

MitoVisualize: A resource for analysis of variants in human mitochondrial RNAs and DNA

Nicole J. Lake^{1,2}, Lily Zhou¹, Jenny Xu¹, and Monkol Lek¹

¹Department of Genetics, Yale School of Medicine, New Haven, CT, USA

²Murdoch Children's Research Institute, Melbourne, Australia

ABSTRACT

We present MitoVisualize, a new tool for analysis of the human mitochondrial DNA (mtDNA). MitoVisualize enables visualization of: (1) the position and effect of variants in mitochondrial transfer RNA (tRNA) and ribosomal RNA (rRNA) secondary structures alongside curated variant annotations, (2) data across RNA structures, such as to show all positions with disease-associated variants or with post-transcriptional modifications, and (3) the position of a base, gene or region in the circular mtDNA map, such as to show the location of a large deletion. All visualizations can be easily downloaded as figures for reuse. MitoVisualize can be useful for anyone interested in exploring mtDNA variation, though is designed to facilitate mtDNA variant interpretation in particular. MitoVisualize can be accessed via <https://www.mitovisualize.org/>. The source code is available at https://github.com/leklab/mito_visualize/.

INTRODUCTION

The human mitochondrial DNA (mtDNA) is a small but essential part of our genome, and encodes for proteins as well as ribosomal RNAs (rRNA) and transfer RNAs (tRNA) (Anderson *et al.*, 1981). Pathogenic mtDNA variants cause a range of phenotypes collectively known as mitochondrial diseases (Gorman *et al.*, 2016). Mitochondrial variants are also involved in other diseases, such as neurological disorders and cancer, and in various traits related to health (Keogh and Chinnery, 2015;

Gammage and Frezza, 2019; Yonova-Doing *et al.*, 2021). However the effect of most mtDNA variants on mitochondrial function, and their role in health and disease, remains unknown.

Several tools have been developed for mtDNA variant analysis, such as MITOMASTER, HmtVar and mvTool, which provide annotations in tabular format (Lott *et al.*, 2013; Preste *et al.*, 2019; Shen *et al.*, 2018). MitImpact provides annotations for nonsynonymous variants, including a visualization in the protein structure, but does not cover RNAs (Castellana *et al.*, 2015). For RNA genes the position of a variant in the secondary structure, and its effect on base pairing, are important considerations when evaluating its impact (McFarland *et al.*, 2004). Analysis of data across the RNA structure, such as to show bases with pathogenic variants, is also useful (Pütz *et al.*, 2007). However tools for visualizing variants and data within RNA structures had not been developed; nor had tools for generating figures labelling bases and regions within the circular mtDNA map.

To address this we developed MitoVisualize, a new tool for visualization of variants and data in RNA secondary structures, and of loci in the circular mtDNA.

IMPLEMENTATION

Various curated data are displayed on MitoVisualize. The secondary structures and domain annotations for the 22 tRNAs are per Mamit-tRNA (Pütz *et al.*, 2007), and as published for the 2 rRNA secondary structures (Amunts *et al.*, 2015; Brown *et al.*, 2014). The mtDNA map is NCBI Reference Sequence NC_012920.1 (Andrews *et al.*, 1999). Annotations provided include (1) population frequency from gnomAD (Laricchia *et al.*, 2021), MITOMAP (Lott *et al.*, 2013) and HelixMTdb (Bolze *et al.*, 2020) databases, (2) maximum heteroplasmy from gnomAD and HelixMTdb, (3) tRNA *in silico* predictions from MitoTIP (Sonney *et al.*, 2017), PON-mt-tRNA (Niroula and Vihinen, 2016), and HmtVar (Preste *et al.*, 2019), (4) disease associations, and their

classification status, from MITOMAP and ClinVar (Landrum *et al.*, 2018), (5) associated haplogroups from Phylotree (van Oven and Kayser, 2009), (5) conservation scores from PhyloP and PhastCons (Pollard *et al.*, 2010), and (6) other data types such as bases with post-transcriptional modifications (Suzuki *et al.*, 2020; Rebelo-Guiomar *et al.*, 2019) or involved in formation of the tertiary structures (tRNA only, from HmtVar). Detailed information on data collection and display is in the Supplementary Table.

