

# MitoBenchStarter: An interactive visual workbench for population genetics on mitochondrial DNA



Judith Neukamm<sup>1,3\*</sup>, Alexander Peltzer<sup>1,2,3\*</sup>, Wolfgang Haak<sup>2</sup>, Johannes Krause<sup>1,2</sup> and Kay Nieselt<sup>3</sup>

<sup>1</sup> Institute for Archaeological Sciences, Archaeo- and Paleogenetics, University of Tuebingen, Germany.

<sup>2</sup> Max Planck Institute for the Science of Human History, Jena, Germany.

<sup>3</sup> Integrative Transcriptomics, Center for Bioinformatics (ZBIT), University of Tuebingen, Germany.

\*These authors contributed equally to the work.

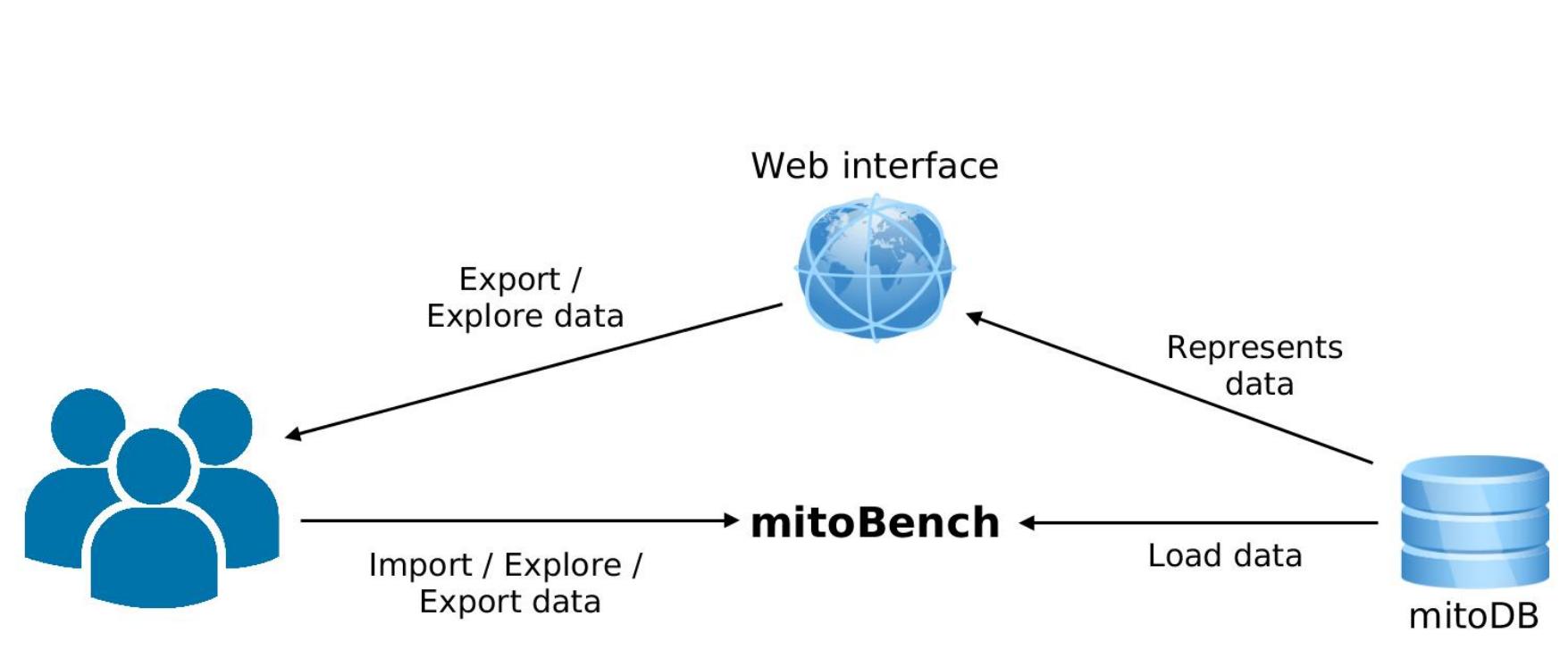


Contact: judith.neukamm@uni-tuebingen.de

## Motivation & Goals

- Especially in the research field of ancient human DNA, mitochondrial DNA (mtDNA) is often the only proxy available to study extinct populations and their relationship with modern populations
- Tools for the analysis typically rely on different file formats → requires manual interaction with the data for downstream analysis
- mitoBench: workbench to ease file conversions, interactively analyze human mitochondrial genomes and visualize results
- mitoDB: database for mitochondrial reference data to provide a central reference database that can be easily accessed via the workbench

