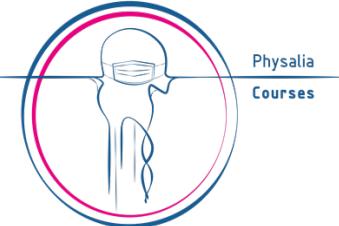
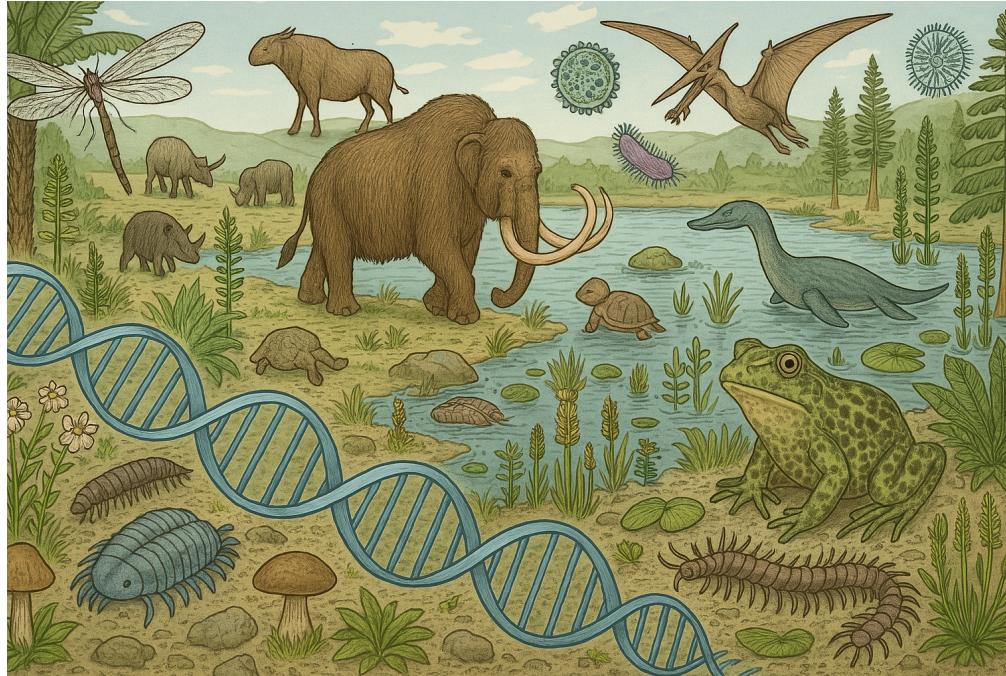


Ancient Metagenomics Analysis

Physalia course, online, 27-29 May 2025

Course outline and practical information

Nikolay Oskolkov, Lund University, NBIS SciLifeLab



Physalia
Courses

About us

Organizer: Carlo Pecoraro, Physalia courses

info@physalia-courses.org



Instructor:

