

# mitoBench: 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.

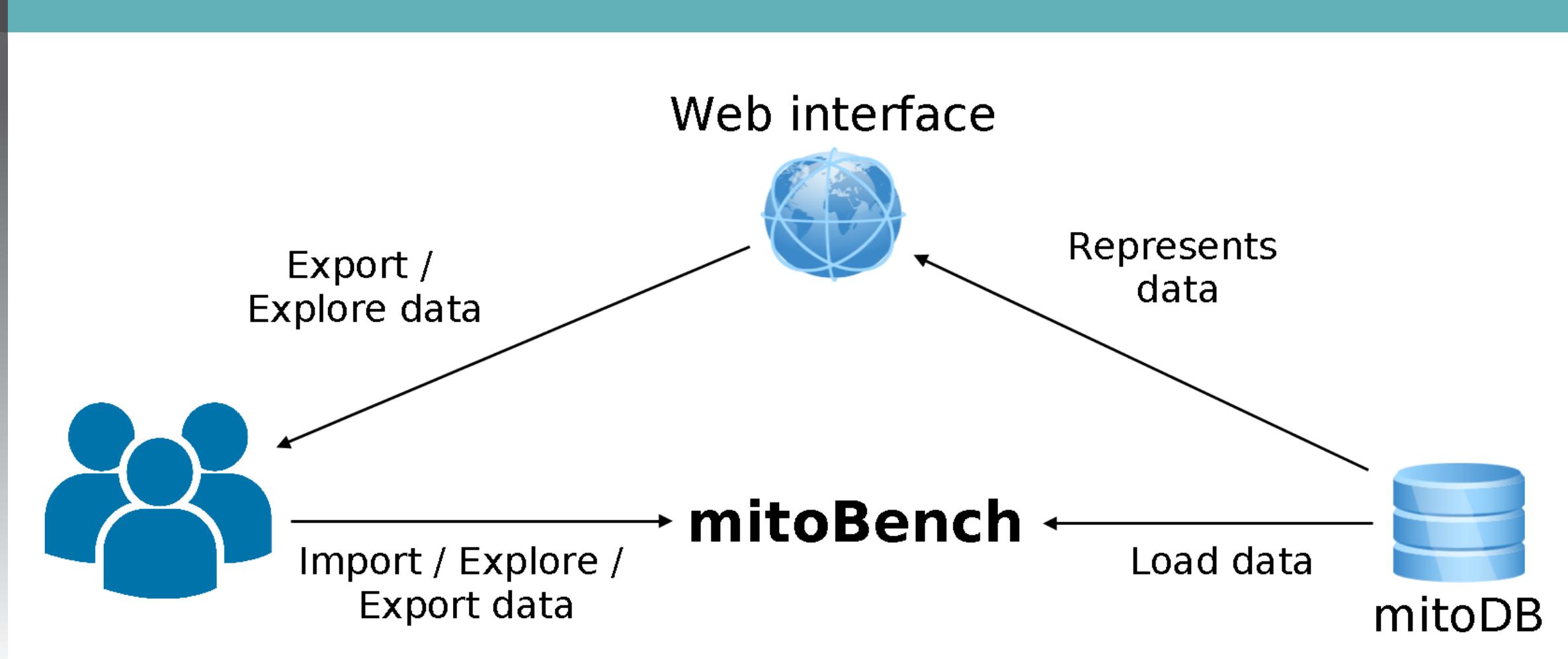


Contact: judith.neukamm@uni-tuebingen.de

## 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 → requiring manual interaction with the data for downstream analysis
- mitoBench: workbench to ease file conversions, interactively analyzing 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