## Workflow



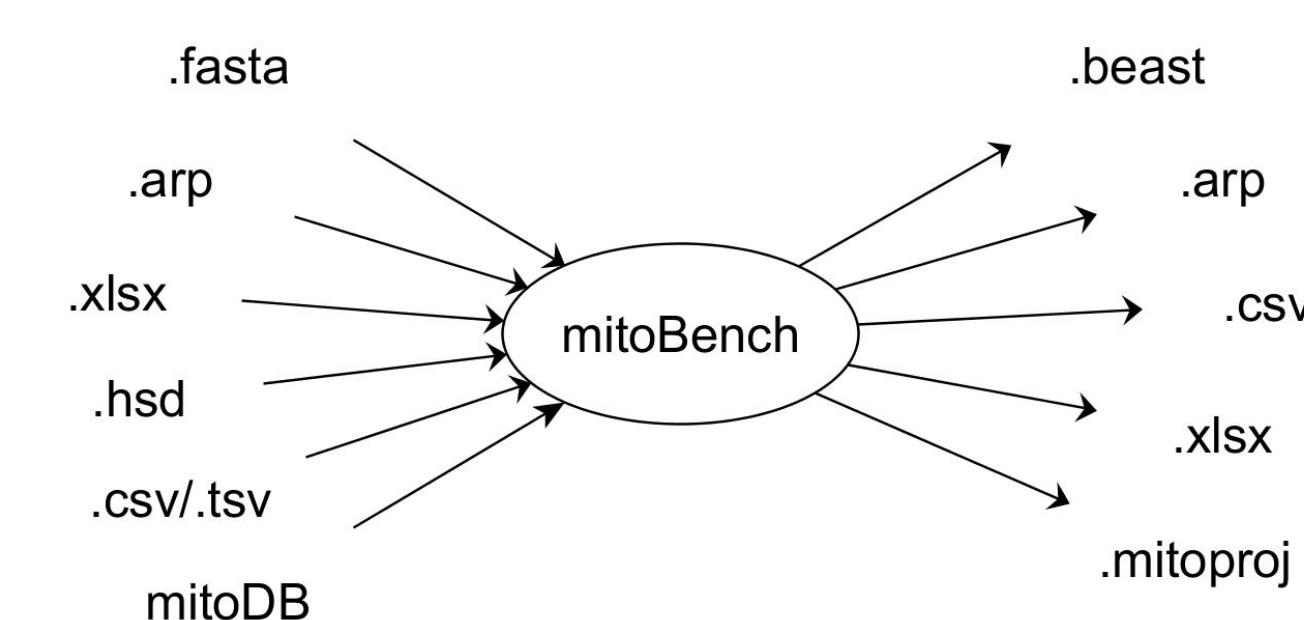
## mitoDB

- Web-Frontend with Vaadin Java Framework
- Backend with PostgreSQL, providing sequence information and meta-data
- Curated data upload with meta-data
- Retrieval possible through WebUI and mitoBench



## mitoBench

### File conversions



⇒ connect the workbench with existing analysis methods such as BEAST, Arlequin, PhyloTree and others