MitoVisualize was built using open source tools. On the server end it runs an Elasticsearch database with data populated from sources described above. Users and other applications can access this processed data directly via the GraphQL interface (<https://mitovisualize.org/api>). The web front-end was implemented in JavaScript using React libraries. Lastly, the RNA and mtDNA visualizations were custom designed in SVG format.

USING MITOVISUALIZE

A search bar for querying mitochondrial RNA variants and genes is located on the landing page, as well as in the navigation bar accessible from every page. The About page summarizes key functionality and data used to build MitoVisualize, and “?” icons throughout the site provide extra information about the various annotations. There are three key functions in MitoVisualize: (1) using the variant tool to analyze a RNA variant, (2) using the gene tool to visualize data across RNA structures, and (3) using the mtDNA tool to annotate a base, gene or region in the circular mtDNA.

The variant tool can be accessed by typing the RNA variant of interest into the search bar, or by selecting ‘Variant’ in the dropdown menu for tRNA and rRNA in the navigation bar. MitoVisualize currently only accepts single nucleotide variants (input format e.g. m.3243A>G). The position of the variant within the RNA secondary structure is shown in red text and yellow highlight; any change to

a base pair type is also displayed (Figure 1A). For rRNAs the region with the variant is shown zoomed-in, while its location in the larger structure is displayed in the corner inset. Various annotations are provided alongside the RNA structure visualization to help the user interpret the impact of the variant (see Supplementary Table). All annotations are hyperlinked for easy navigation to the data source.

The gene tool can be accessed by typing the RNA gene of interest into the search bar, or by selecting 'Gene' in the dropdown menu for tRNA and rRNA in the navigation bar. A user can select a data type of interest for visualization across the RNA structure; note the mtDNA coordinate of a base can be revealed by hovering over it. Quantitative data types (i.e. maximum heteroplasmy of base variants in population databases, Figure 1B) are visualized as a color gradient, while qualitative data (i.e. bases with disease-associated variants) are shown in red.

The mtDNA tool is accessible via the navigation bar and displays the circular mtDNA molecule. Protein-coding genes are colored blue, rRNA genes yellow, tRNA genes in green, and non-coding regions in red. A user can label the position of a gene, base/variant, or region within the mtDNA (Figure 1C). For regions, overlapping genes are also listed.

SUMMARY

MitoVisualize is a free, user-friendly tool for visualizing the position and effect of variants in human mitochondrial tRNA and rRNA secondary structures, and in the circular mtDNA map. It also enables visualization of data across the RNA secondary structures. All visualizations can be downloaded as png figures for reuse. By including the variant or gene name in the URL, we enable easy linking to MitoVisualize from other resources; users can also access data directly via API. Version and date of

access for all data sources are listed on the About page of MitoVisualize, and will be updated periodically.

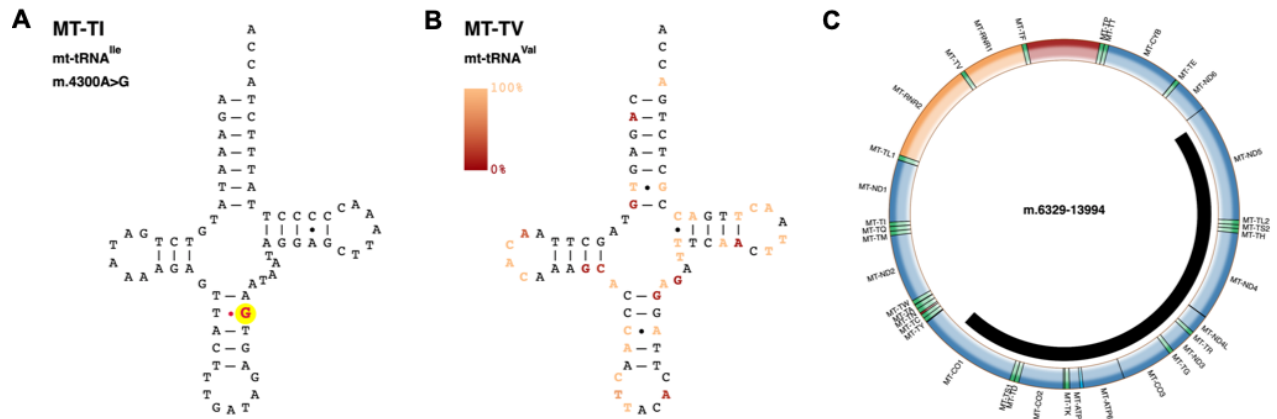


Figure 1: Example visualizations in MitoVisualize. (A) Using the variant tool to visualize the position and effect of m.4300A>G. (B) Using the gene tool to visualize maximum heteroplasmy in gnomAD across *MT-TV*. (C) Using the mtDNA tool to visualize a deletion spanning m.6329-13994.

