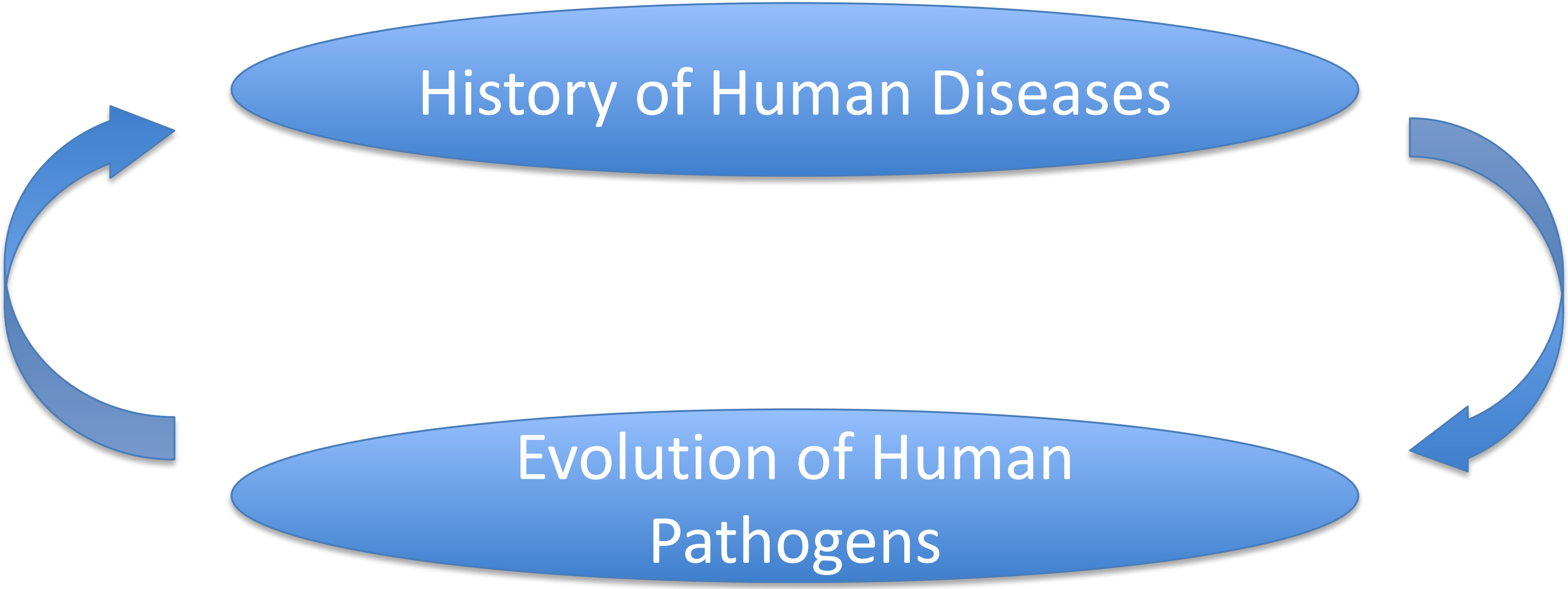
SPAAM

Standards,
Precautions &
Advances in
Ancient
Metagenomics

Lecture 4A: Introduction to Microbial Genomics

Alexander Herbig

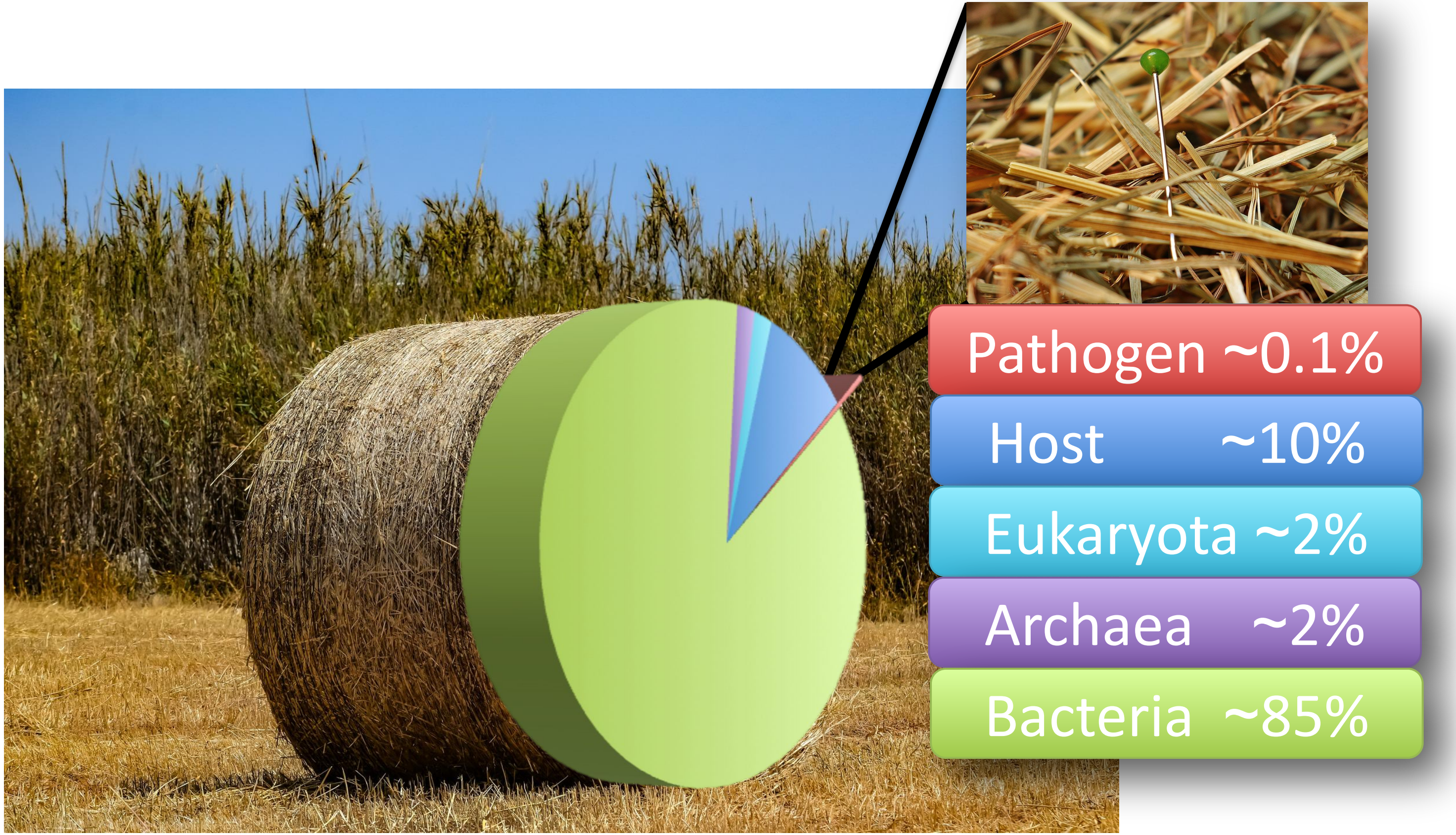


History of Human Diseases

Evolution of Human Pathogens

Pathogen Screening of Human Archaeological Remains

Pathogen Screening of Human Archaeological Remains



Pathogen ~0.1%

Host ~10%

Eukaryota ~2%

Archaea ~2%

Bacteria ~85%

Authentication

Is the recovered DNA of ancient Origin?

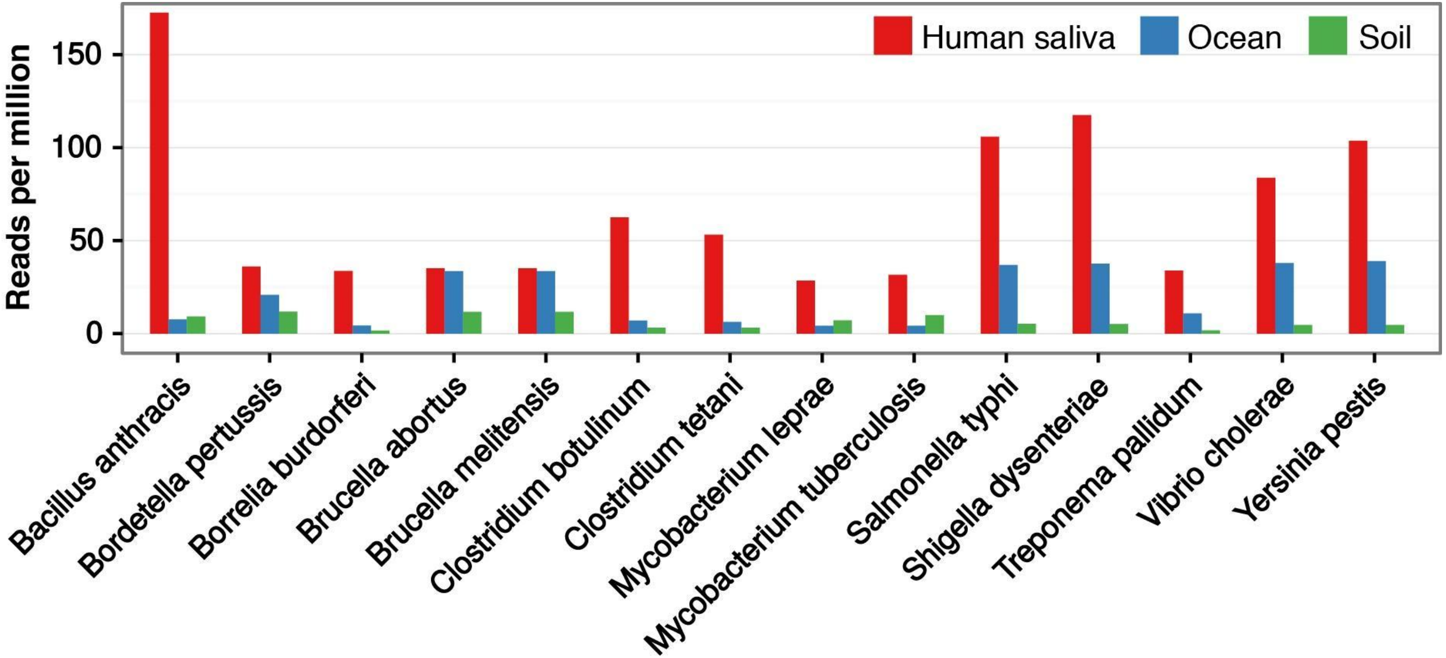
Can we differentiate Species?