### Data representation as Table

ID	MTSequence	Haplogroup	Time Period	continent	country	country_region	culture_type	dating_lower
I2380	GATCA..	T1a2	Pre-Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient	-469
egyp5.5A129	GATCA..	X1c	Egypt Modern	Africa	Egypt	Other	Modern	2015
egyp5.6A9	GATCA..	H3w	Egypt Modern	Africa	Egypt	Other	Modern	2015
egyp5.3058	GATCA..	L37a1	Ethiopia Modern	Africa	Ethiopia	Other	Modern	2015
egyp5.3A63	GATCA..	R0a3	Egypt Modern	Africa	Egypt	Other	Modern	2015
egyp5.30583	GATCA..	M1a5	Ethiopia Modern	Africa	Ethiopia	Other	Modern	2015
egyp5.30586	GATCA..	R0a2	Ethiopia Modern	Africa	Ethiopia	Other	Modern	2015
egyp5.30587	GATCA..	R0a2g	Ethiopia Modern	Africa	Ethiopia	Other	Modern	2015
egyp5.30602	GATCA..	L3	Ethiopia Modern	Africa	Ethiopia	Other	Modern	2015
egyp5.30603	GATCA..	L241d1	Ethiopia Modern	Africa	Ethiopia	Other	Modern	2015
egyp5.306030	GATCA..	I6	Egypt Modern	Africa	Egypt	Other	Modern	2015
I2379	GATCA..	U3b	Roman Period	Africa	Egypt	Beni Suef	Ancient	2015
I2378	GATCA..	T1a7	Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient	-235
egyp5.8A134	GATCA..	H20	Egypt Modern	Africa	Egypt	Other	Modern	2015

### Data grouping

- group data based on shared feature (e.g. time period, location)
- user defined grouping