Dr. Nikolay Oskolkov, Lund University, NBIS SciLifeLab

Nikolay.Oskolkov@biol.lu.se



@NikolayOskolkov



@oskolkov.bsky.social

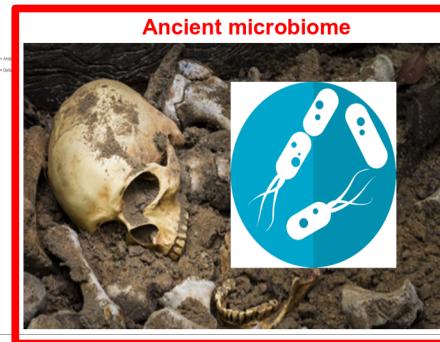
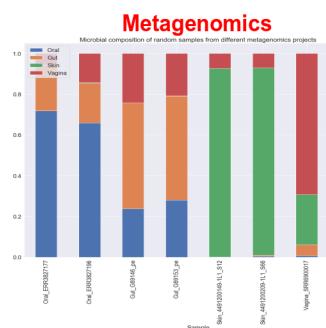
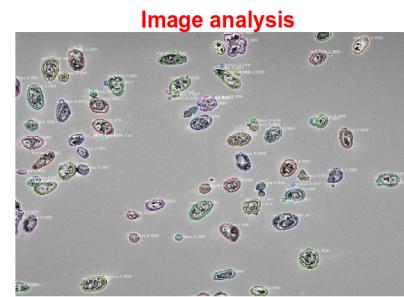
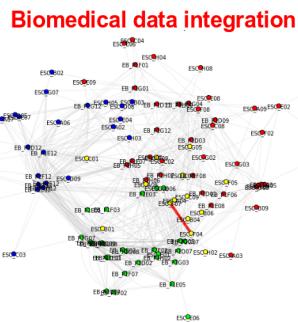
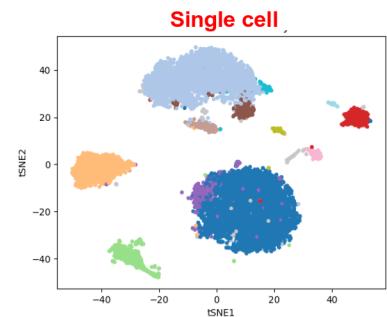


Personal homepage:
<https://nikolay-oskolkov.com>

2007 PhD in theoretical physics

2011 medical genomics at Lund University

2016 working at NBIS SciLifeLab, Sweden



Pochon et al. *Genome Biology* (2023) 24:242
<https://doi.org/10.1186/s13059-023-03083-9>

METHOD

Genome Biology

Open Access



aMeta: an accurate and memory-efficient ancient metagenomic profiling workflow

Zoé Pochon^{1,2†}, Nora Bergfeldt^{1,3,4†}, Emrah Kirdök⁵, Mário Vicente^{1,2}, Thijessen Naidoo^{1,2,6,7}, Tom van der Valk^{1,4}, N. Ezgi Altıntışık⁸, Maja Krzewińska^{1,2}, Love Dalén^{1,3}, Anders Götherström^{1,2†}, Claudio Mirabello^{9†}, Per Unneberg^{10†} and Nikolay Oskolkov^{11†}

¹Zoé Pochon, Nora Bergfeldt, Anders Götherström, Claudio Mirabello, Per Unneberg, and Nikolay Oskolkov shared authorship.

*Correspondence:
Nikolay.Oskolkov@biol.lu.se

¹¹Department of Biology,
Science for Life Laboratory,
National Bioinformatics
Infrastructure Sweden, Lund
University, Lund, Sweden
Full list of author information is
available at the end of the article

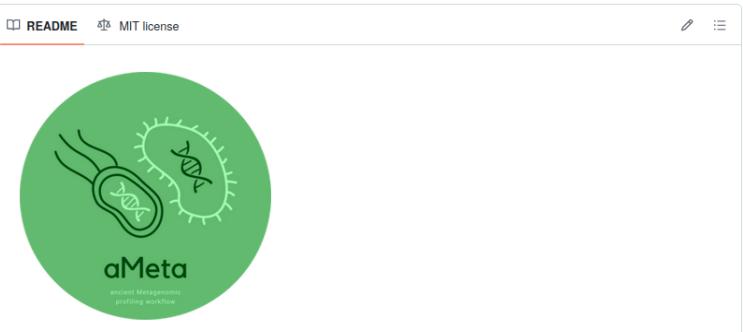
Abstract

Analysis of microbial data from archaeological samples is a growing field with great potential for understanding ancient environments, lifestyles, and diseases. However, high error rates have been a challenge in ancient metagenomics, and the availability of computational frameworks that meet the demands of the field is limited. Here, we propose aMeta, an accurate metagenomic profiling workflow for ancient DNA designed to minimize the amount of false discoveries and computer memory requirements. Using simulated data, we benchmark aMeta against a current state-of-the-art workflow and demonstrate its superiority in microbial detection and authentication, as well as substantially lower usage of computer memory.