Influence of Contamination?

Multiple Infections?

Species Identification

Close Environmental Relatives



Warinner et al, 2017

Species Identification

Database Biases

Issue: Databases are incomplete and biased towards pathogenic organisms

Suggested Measures:

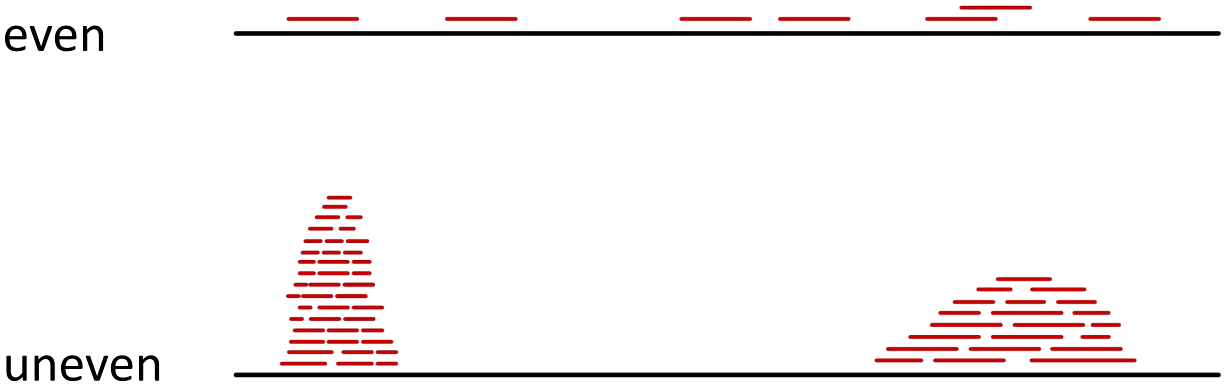
- Evenness of Genome Coverage
  (Random distribution vs. accumulation in conserved regions)

- Percent identity distributions
  (Distinguishing foreground and background)

Species Identification

Evenness of Coverage

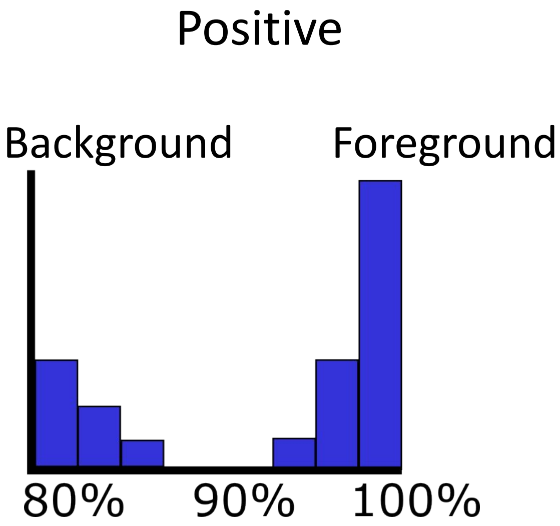
Distribution of reads across the reference genome

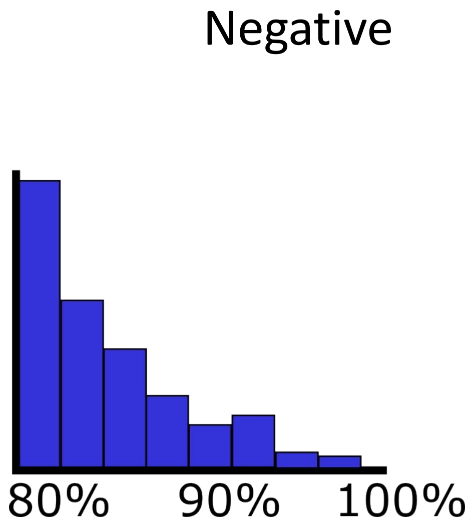


Species Identification

Similarity Distributions

Distributions of **%identity values** for aligned reads



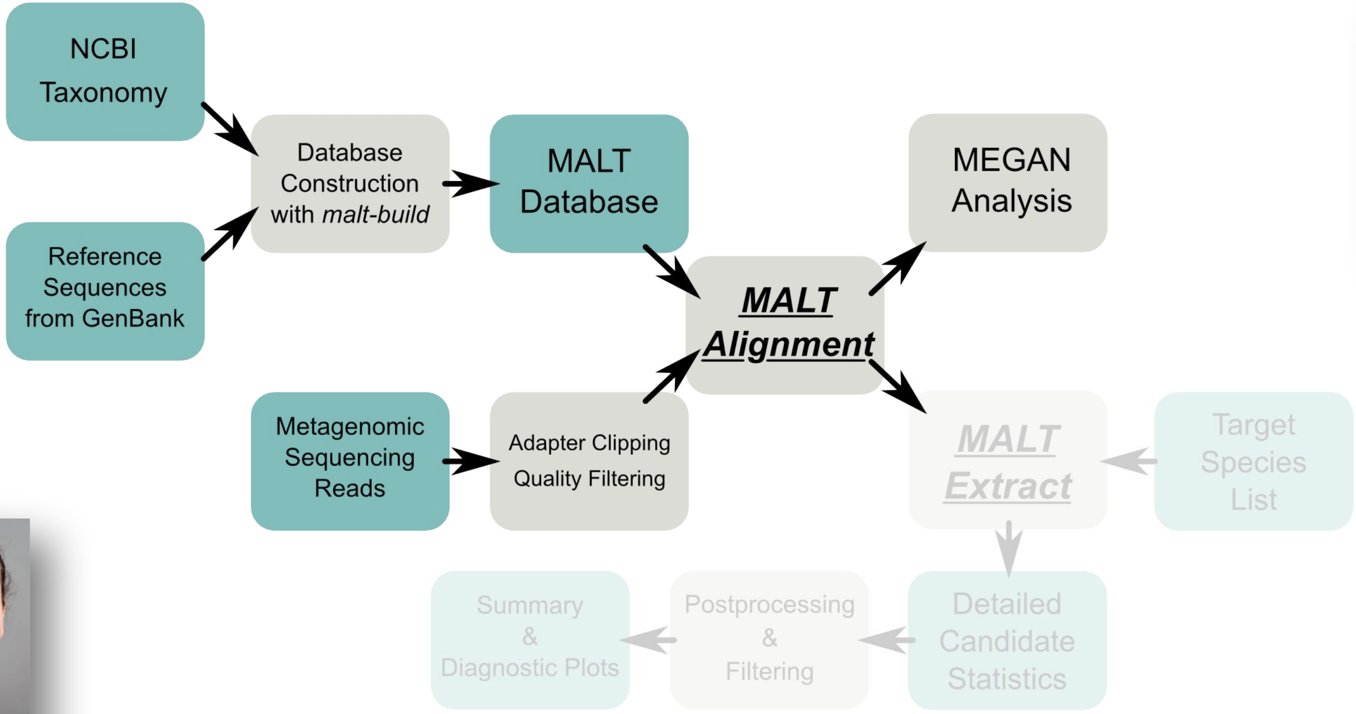


Authentication

- Ancient DNA?
    - DNA Damage patterns

- Correct species?
    - Taxonomy vs. Phylogeny
    - Database Biases
    - Close Environmental Relatives