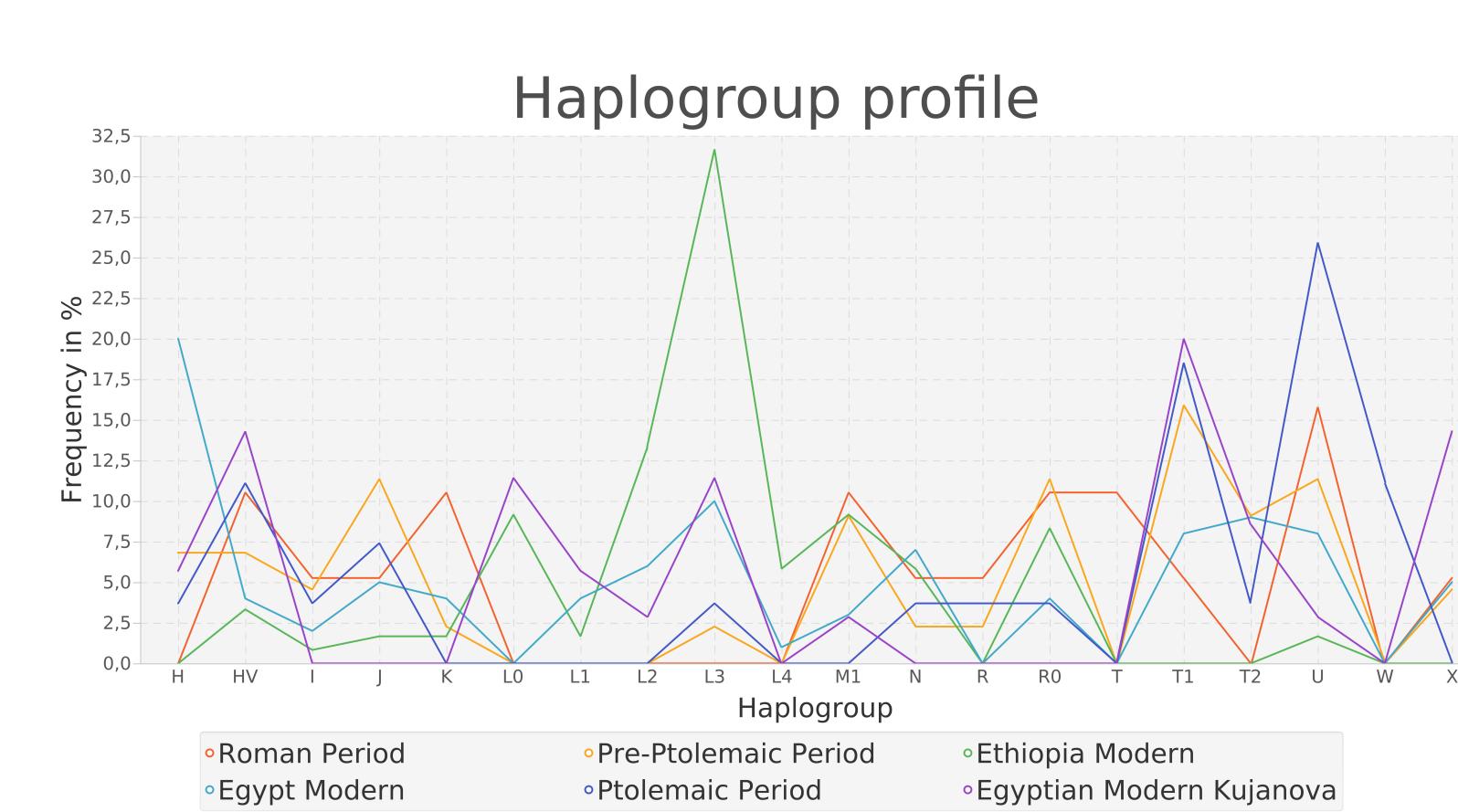
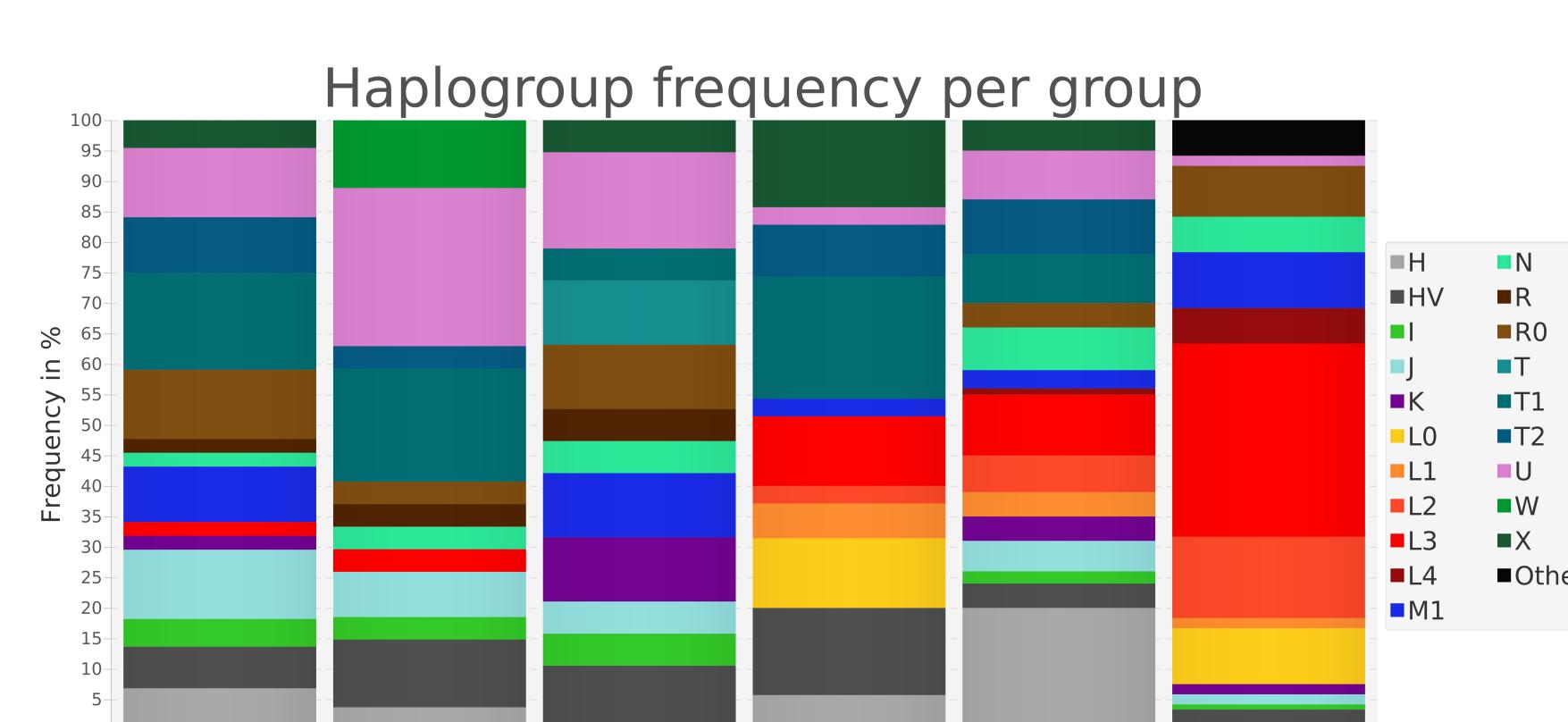
⇒ allows analysis of Haplogroup distribution between different groups