Keywords: Ancient metagenomics, Pathogen detection, Microbiome profiling, Ancient DNA

Background

Historically, ancient DNA (aDNA) studies have focused on human and faunal evolution and demography, extracting and analyzing predominantly eukaryotic aDNA [1–3]. With the development of next-generation sequencing (NGS) technologies, it was demonstrated that host-associated microbial aDNA from eukaryotic remains, which was previously treated as a sequencing by-product, can provide valuable information about ancient pandemics, lifestyle, and population migrations in the past [4–6]. Modern technologies have made it possible to study not only ancient microbiomes populating eukaryotic hosts, but also sedimentary ancient DNA (sedaDNA), which has rapidly become an independent branch of palaeogenetics, delivering unprecedented information about hominin and animal evolution without the need to analyze historical bones and teeth [7–12]. Previously available in microbial ecology, meta-barcoding methods lack validation and authentication power, and therefore, shotgun metagenomics has become the *de facto* standard in ancient microbiome research [13]. However, accurate detection,



aMeta: an accurate and memory-efficient ancient Metagenomic profiling workflow

≥6.16.0 Tests passing

About

aMeta is a Snakemake workflow for identifying microbial sequences in ancient DNA shotgun metagenomics samples. The workflow performs:

- trimming adapter sequences and removing reads shorter than 30 bp with Cutadapt
- quality control before and after trimming with FastQC and MultiQC
- taxonomic sequence kmer-based classification with KrakenUniq
- sequence alignment with Bowtie2 and screening for common microbial pathogens
- deamination pattern analysis with MapDamage2
- Lowest Common Ancestor (LCA) sequence alignment with Malt
- authentication and validation of identified microbial species with MaltExtract

When using aMeta and / or pre-built databases provided together with the workflow for your research projects, please cite our preprint: <https://www.biorxiv.org/content/10.1101/2022.10.03.510579v1>

Authors

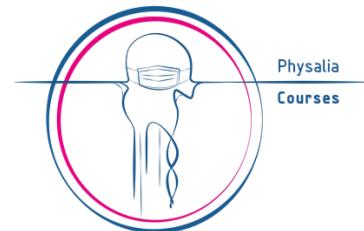
- Nikolay Oskolkov (@LeandroRitter) nikolay.oskolkov@scilifelab.se
- Claudio Mirabello (@clami66) claudio.mirabello@scilifelab.se
- Per Unneberg (@percyfal) per.unneberg@scilifelab.se

<https://github.com/NBISweden/aMeta>



About you

- Name
- University / Institute / Company
- Research interest(s)
- Previous experience with computational analysis and bioinformatics
- Motivation to join the course
- Expectations from the course

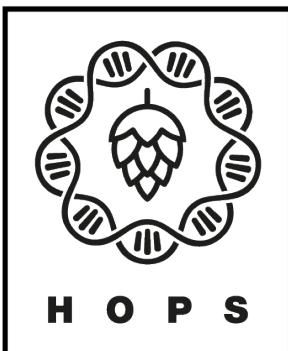


Typical analysis methods used in ancient metagenomics

1) Alignment:



BWA
stands for
Burrows Wheeler Aligner
 Abbreviations.com



2) Classification:



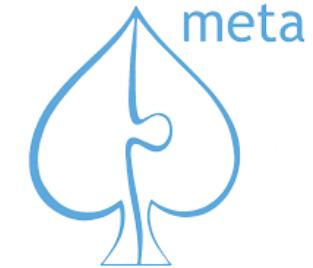
Centrifuge

MetaPhlan

Clark

Reference based:
assume similarity to reference

3) De-novo assembly:



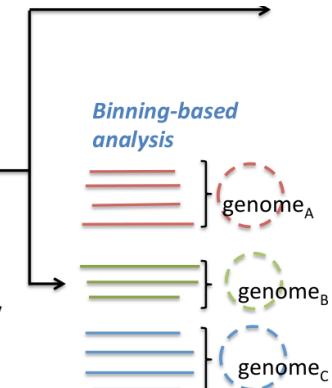
>seq1
GCCGTAGTCC...
>seq2
...