HOPS – Heuristic Operations for Pathogen Screening

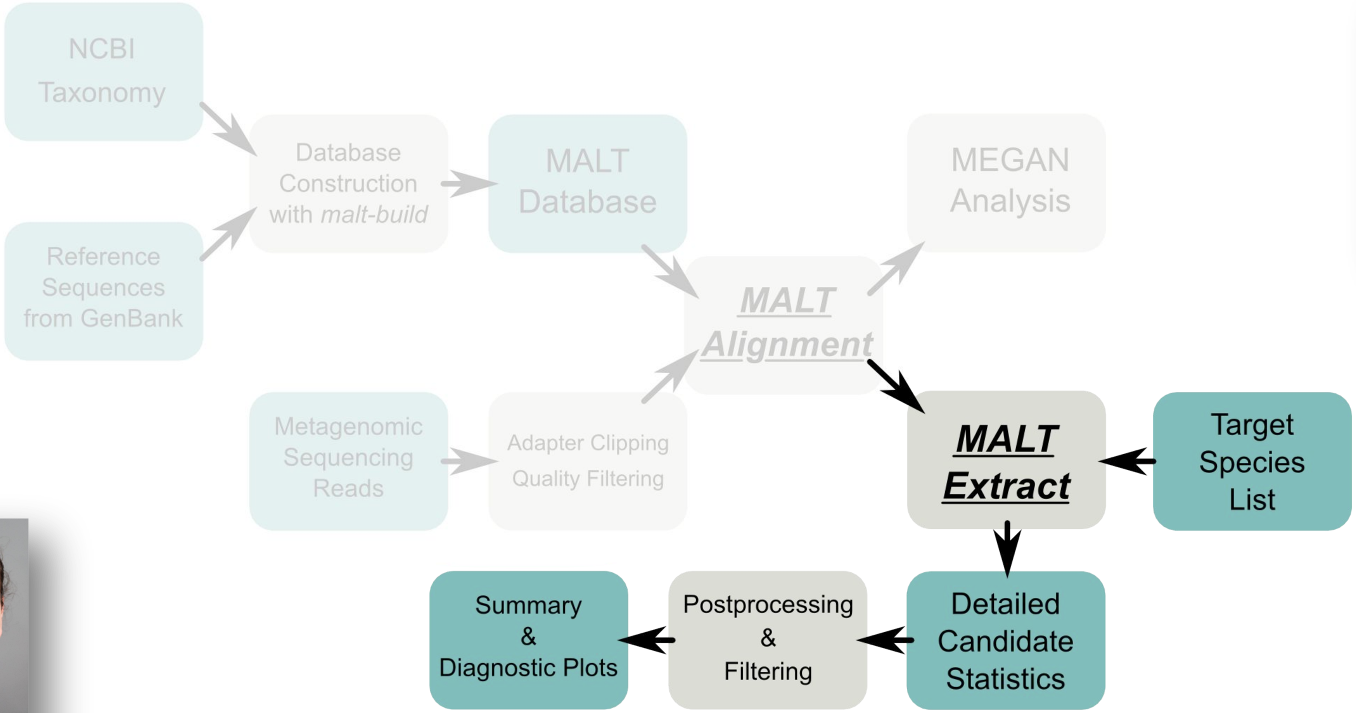


Ron Hübler

Felix Key

Hübler, Key et al., Genome Biology 2019

HOPS – Heuristic Operations for Pathogen Screening



NCBI Taxonomy

Reference Sequences from GenBank

Database Construction with *malt-build*

MALT Database

Metagenomic Sequencing Reads

Adapter Clipping Quality Filtering

MALT Alignment

MEGAN Analysis

MALT Extract

Target Species List

Detailed Candidate Statistics

Postprocessing & Filtering

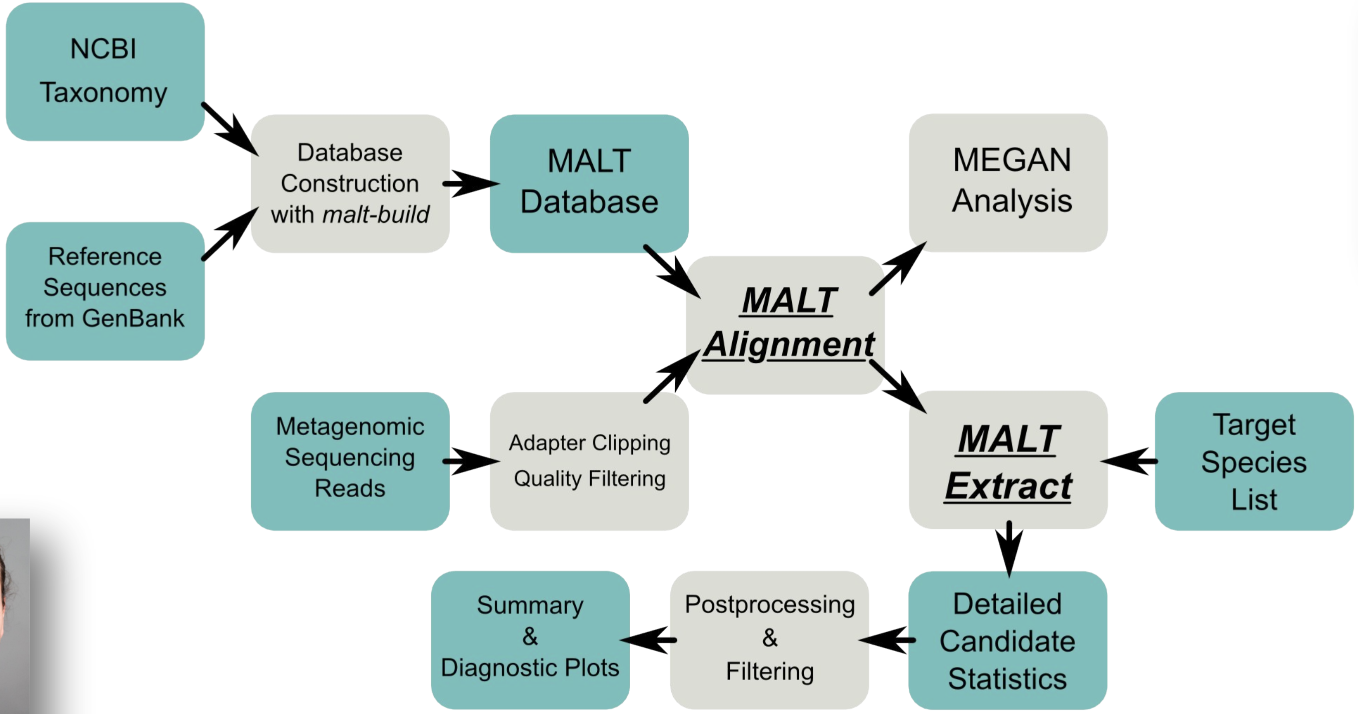
Summary & Diagnostic Plots

Ron Hübler

Felix Key

Hübler, Key et al., Genome Biology 2019

HOPS – Heuristic Operations for Pathogen Screening



NCBI Taxonomy

Reference Sequences from GenBank

Database Construction with *malt-build*

MALT Database

MEGAN Analysis

MALT Alignment

Metagenomic Sequencing Reads

Adapter Clipping Quality Filtering

MALT Extract

Target Species List

Detailed Candidate Statistics

Postprocessing & Filtering

Summary & Diagnostic Plots

Ron Hübler

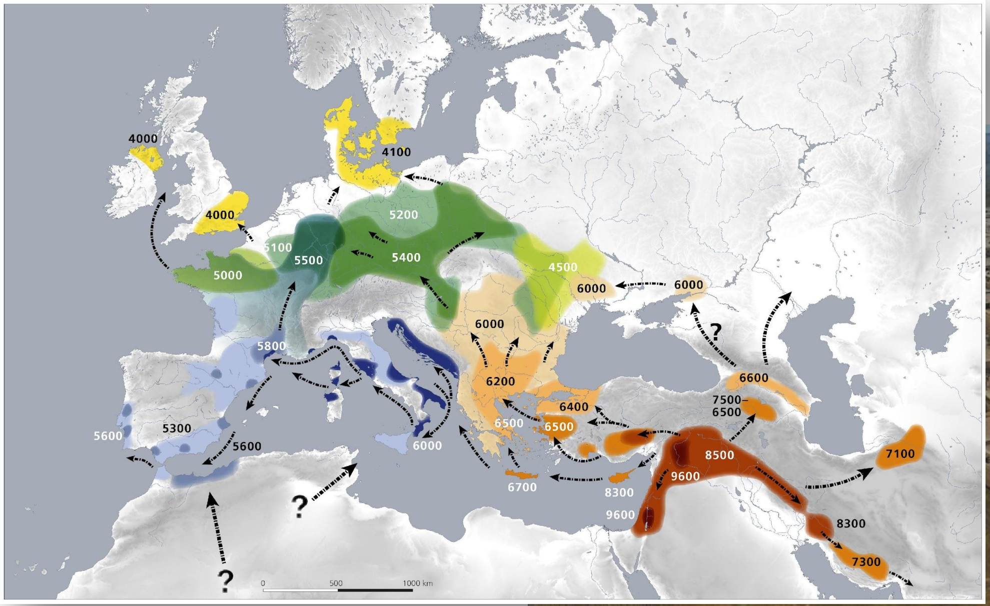
Felix Key

Hübler, Key et al., Genome Biology 2019

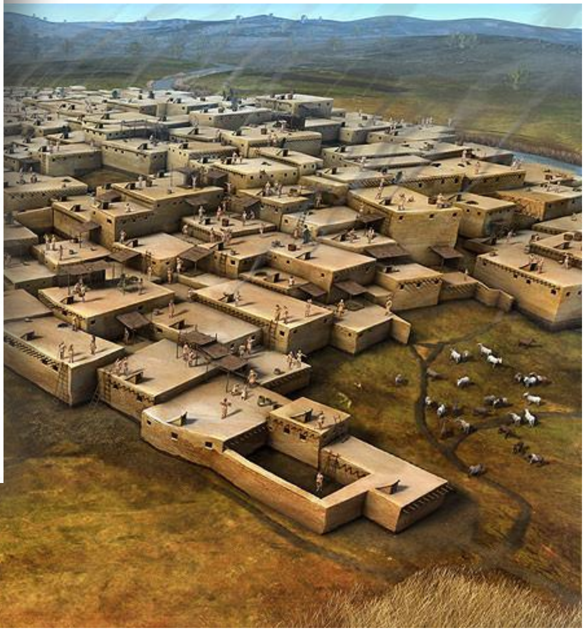
Evolution of Human Pathogens

Where to start?