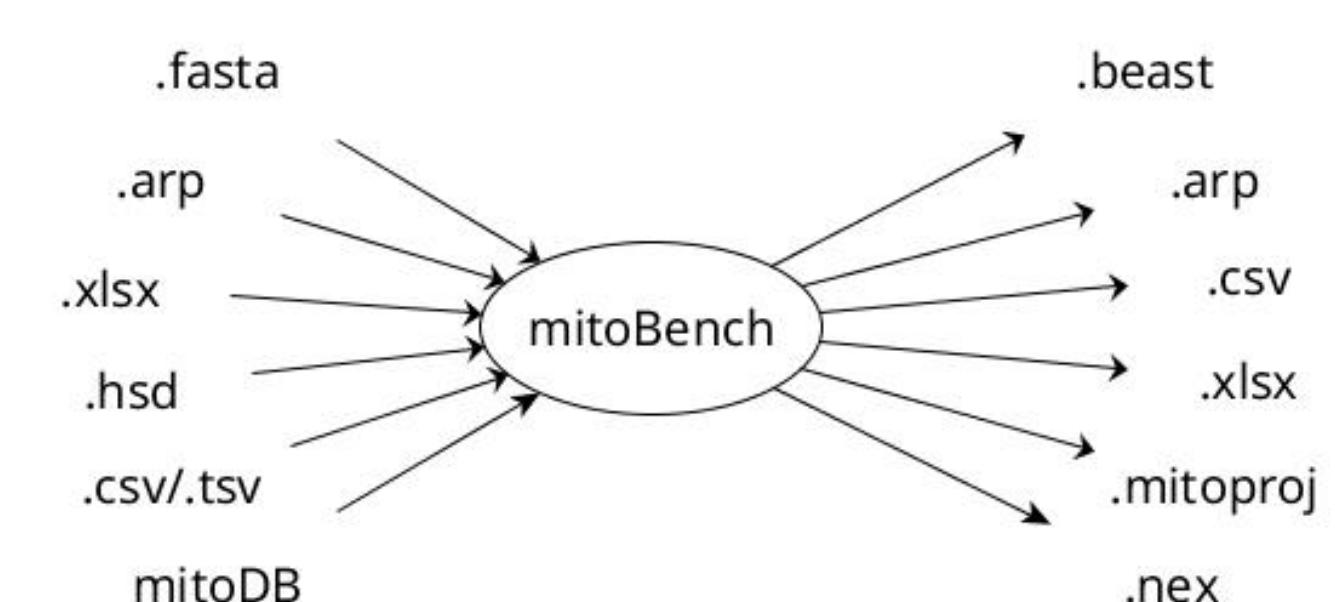
## mitoDB

- 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
J2975	GATCA...	R	Roman Period	Africa	Egypt	Beni Suef	Ancient
J2976	GATCA...	U	Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient
J2977	GATCA...	T1a5	Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient
J2916	NNAGA...	R0	Pre-Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient
J2972	GATCA...	HT1a2a	Roman Period	Africa	Egypt	Beni Suef	Ancient
J2918	GATCA...	J2a2a	Roman Period	Africa	Egypt	Beni Suef	Ancient
J2984	GATCA...	T1a5	Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient
J2137	GATCA...	J2a2b1	Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient
J2962	GATCA...	M2a1	Pre-Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient
J2989	GATCA...	K2a1	Pre-Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient
J2807	GATCA...	J2a1a1	Pre-Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient
J2153	GATCA...	R0a1a	Roman Period	Africa	Egypt	Beni Suef	Ancient
J2132	GATCA...	T	Roman Period	Africa	Egypt	Beni Suef	Ancient
J2128	GATCA...	HT2a1	Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient
J2130	GATCA...	M1a1	Roman Period	Africa	Egypt	Beni Suef	Ancient
J2952	GATCA...	M2a1	Pre-Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient
J2920M	GATCA...	R0a2f	Pre-Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient
J2896	GATCA...	HT1a2	Pre-Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient
J2883	GATCA...	T1a	Pre-Ptolemaic Period	Africa	Egypt	Beni Suef	Ancient
J2879	GATCA...	U2b	Roman Period	Africa	Egypt	Beni Suef	Ancient