### Data filtering / Statistics

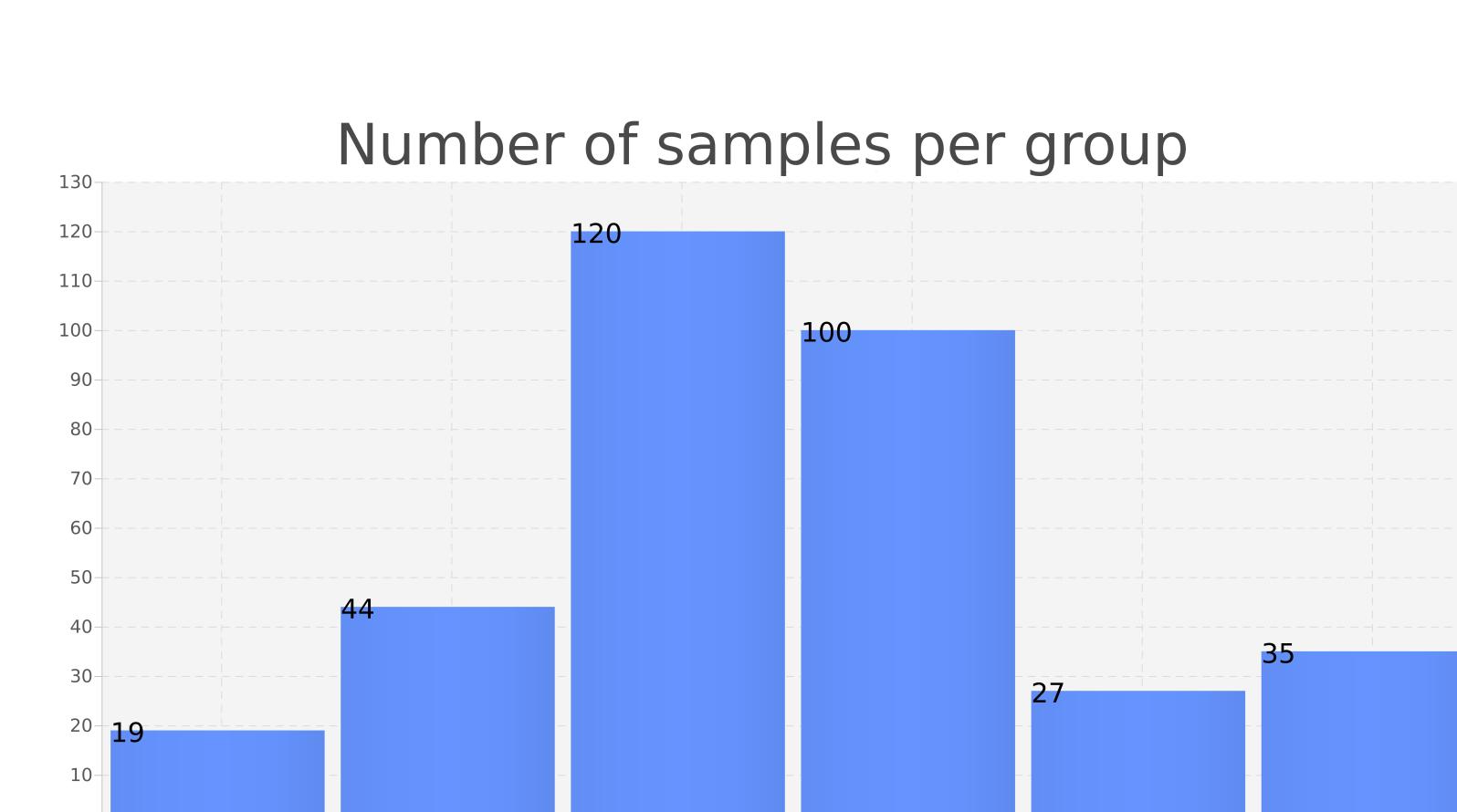