The Neolithic Revolution

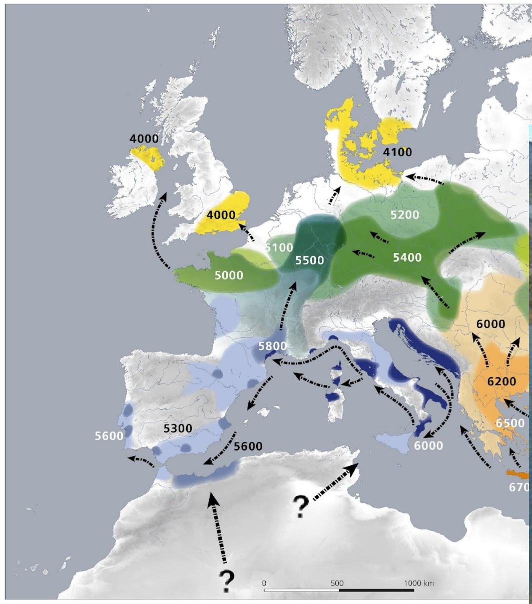


D. Gronenborn/ M. Ober, RGZM [CC BY 4.0]

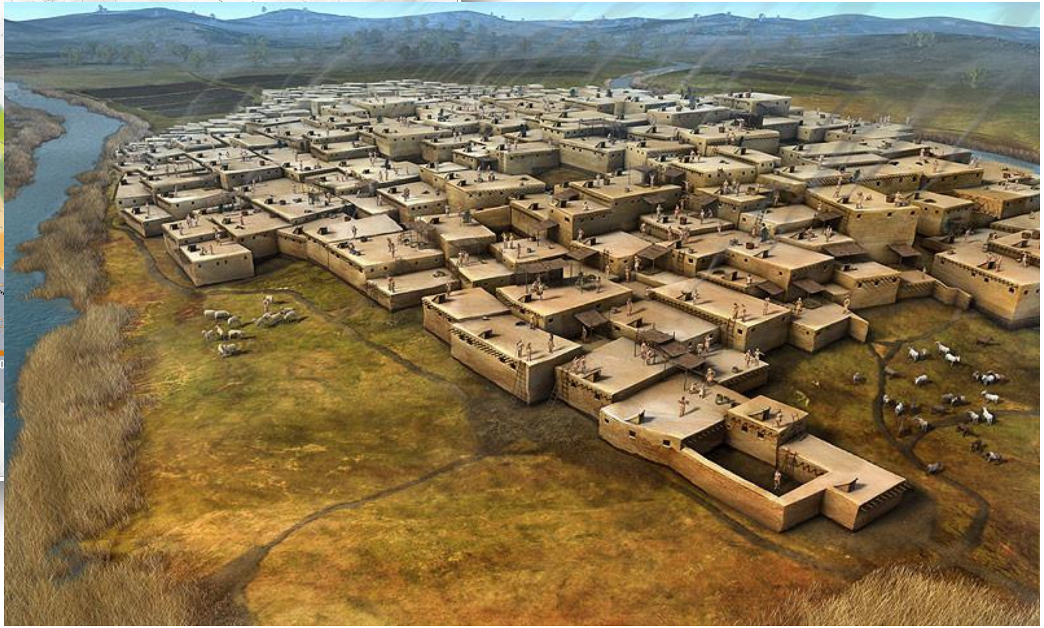


Çatalhöyük by Dan Lewandowski

The Neolithic Revolution

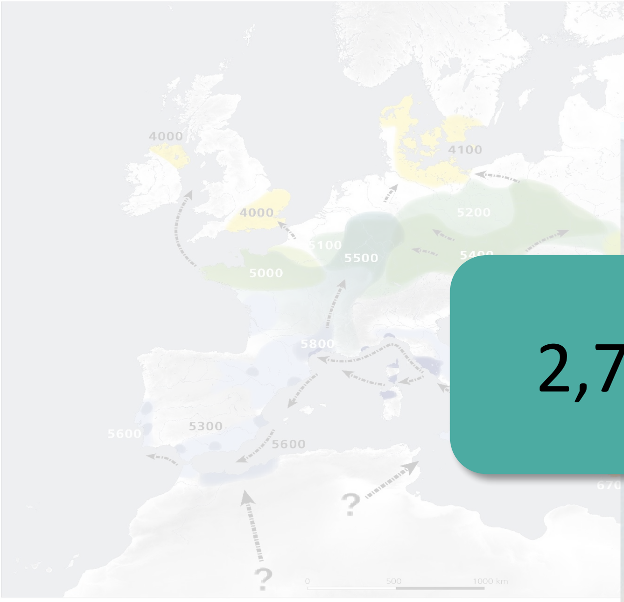


D. Gronenborn/ M. Ober, RGZM [CC BY 4.0]



Çatalhöyük by Dan Lewandowski

The Neolithic Revolution

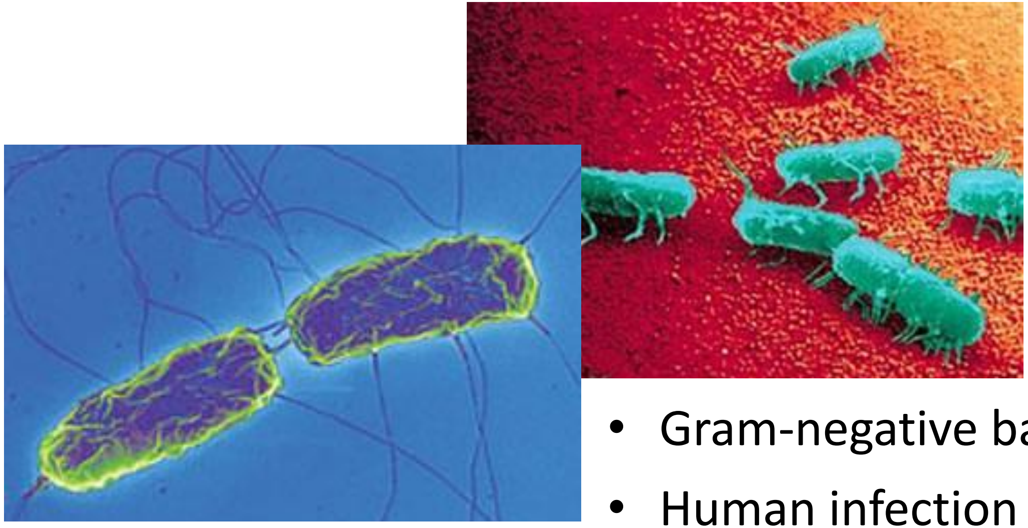


D. Gronenborn/ M. Ober, RGZM [CC BY 4.0]