ACKNOWLEDGEMENTS

We gratefully acknowledge the contribution of Tianxiao (Lisa) Wang who helped create the RNA SVG files.

FUNDING

This work was supported by the National Health and Medical Research Council [1159456 to N.J.L.].

Conflict of interest: None declared.

REFERENCES

- Amunts,A. *et al.* (2015) Ribosome. The structure of the human mitochondrial ribosome. *Science*, **348**, 95–98.
- Anderson,S. *et al.* (1981) Sequence and organization of the human mitochondrial genome. *Nature*, **290**, 457–465.
- Andrews,R.M. *et al.* (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.*, **23**, 147.
- Bolze,A. *et al.* (2020) A catalog of homoplasmic and heteroplasmic mitochondrial DNA variants in humans. *BioRxiv*.
- Brown,A. *et al.* (2014) Structure of the large ribosomal subunit from human mitochondria. *Science*, **346**, 718–722.
- Castellana,S. *et al.* (2015) MitImpact: an exhaustive collection of pre-computed pathogenicity predictions of human mitochondrial non-synonymous variants. *Hum. Mutat.*, **36**, E2413-22.
- Gammage,P.A. and Frezza,C. (2019) Mitochondrial DNA: the overlooked oncogenome? *BMC Biol.*, **17**, 53.
- Gorman,G.S. *et al.* (2016) Mitochondrial diseases. *Nat. Rev. Dis. Primers*, **2**, 16080.
- Keogh,M.J. and Chinnery,P.F. (2015) Mitochondrial DNA mutations in neurodegeneration. *Biochim. Biophys. Acta*, **1847**, 1401–1411.
- Landrum,M.J. *et al.* (2018) ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.*, **46**, D1062–D1067.
- Laricchia,K.M. *et al.* (2021) Mitochondrial DNA variation across 56,434 individuals in gnomAD. *BioRxiv*.
- Lott,M.T. *et al.* (2013) mtDNA Variation and Analysis Using MITOMAP and MITOMASTER. *Curr. Protoc. Bioinformatics*, **1**, 1.23.1-1.23.26.
- McFarland,R. *et al.* (2004) Assigning pathogenicity to mitochondrial tRNA mutations: when “definitely maybe” is not good enough. *Trends Genet.*, **20**, 591–596.

- Niroula,A. and Vihinen,M. (2016) PON-mt-tRNA: a multifactorial probability-based method for classification of mitochondrial tRNA variations. *Nucleic Acids Res.*, **44**, 2020–2027.
- van Oven,M. and Kayser,M. (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.*, **30**, E386-94.
- Pollard,K.S. *et al.* (2010) Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.*, **20**, 110–121.
- Prete,R. *et al.* (2019) HmtVar: a new resource for human mitochondrial variations and pathogenicity data. *Nucleic Acids Res.*, **47**, D1202–D1210.
- Pütz,J. *et al.* (2007) Mamit-tRNA, a database of mammalian mitochondrial tRNA primary and secondary structures. *RNA*, **13**, 1184–1190.
- Rebelo-Guiomar,P. *et al.* (2019) The mammalian mitochondrial epitranscriptome. *Biochim. Biophys. Acta Gene Regul. Mech.*, **1862**, 429–446.
- Shen,L. *et al.* (2018) MSeqDR mvTool: A mitochondrial DNA Web and API resource for comprehensive variant annotation, universal nomenclature collation, and reference genome conversion. *Hum. Mutat.*, **39**, 806–810.
- Sonney,S. *et al.* (2017) Predicting the pathogenicity of novel variants in mitochondrial tRNA with MitoTIP. *PLoS Comput. Biol.*, **13**, e1005867.
- Suzuki,Takeo *et al.* (2020) Complete chemical structures of human mitochondrial tRNAs. *Nat. Commun.*, **11**, 4269.
- Yonova-Doing,E. *et al.* (2021) An atlas of mitochondrial DNA genotype-phenotype associations in the UK Biobank. *Nat. Genet.*, **53**, 982–993.