	0.560999	0.96748	0.603952	0.206597	0.793403	0.29413	0.098153
SIFT	27	948	94	96	0.909789	0.219512	0.837629	0.22314	0.77686	0.045881	0.000841
FatHmm	1	1041	0	122	1	0.00813	0.895189	1	0	0.002643	0
FatHmm_W	86	887	154	37	0.852065	0.699187	0.835911	0.358333	0.641667	0.208035	0.003576
PROVEAN	113	620	421	10	0.595581	0.918699	0.629725	0.21161	0.78839	0.270804	0.036168
MutationAssessor	117	658	379	10	0.634523	0.92126	0.665808	0.235887	0.764113	0.280322	0.03256
EFIN_SP	69	883	158	54	0.848223	0.560976	0.817869	0.303965	0.696035	0.155033	0.002514
EFIN_HD	76	915	126	47	0.878963	0.617886	0.851375	0.376238	0.623762	0.182086	0.002303
CADD	121	328	713	2	0.315082	0.98374	0.385739	0.145084	0.854916	0.277786	0.306271
PANTHER	78	634	270	182	0.701327	0.3	0.611684	0.224138	0.775862	0.000844	0.001274
PhD-SNP	118	535	506	5	0.513929	0.95935	0.560997	0.189103	0.810897	0.285195	0.086942
SP	114	686	355	9	0.658982	0.926829	0.687285	0.24307	0.75693	0.281124	0.033887
MetaSNP	115	669	372	8	0.642651	0.934959	0.67354	0.23614	0.76386	0.283701	0.039948
CAROL	107	736	305	16	0.707012	0.869919	0.724227	0.259709	0.740291	0.259286	0.016377
Condel	6	435	606	117	0.417867	0.04878	0.378866	0.009804	0.990196	-0.38681	0.00445
COVEC WMV	113	739	271	41	0.731683	0.733766	0.731959	0.294271	0.705729	0.225178	0.005678
MToolBox	119	543	498	4	0.521614	0.96748	0.568729	0.192869	0.807131	0.290835	0.106959
APOGEE Bootstrap	102	854	187	21	0.820365	0.829268	0.821306	0.352941	0.647059	0.254164	0.00765

Table 1 – Computed statistics for 18 annotation tools – We ran our test set against the available annotations of 18 software predictors, and determined the resulting number of true positive, true negative, false positive, and false negative predictions. We then calculated the specificity, sensitivity, accuracy, and precision of those predictions, as well as the False Discovery Rate, the Misclassification Rate, and Matthew's Correlation Coefficient.





RESULTS

The 18 annotation tools tested presented varying degrees of accuracy, precision, sensitivity, and specificity.

- FatHmm scored the highest number of true negatives, and the least amount of false negatives.
- CADD scored the highest number of true positives and the least amount of false negatives.
- FatHmm had the highest specificity, accuracy, precision, and the lowest False Discovery Rate as well as the lowest Misclassification Rate.
- Annotation tools that with best performance were FatHmm, FatHmm_W, EFIN_SP, EFIN_HD, and APOGEE.

DISCUSSION

Our results contradict the current standard for variant annotation that relies mostly on software predictors built specifically for the nuclear genome. Those particular annotators such as SIFT, PolyPhen, PANTHER, and PROVEAN were retrofitted for mitochondrial DNA, and, based on the results of both our tests and the experiments of others, fall short in a number of statistically significant areas. The APOGEE Bootstrap model also presents convincing results to support the idea that poor scoring annotation tools, such as PolyPhen2, are not capable of accurately predicting the pathogenicity of mitochondrial variants [2]. A number of annotations tools base their predictions on the results of other scoring predictors. For example, CAROL annotation software provides results based almost entirely off the combined predictions of SIFT and Polyphen2. FatHmm, EFIN, and APOGEE algorithms ultimately outperformed conventional annotation tools, and we recommend these tools for mitochondrial annotations and variant prediction.

With these results we intend to build our own annotation tool for mitochondrial DNA specifically, that provides pathogenicity predictions from the tools we have deemed best fit to provide accurate results. We would like to create a software predictor that addresses the prediction of tRNA and rRNA variants, as there are still a limited number of prediction tools that include tRNA variants, and almost none that include rRNA variants.

REFERENCES

[1] IA Adzhubei, S Schmidt, L Peshkin, VE Ramensky, A Gerasimova, P Bork, AS Kondrashov, and SR Sunyaev. 2010. A method and server for predicting damaging missense mutations. *Nature Methods* (April 2010). http://dx.doi.org/10.1038/ nmeth0410-248

[2] S Castellana, C Fusilli, G Mazzoccoli, T Biagini, D Capocefalo, M Carella, AL Vescovi, and T Mazza. 2017. High-confidence assessment of functional impact of human mitochondrial non-synonymous genome variations by APOGEE. *PLoS Computational Biology* (June 2017). http://dx.doi.org/10.1371/journal.pcbi.1005628

[3] P Kumar, S Henikoff, and PC Ng. 2009. Predicting the effects of coding nonsynonymous variants on protein function using the SIFT algorithm. *Nature Protocols* (June 2009). http://dx.doi.org/10.1038/nprot.2009.86

[4] S Sonney, J Leipzig, MT Lott, S Zhang, V Procaccio, DC Wallace, and N Sondheimer. 2017. Predicting the pathogenicity of novel variants in mitochondrial tRNA with MitoTIP. *PLoS Computational Biology* (Dec. 2017).http://dx.doi.org/10.1371/ jounral.pcbi.1005867

[5] http://mitimpact.css-mendel.it/