Çatalhöyük by Dan Lewandowski

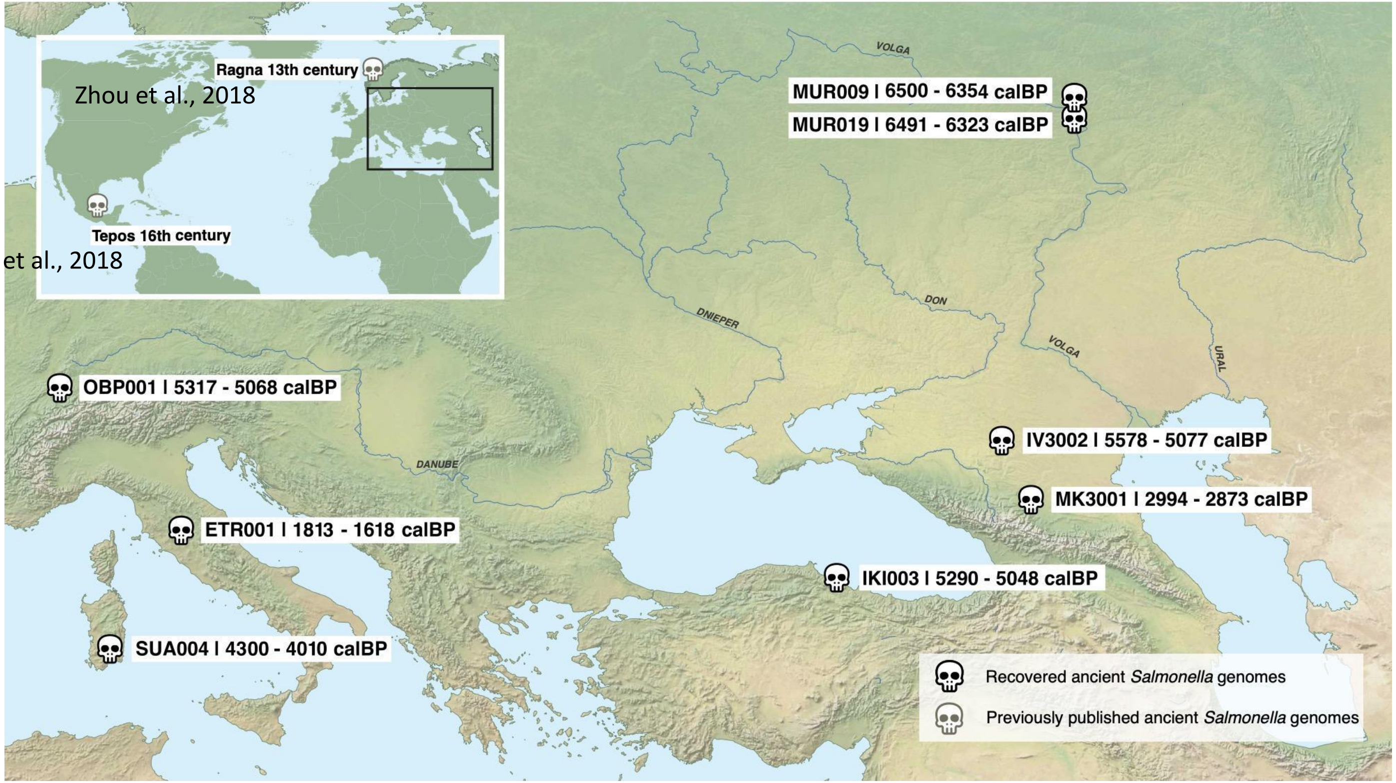
2,726 Data Sets Screened

Salmonella enterica



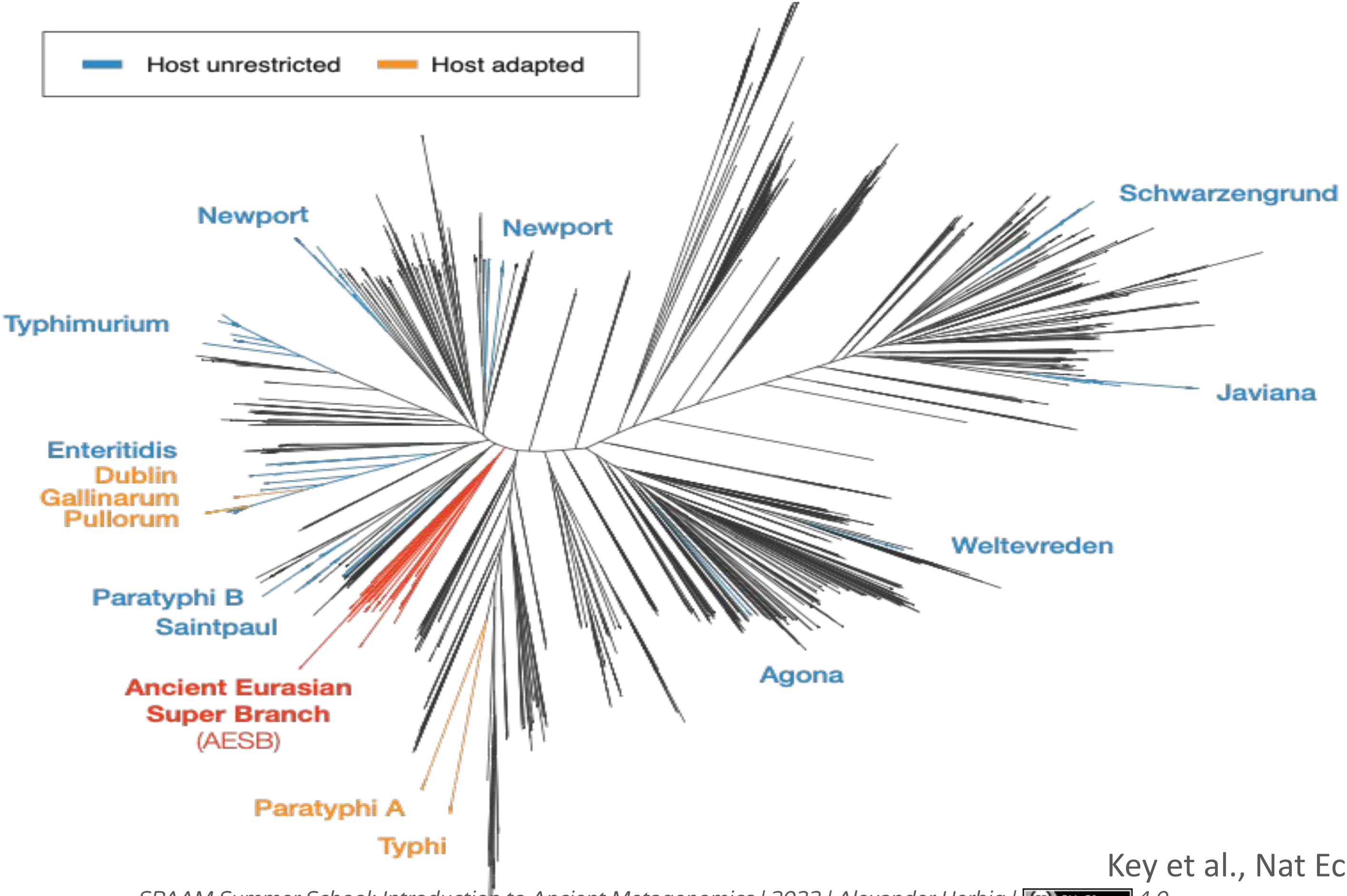
- Gram-negative bacterium
- Human infection often through **contaminated food**
- High number of **environmental** isolates
- Several rather **host-specific** serovars
- Human-specific serovars: Typhi, Paratyphi A,B,C (**typhoid, paratyphoid fever**)