### 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

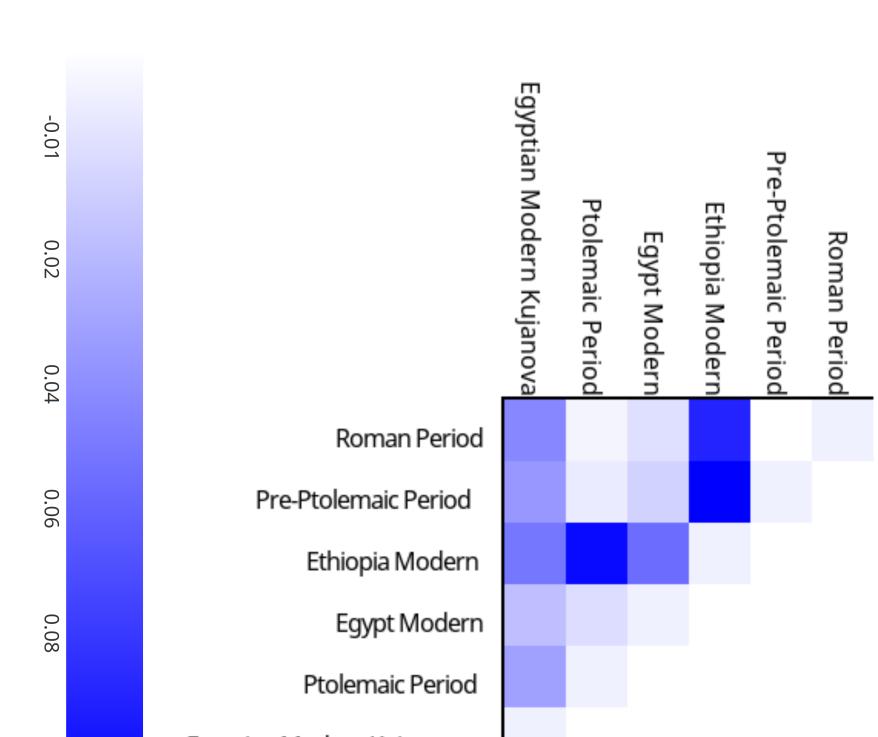
- Haplogroup filtering / frequencies
- Haplotype filtering / frequencies

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

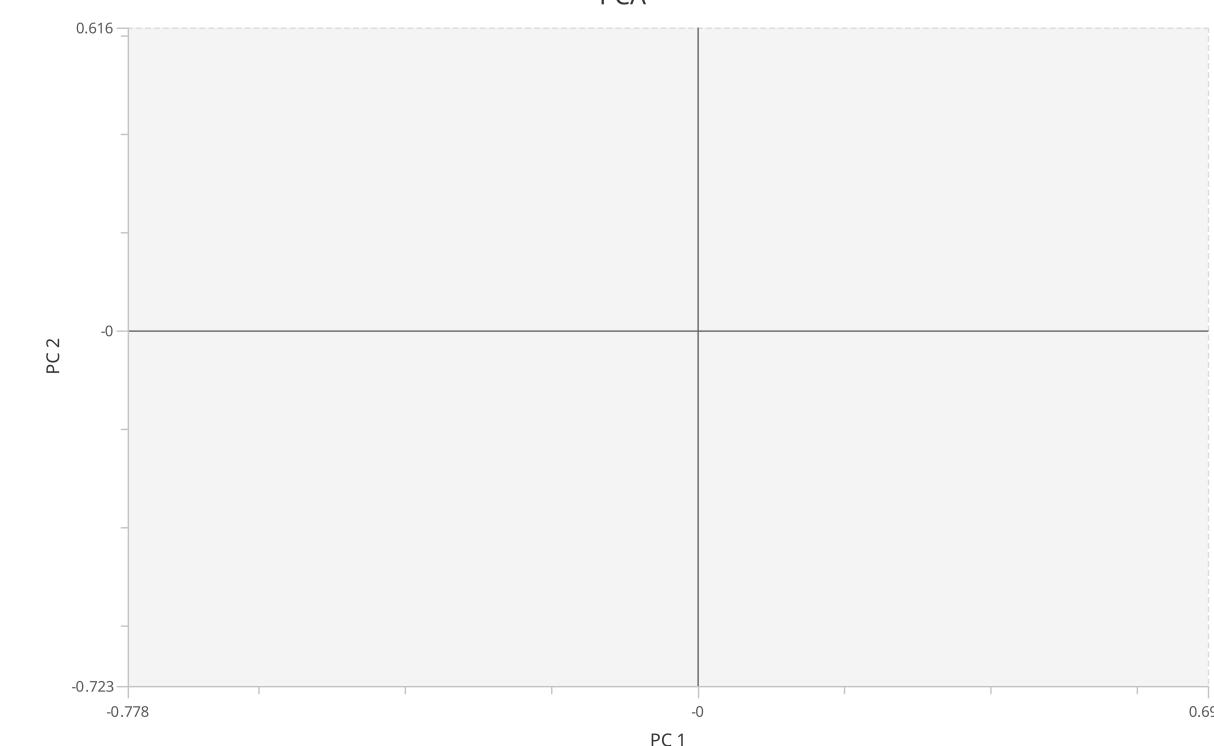
### Data analysis

Investigate correlations between populations with Fst and PC analysis.

### Fst analysis



### PCA



## Outlook

### mitoBench

- Provide methods for downstream analysis (e.g. Founder Analysis, Haplogroup determination)

### 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., ... & Schönherr, S. (2016). Haplotype 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic acids research*, gkw233.