

Introduction to Ancient Metagenomics

Summer School

Supported by SPAAM, Werner Siemens-Stiftung, and JSMC

Description

Ancient metagenomics applies cutting-edge metagenomic methods to the degraded DNA content of archaeological and paleontological specimens. The rapidly growing field is currently uncovering a wealth of novel information for both human and natural history, from identifying the causes of devastating pandemics such as the Black Death, to revealing how past ecosystems changed in response to long-term climatic and anthropogenic change, to reconstructing the microbiomes of extinct human relatives. However, as the field grows, the techniques, methods, and workflows used to analyse such data are rapidly changing and improving.

In this summer school (block practicum) we will go through the main steps of ancient metagenomic bioinformatic workflows, familiarising students with the command line, demonstrating how to process next-generation-sequencing (NGS) data, and showing how to perform *de novo* metagenomic assembly. Focusing on host-associated ancient metagenomics, the course consists of a combination of lectures and hands-on exercises, allowing participants to become familiar with the types of questions and data researchers work with. Round table discussions with experts at each stage of the workflow will be held to allow participants to get advice on their own projects and research.

By the end of the course, participants will have an understanding of how to effectively carry out the major bioinformatic components of an ancient metagenomic project in an open and transparent manner.

Organisation

What

The summer school will be a 5 day practical workshop (block practicum). Attendees who participate in the entire course will be eligible for 1.5 ECTS points (40 hours of instruction and practicals, 5 hours of coursework).

Who

The course is aimed at masters students and first- or second-year level PhD students, to a maximum of 40 participants.

Where

This summer school will be held virtually on a platform such as gather.town.

When

The summer school will be provisionally held between 31st July-4th August, 09:00-17:00 CEST (Berlin Time).

Content

- **Lecture** Introduction to NGS data: *TBC*
 - What is NGS sequencing and what does its data look like?
- **Practical** BareBonesBash: *Aida Andrades Valtueña and Theseas Lamnidis*
 - Introduction to the UNIX command line.
- **Practical** Introduction to R and the tidyverse / Introduction to Python and Pandas: *Clemens Schmid and James Fellows Yates / Kevin Nota and Robin Warner*
 - How to read/write files and run statistical analyses? How to make figures with ggplot2/plotnine
 - (Note: Parallel sessions participants can choose which to attend based on their prior experience)
- **Round table** Introductions
 - Welcome - let's get to know each other

DAY 2

- **Lecture** Introduction to ancient DNA: *Christina Warinner*
 - How to identify and handle DNA from ancient contexts
- **Practical** Introduction to R and the tidyverse / Introduction to Python and Pandas: *Clemens Schmid and James Fellows Yates / Kevin Nota and Robin Warner*
 - How to read/write files and run statistical analyses? How to make figures with ggplot2/plotnine
 - (Note: Parallel sessions participants can choose which to attend based on their prior experience)
- **Practical** Bytesize git: *James Fellows Yates*
 - What is git and github? How to clone a repository? What does commit/pull/push mean?
- **Practical** ancientMetagenomeDir: *James Fellows Yates*
 - Where to get ancient metagenomic comparative data, and why is metadata important?
 - Git practice
- **Round table** Project organisation
 - How to structure my project analysis folders? How to reproducibly document my work?

DAY 3

- **Lecture** Introduction to Metagenomics: *Christina Warinner*
 - What are the questions tackled in metagenomics? What are the challenges?
- **Practical** Taxonomic Profiling, OTU Tables and Visualisation: *Nikolay Oskolkov and Kevin Nota*
 - Python practice. How to normalise OTU tables and perform microbial ecology analyses.
- **Practical** Genome mapping: *Alina Hiss and Alexander Herbig*

- Why and how do we map against a reference? What parameters are important for aDNA?
- **Round table** Taxonomic Classifiers
 - How do I pick a taxonomic classifier? What is best for aDNA?

DAY 4

- **Lecture** Introduction to microbial genomics: *Meriam Guellil*
 - How can we learn from studying the genomes of microbes? What questions can we ask?
- **Practical** Genome assembly: *Nikolay Oskolkov and Aida Andrades Valtueña*
 - De novo methods for obtaining metagenomically assembled genomes (MAGs). What are the best pipelines and parameters for ancient DNA?
- **Practical** Phylogenomics: *Arthur Kocher and Alexander Herbig*
 - How to perform phylogenetic analysis. What to consider when dealing with low coverage data
- **Roundtable** Databases
 - Why are databases so important? What impact does database selection have on my results? What pitfalls should I watch out for?

DAY 5

- **Lecture** Evolutionary Biology: *Alexander Herbig*
 - How have microbes and microbial communities evolved and changed through time? How does this inform our current understanding of the relationships among microbes?
- **Practical** Authentication and Decontamination: *James Fellows Yates and Nikolay Oskolkov*
 - How can I validate my DNA is ancient? What sort of analyses can convince ourselves and a reviewer we have good preservation?
- **Practical** Ancient Metagenomic Pipelines: *James Fellows Yates and Nikolay Oskolkov*
 - What pipelines exist for specifically ancient DNA and ancient metagenomic data?
- **Round table** Workshop recap