Eurasian Time Transect



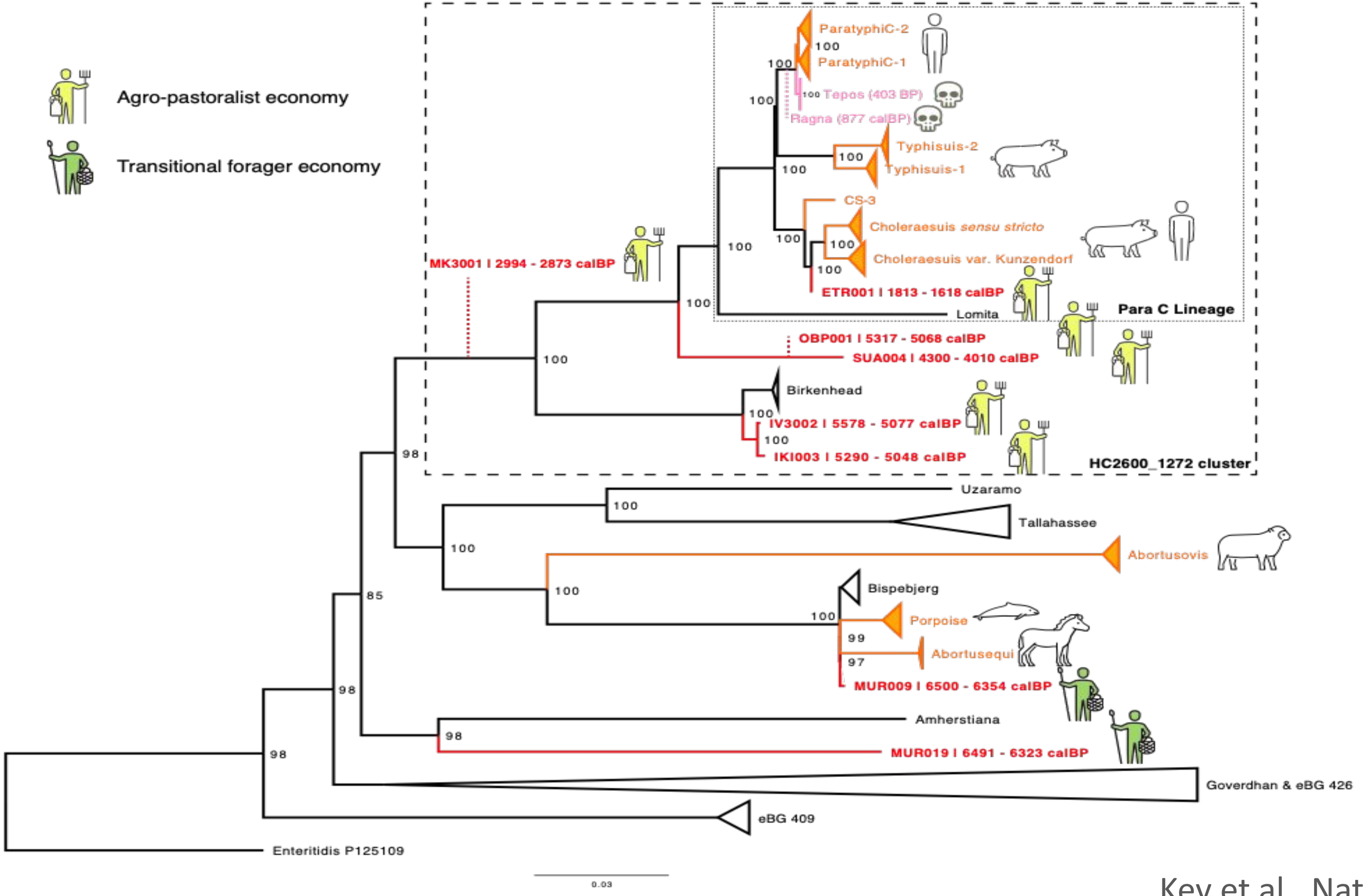
Key et al., Nat Ecol Evol 2020

Eurasian *S. enterica* during the Neolithisation



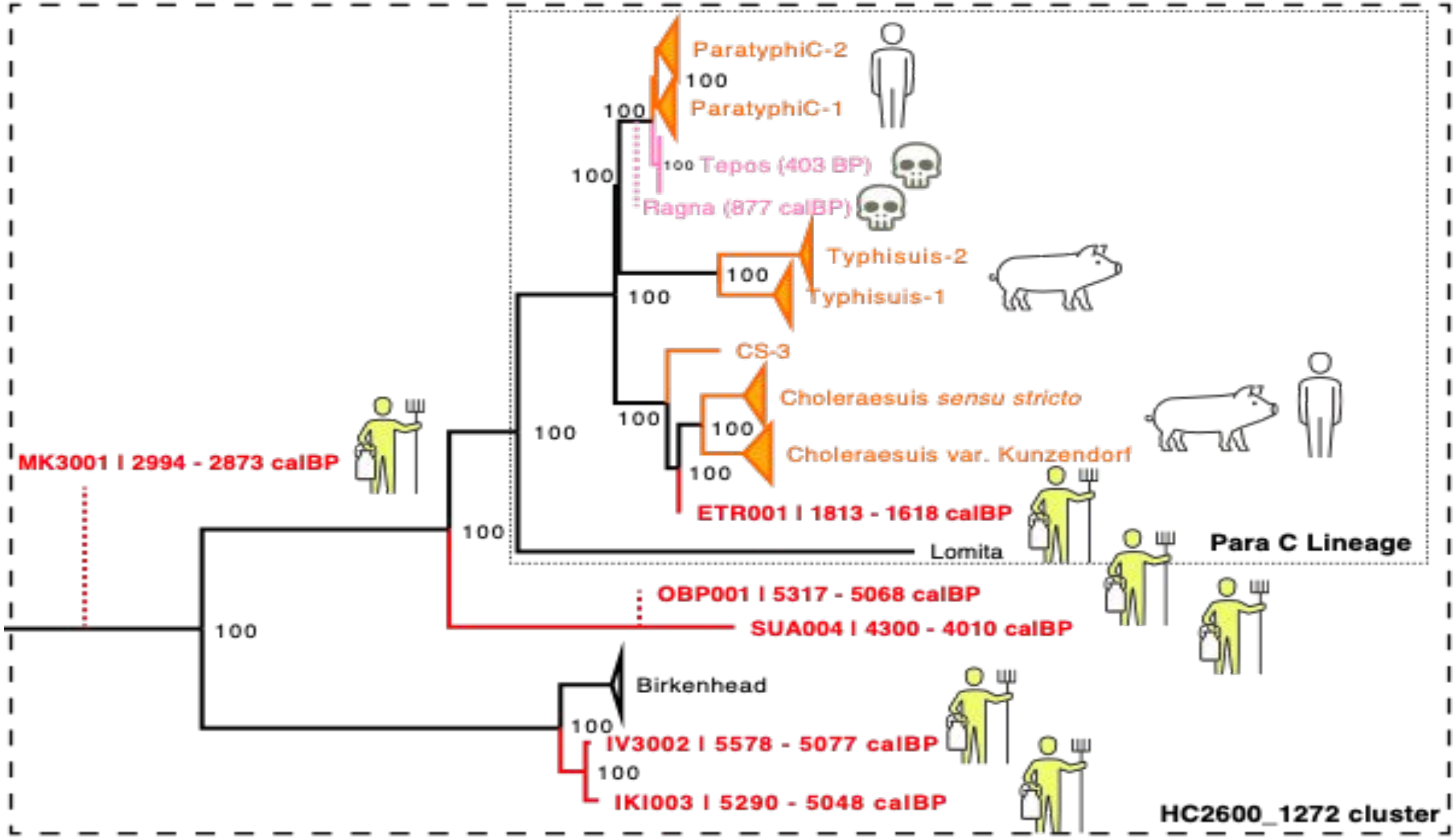
Key et al., Nat Ecol Evol 2020

Eurasian *S. enterica* during the Neolithisation



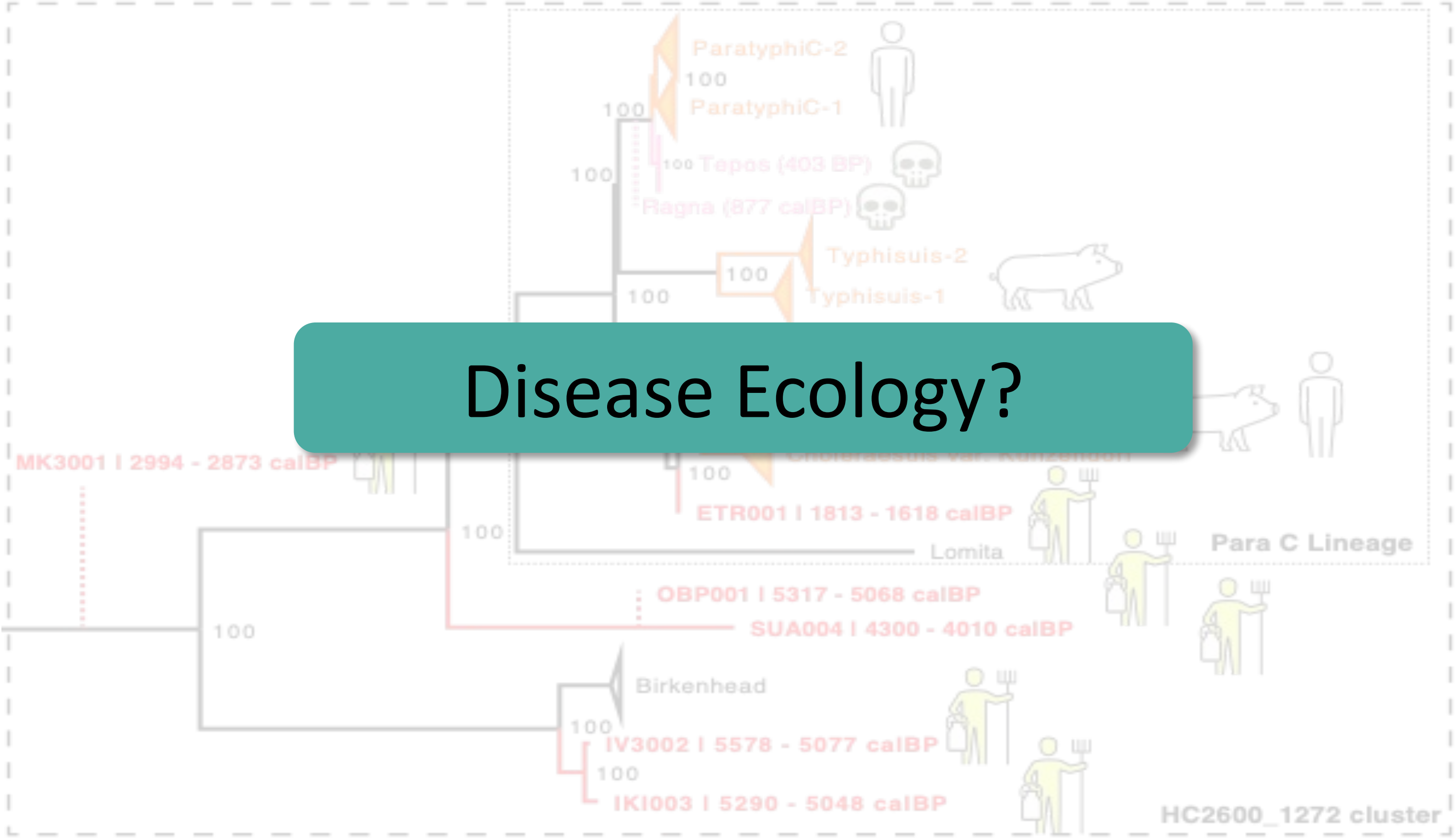
Key et al., Nat Ecol Evol 2020

Eurasian *S. enterica* during the Neolithisation



Key et al., Nat Ecol Evol 2020

Eurasian *S. enterica* during the Neolithisation



Disease Ecology?