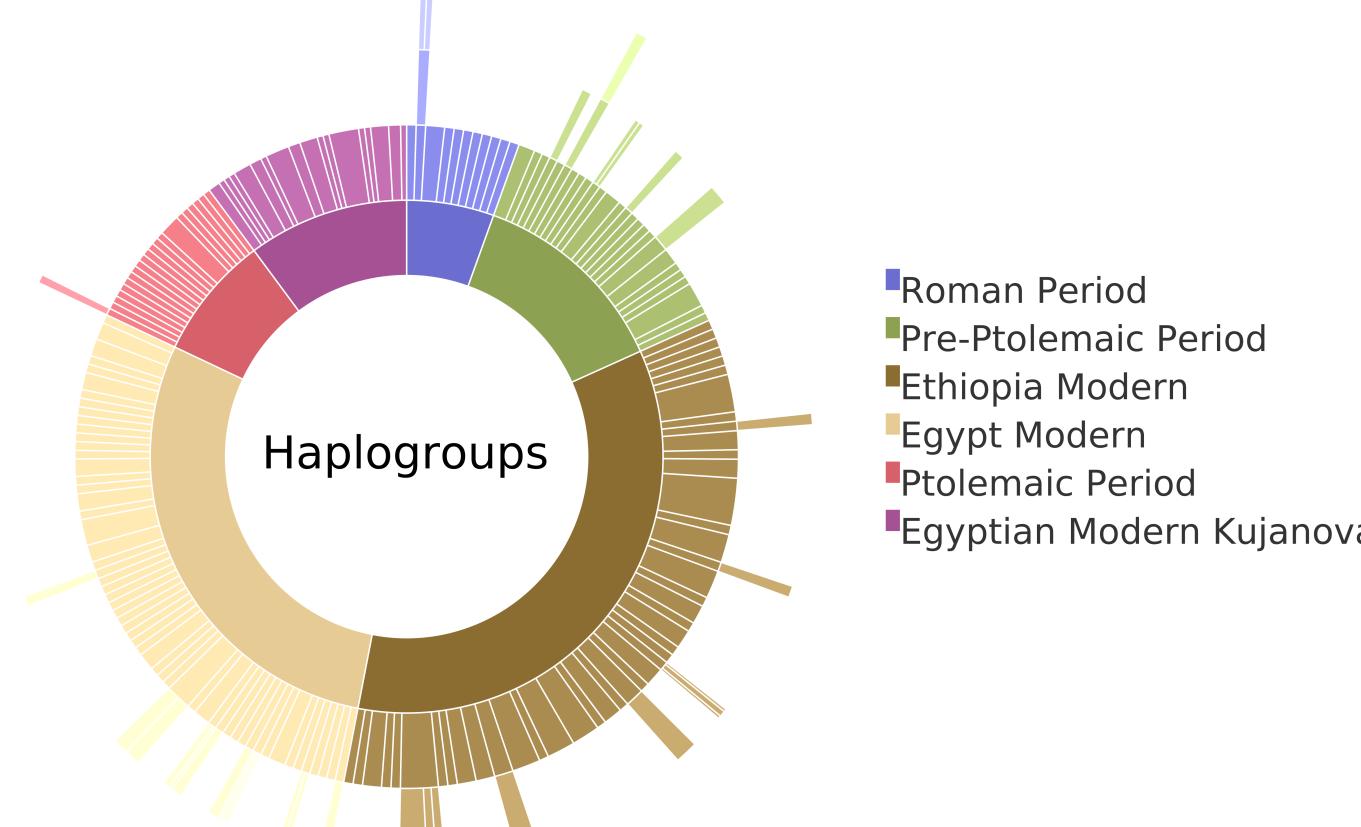
- Haplogroup filtering / frequencies
- Mutation filtering / frequencies