Assembly

Assembly-based
analysis

gene prediction/
annotation



Binning-based
analysis

genome_A

genome_B

genome_C

Phylogenetic binning

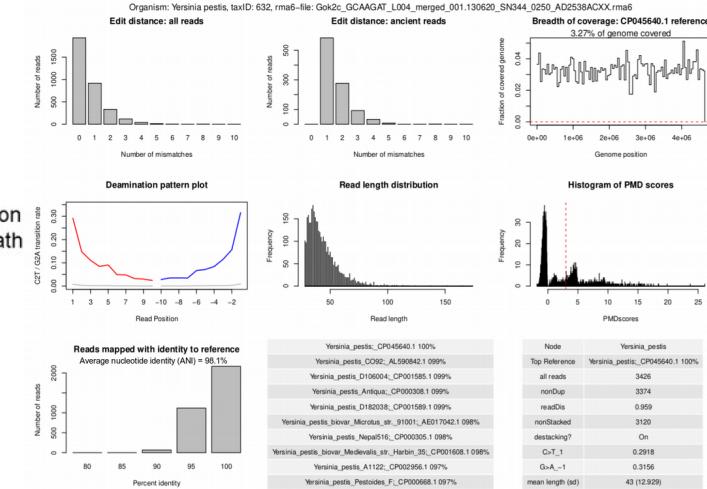
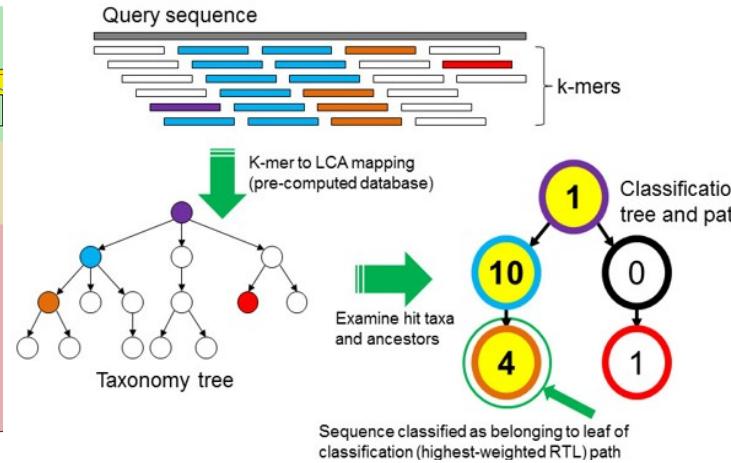
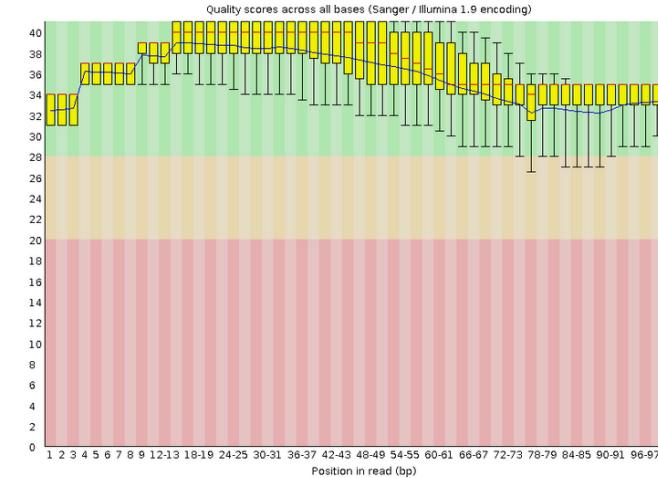
Reference free:
unbiased but challenging

Course outline

Day 1: introduction to ancient microbial and environmental metagenomics, quality control, adapter removal, taxonomic profiling

Day 2: alignment, assessing mapping quality, decontamination, source tracking and authentication analyses

Day 3: aMeta workflow, *de novo* assembly and authentication



- To ask questions please **raise your hand, unmute yourself and ask**. You can also ask questions in the **zoom chat** or the **slack channel** for this course.
- I would greatly appreciate if you **keep your camera on** as much as possible for better contact and communication.
- The course includes **6 lectures (~1h each) and 7 practicals (~1.5h each)**, there is a **15 min break** after each lecture and each practical
- During practicals, we will use **bash command line tools**. At the end of each practical, I will go through command lines with my explanations.
- The course assumes **some familiarity with basics of R / Python / aDNA**, you do not have to be fluent though. In case any terminology is unclear, please ask, I am ready to provide more introduction for non-experts.