Pseudogenisation

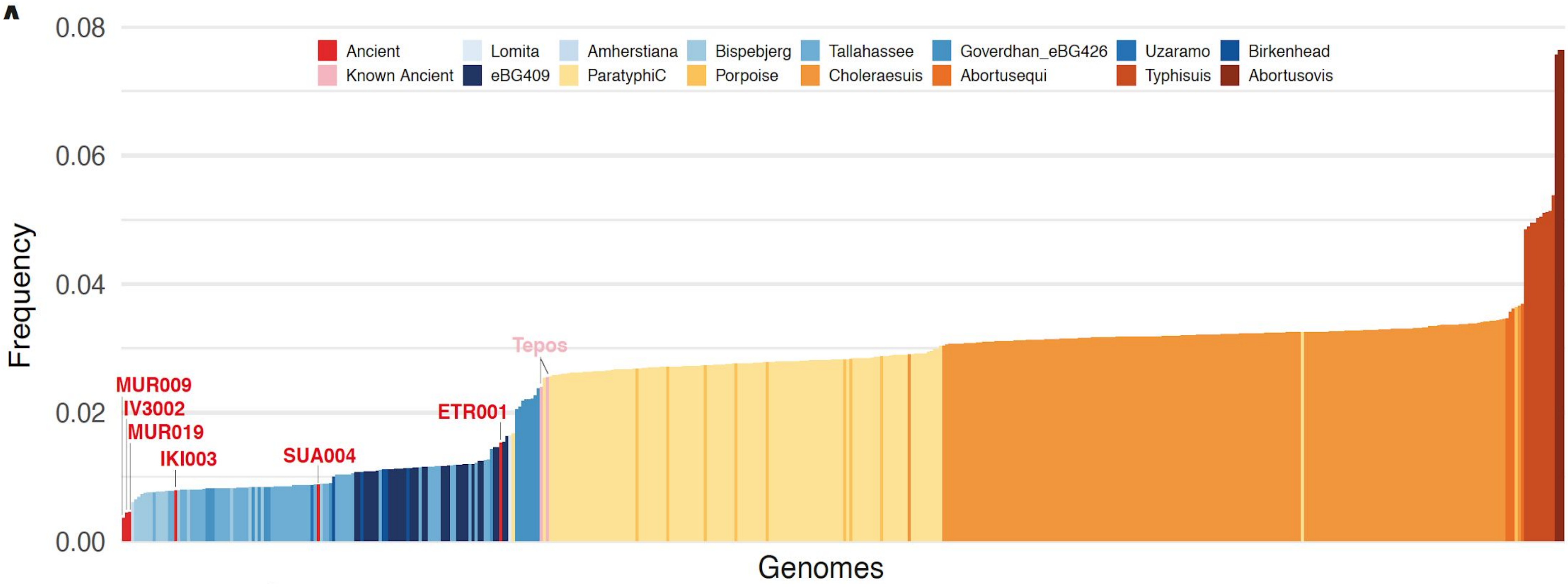
Process by which genes become non-functional

- Point mutation resulting in a stop-codon
- Insertion/deletion resulting in a frame shift

Not negatively selected because...

- Function not needed anymore (change in ecology)
- Function has even a negative effect (-> positive selection)
- Duplication (one functional copy is enough)

Pseudogenisation

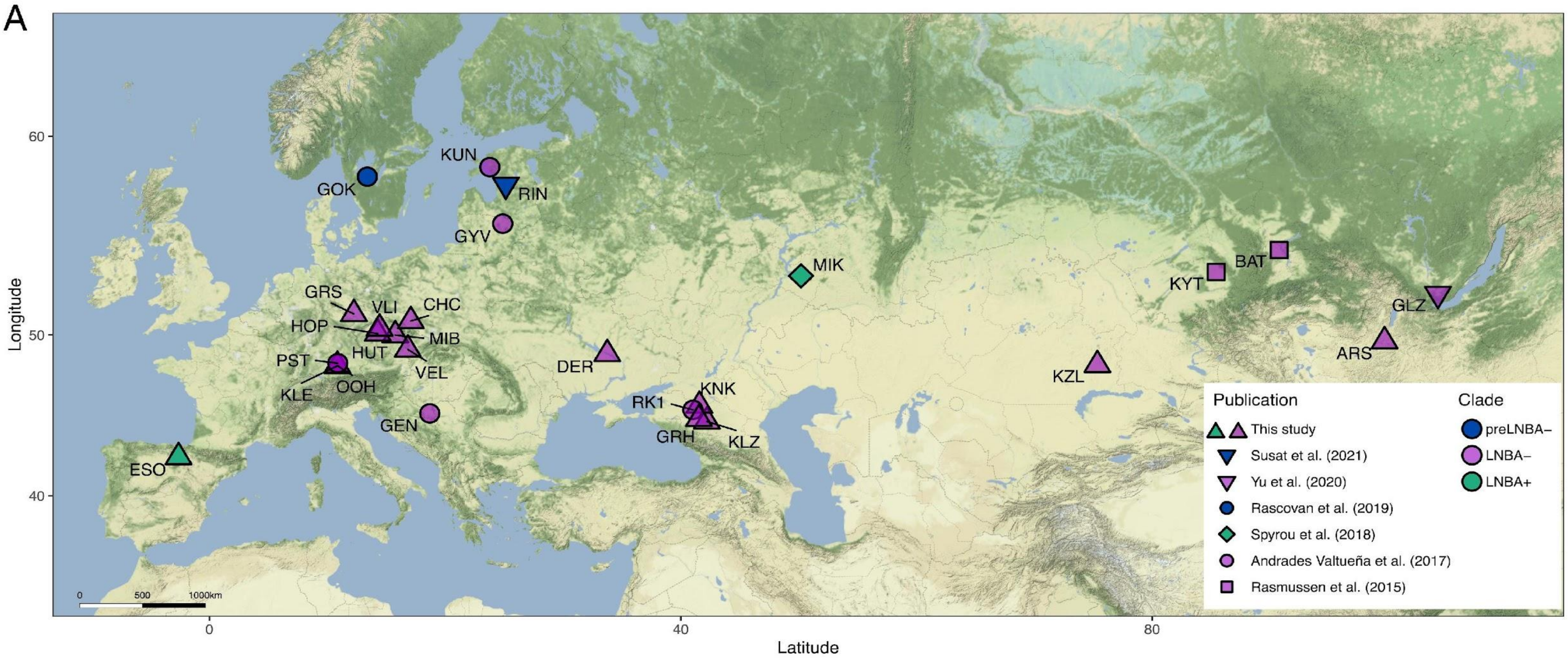


Key et al., Nat Ecol Evol 2020

The Plague

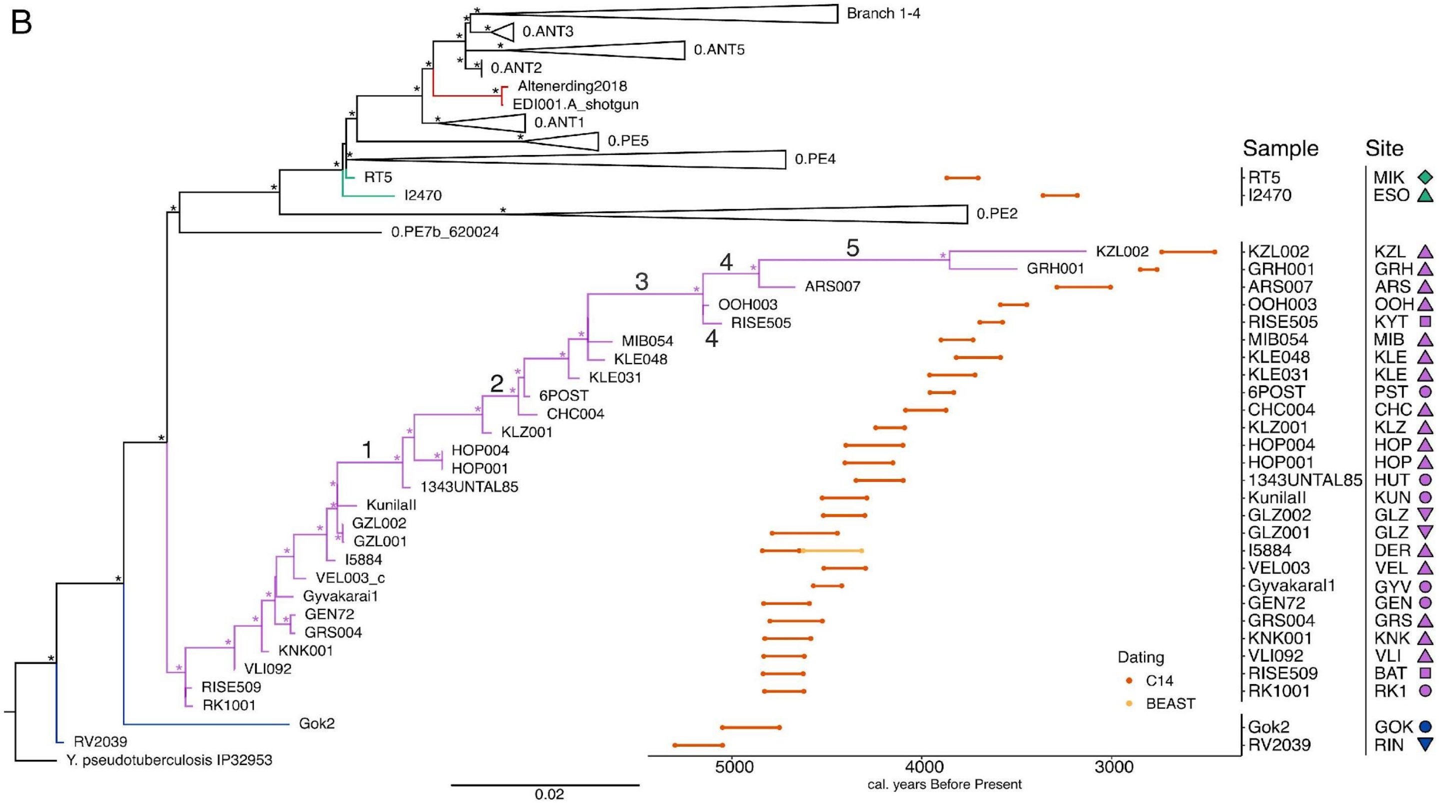
Scène de la peste de 1720 à la Tourette (Marseille)