Haplogroups based on **PhyloTree<sup>mt</sup>**

### Data visualizations



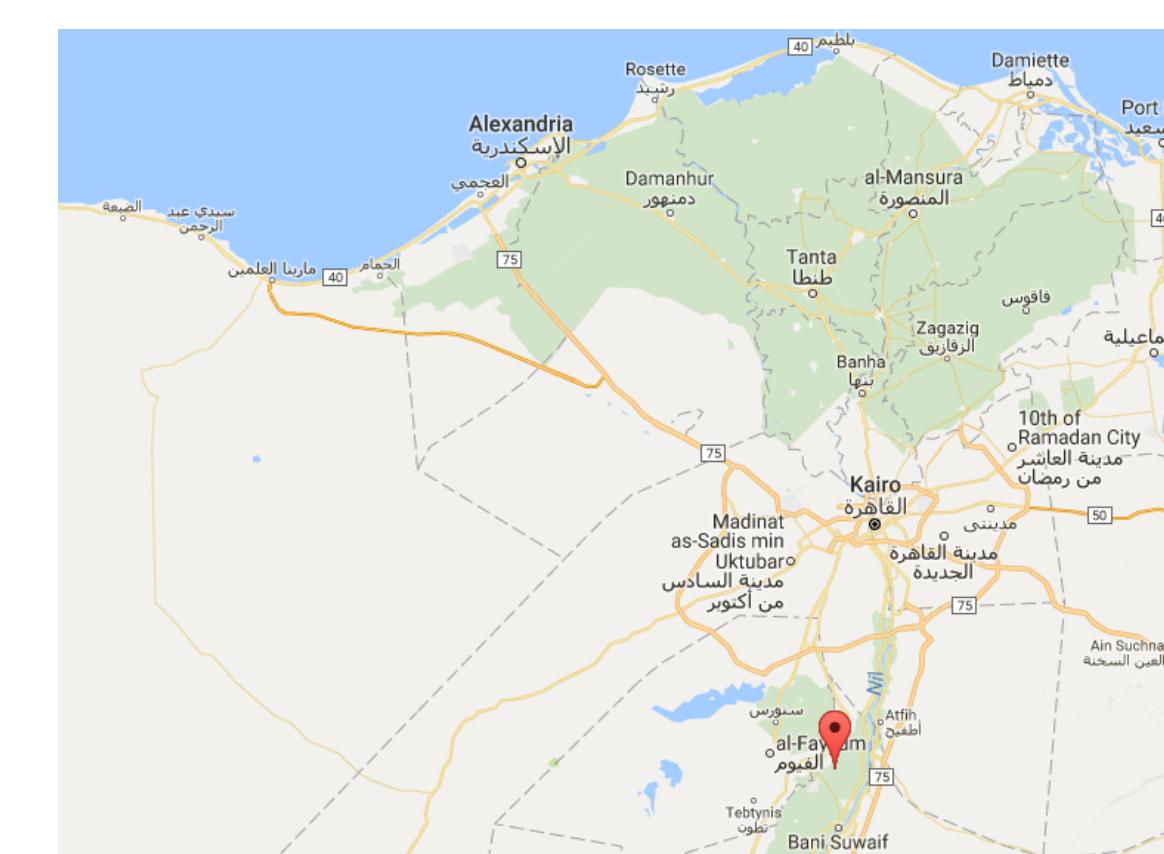
### Haplogroup hierarchy of dataset



## Outlook

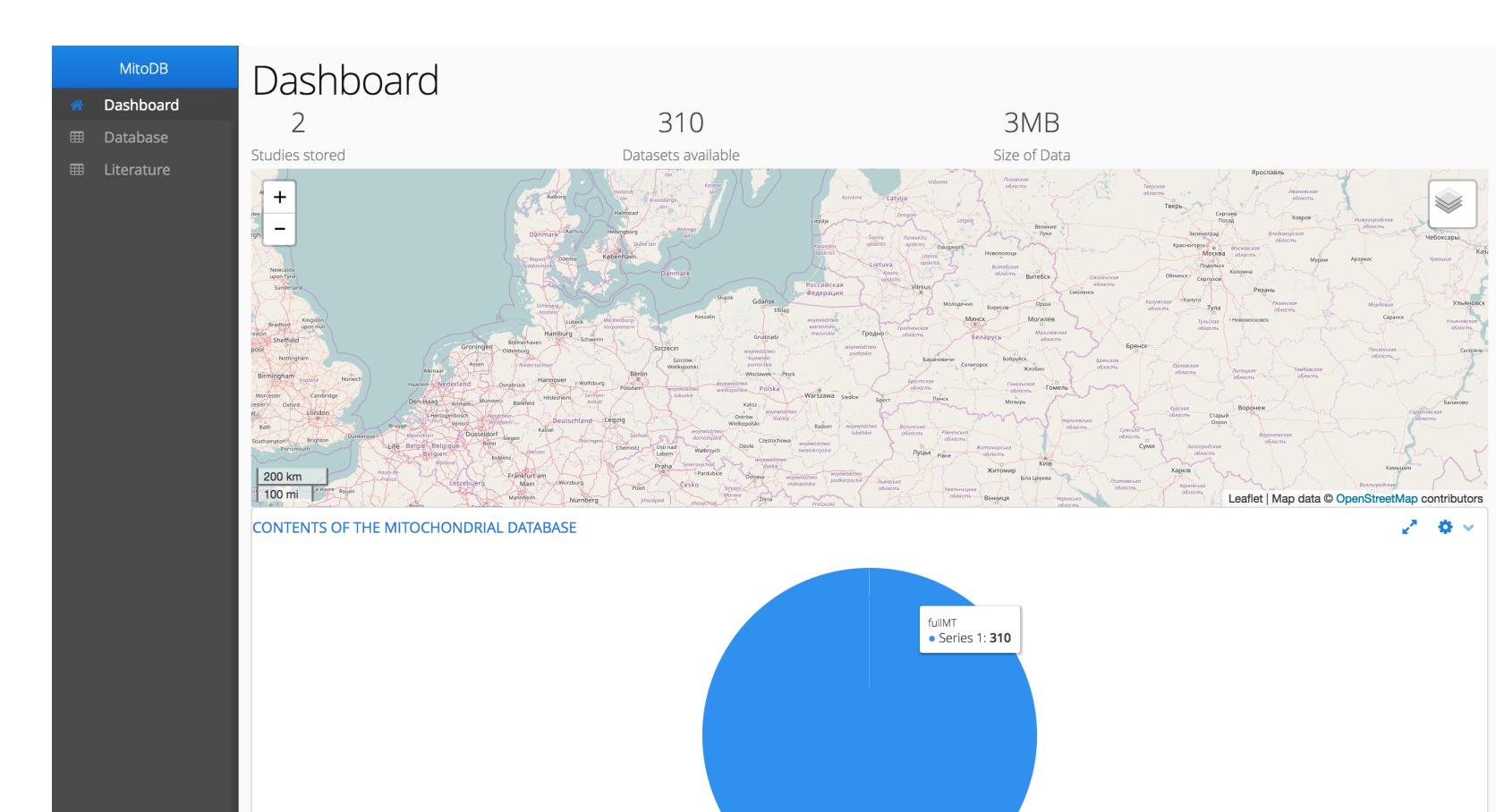
### mitoBench

- Provide methods for downstream analysis (e.g.  $F_{ST}$  calculations, Founder Analysis)
- Offer more visualizations (e.g. geographical maps)



### mitoDB & web interface

- Export/import functionality
- web-based dashboard to explore database



## References

- Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular biology and evolution*, 29(8), 1969-1973.
- Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources*, 10(3), 564-567.
- Van Oven, M., & Kayser, M. (2009). Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Human mutation*, 30(2), E386-E394.
- Weissensteiner, H., Pacher, D., Kloss-Brandstätter, A., Forer, L., Specht, G., Bandelt, H. J., ... & Schönherz, S. (2016). Haplotype 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic acids research*, gkw233.