Practical information: GitHub and Zoom

The course will take place in Zoom from 2 pm to 8 pm (CET, Berlin time)

Links to the Zoom room will be posted in the Slack channel

The course GitHub repository containing lectures and exercises is:

https://github.com/NikolayOskolkov/Physalia_AncientMetagenomics_2025

Please bookmark this address! To copy the material to your computer, please do:

git clone https://github.com/NikolayOskolkov/Physalia_AncientMetagenomics_2025





Code

Issues

Pull requests

Actions

Projects

Wiki

Security

Insights

Settings



Pin

Unwatch 1

Fork 0

Star 0

master · 1 Branch · 0 Tags

Go to file

Add file

Code

About



No description, website, or topics provided.

Readme

Activity

0 stars

1 watching

0 forks

Releases

No releases published

[Create a new release](#)

Packages

No packages published

[Publish your first package](#)

Languages

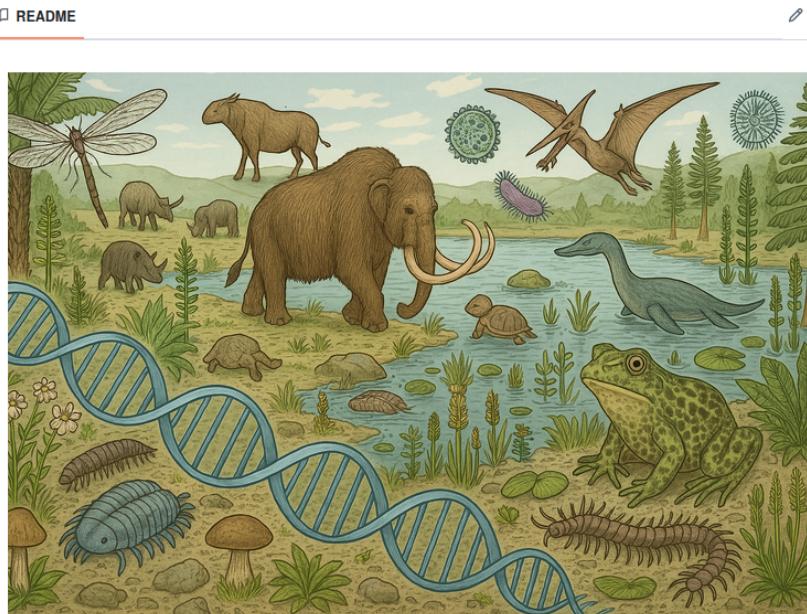


Suggested workflows

Based on your tech stack

**SLSA Generic generator**[Configure](#)

Generate SLSA3 provenance for your existing release workflows

**Jekyll using Docker image**[Configure](#)Package a Jekyll site using the `jekyll/builder` Docker image.[More workflows](#)[Dismiss suggestions](#)

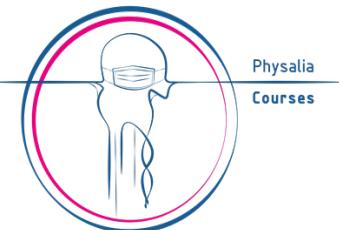
Practical information: Amazon Cloud (AWS EC2)



We will use the Cloud Computing service from Amazon, which we will access via `ssh` (secure shell protocol)

```
ssh -i ameta.pem -X ubuntu@54.244.40.220
```

- The IP address changes every day
- Everyone is given a username, with a `home` and `shared` folders
 - List of usernames can be found in Slack
 - The `shared` folder is copy-only: do not delete, move, rename, or write





File Edit View Search Terminal Help

(base) nikolay@dell:~/Documents/Consultation/Physalia/ancient_metagenomics\$ ssh -i ameta25.pem ubuntu@54.202.33.70
Welcome to Ubuntu 24.04.1 LTS (GNU/Linux 6.8.0-1021-aws x86_64)

* Documentation: <https://help.ubuntu.com>
* Management: <https://landscape.canonical.com>
* Support: <https://ubuntu.com/pro>

System information as of Thu May 22 08:18:12 UTC 2025

System load: 0.0 Processes: 165
Usage of /: 18.8% of 967.64GB Users logged in: 0
Memory usage: 1% IPv4 address for enX0: 172.31.36.154
Swap usage: 0%

Expanded Security Maintenance for Applications is not enabled.