The Stone Age Plague (Late Neolithic, Bronze Age, Iron Age)



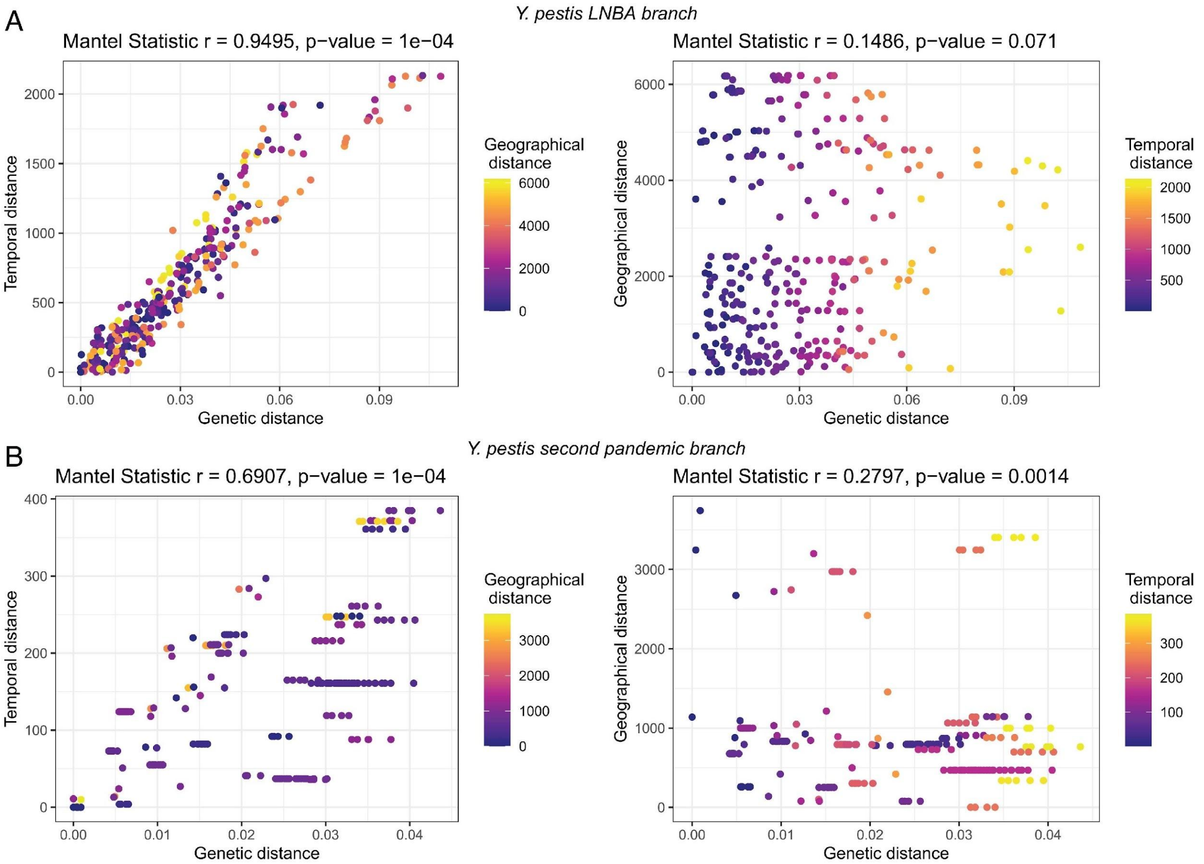
Andrades Valtueña et al, 2022

The Stone Age Plague – Phylogeny



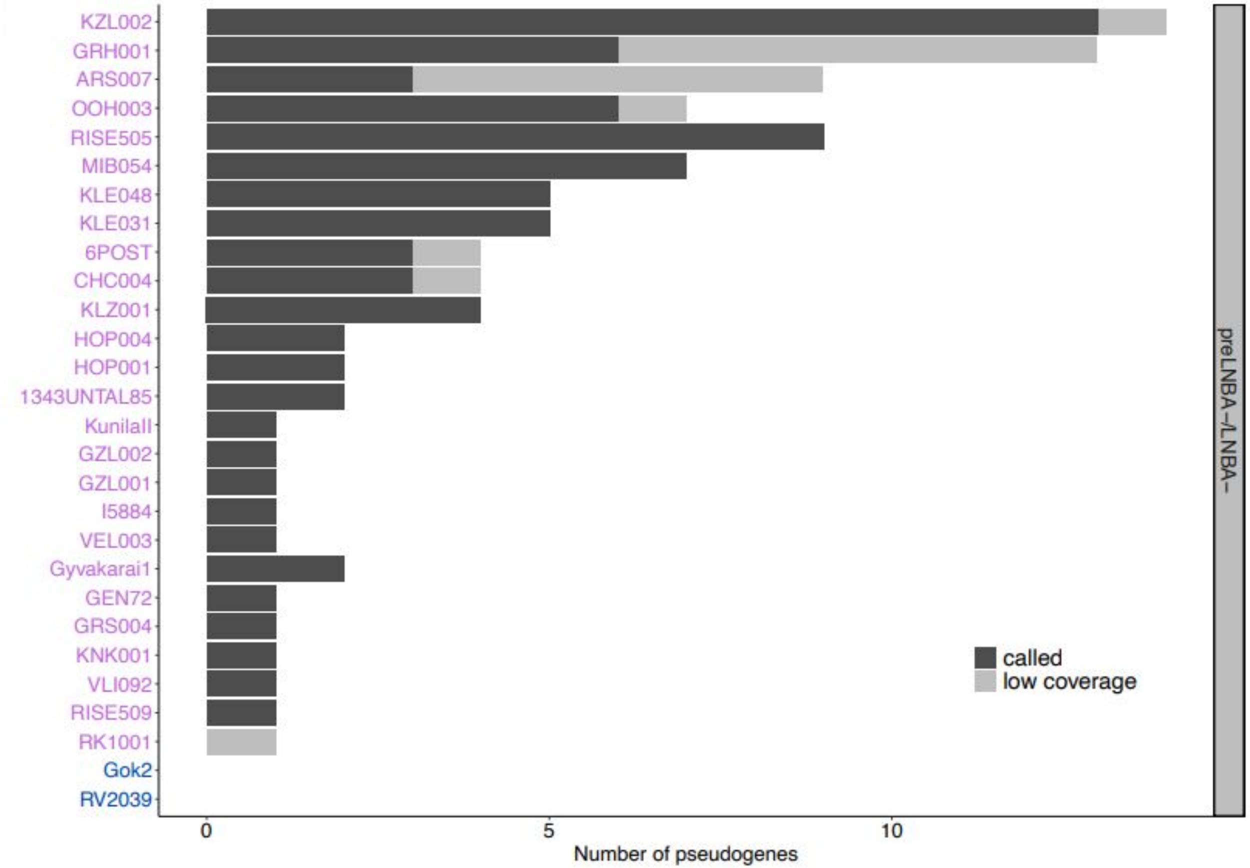
Andrades Valtueña et al, 2022

The Stone Age Plague – Genetics vs Time/Geography



Andrades Valtueña et al, 2022

The Stone Age Plague – Pseudogenes



Andrades Valtueña et al, 2022

The Stone Age Plague – Conclusions

- Pre-historic *Yersinia pestis* wide spread and highly mobile
- Evolutionary pattern points to single well connected reservoir
- Parallel spread of lineages with different ecological background over millennia

Literature

Screening and Authentication:

A Robust Framework for Microbial Archaeology
Annual Review of Genomics and Human Genetics 2017

Mining Metagenomic Data Sets for Ancient DNA: Recommended Protocols for Authentication
Trends in Genetics 2017

HOPS: automated detection and authentication of pathogen DNA in archaeological remains
Genome Biology 2019

Human Pathogen Evolution:

Paleomicrobiology: Diagnosis and Evolution of Ancient Pathogens
Annual Review of Microbiology 2019

Ancient pathogen genomics as an emerging tool for infectious disease research
Nature Reviews Genetics 2019

Emergence of human-adapted *Salmonella enterica* is linked to the Neolithization process
Nature Ecology & Evolution 2020