63 updates can be applied immediately.
To see these additional updates run: apt list --upgradable

8 additional security updates can be applied with ESM Apps.
Learn more about enabling ESM Apps service at <https://ubuntu.com/esm>

Last login: Thu May 22 08:18:13 2025 from 130.235.244.193

(base) ubuntu@ip-172-31-36-154:~\$ ll
total 396508
drwxr-xr-x 14 ubuntu ubuntu 4096 May 21 14:59 ./
drwxr-xr-x 40 root root 4096 Jan 8 13:13 ../
-rw----- 1 ubuntu ubuntu 61 Jan 27 11:34 .Xauthority
-rw----- 1 ubuntu ubuntu 20865 May 21 16:25 .bash_history
-rw-r--r-- 1 ubuntu ubuntu 220 Mar 31 2024 .bash_logout
-rw-r--r-- 1 ubuntu ubuntu 4333 Jan 8 11:11 .bashrc
drwx----- 8 ubuntu ubuntu 4096 May 21 12:52 .cache/
drwxrwxr-x 2 ubuntu ubuntu 4096 Jan 9 12:13 .conda/
drwx----- 6 ubuntu ubuntu 4096 Jan 28 14:35 .config/
drwxrwxr-x 4 ubuntu ubuntu 4096 May 21 12:52 .java/
-rw----- 1 ubuntu ubuntu 40 Jan 28 11:48 .lessht
drwxrwxr-x 3 ubuntu ubuntu 4096 Jan 8 11:24 .local/
drwxrwxr-x 3 ubuntu ubuntu 4096 May 21 12:51 .parallel/
-rw-r--r-- 1 ubuntu ubuntu 807 Mar 31 2024 .profile
-rw----- 1 ubuntu ubuntu 11 Jan 8 10:30 .python_history
drwx----- 2 ubuntu ubuntu 4096 Jan 8 09:52 .ssh/
-rw-r--r-- 1 ubuntu ubuntu 0 Jan 8 10:12 .sudo_as_admin_successful
-rw----- 1 ubuntu ubuntu 16260 Jan 28 18:58 .viminfo
-rw-rw-r-- 1 ubuntu ubuntu 381 May 21 15:28 .wget-hsts
-rw-rw-r-- 1 ubuntu ubuntu 147784736 Jan 8 11:08 Miniconda3-latest-Linux-x86_64.sh
drwxr-xr-x 3 ubuntu ubuntu 4096 Jan 8 13:26 R/
drwxr-xr-x 4 ubuntu ubuntu 4096 May 21 15:04 Share/
drwxrwxr-x 12 ubuntu ubuntu 4096 May 21 12:44 aMeta/
-rw-rw-r-- 1 ubuntu ubuntu 1318204 Apr 21 2020 libssl1.1_1.1.1f-1ubuntu2_amd64.deb
drwxrwxr-x 19 ubuntu ubuntu 4096 Jan 8 11:11 miniconda3/
-rw-r--r-- 1 ubuntu ubuntu 574 Jan 13 10:57 r_packages_install.R
-rw-rw-r-- 1 ubuntu ubuntu 128381702 Dec 16 19:50 rstudio-server-2024.12.0-467-amd64.deb
-rw-rw-r-- 1 ubuntu ubuntu 128381702 Dec 16 19:50 rstudio-server-2024.12.0-467-amd64.deb.1
-rw-rw-r-- 1 ubuntu ubuntu 205 Jan 8 14:42 cat_passwords.sh*