

Standards,  
Precautions &  
Advances in  
Ancient  
Metagenomics

SPAAM2: Final stretch!

---



# Session Summaries

## Session 1

- Identified research gaps relative to in the **data generation process can cause biases in metagenomic data**. Remain to be tested: different extraction methods? lab-derived specific contamination? Potential common projects to fill these gaps: same samples, same protocols, comparison of the data. To discuss.
- Acknowledged the **necessity** (monitoring contamination) and **limitations** (cost, low biomass) of including and **sequencing blanks**. Identified a need for a consensus relative to how to sequence blanks (number of cycles, depth) and how to upload sequencing data onto public repositories.
- Presented our **sampling kits**: variability relative to sampling sites (field vs museums), sample types (sediment cores, human/animal remains) and time constraints. Optimized to limit contamination. Identified the need to report detailed sampling protocols in papers.
- Discussed the **pros** (cleans data for downstream genotype calling) and **cons** (removes damage signal used to identify damaged reads) of **UDG treatment**. Slack pool (N=22): 23% USER+, 50% USER-, 27% Sometimes



# Session Summaries

## Session 2

- Genotyping: trying to identify majority call (string of SNPs) or looking at gene content
- Reference bias, try new approaches such as vg graphs. Map to multiple references. Lack of modern genomes for comparison is often a problem.
- Using probabilities of variant calls instead of majority call. This is something we could shift to doing, but many tools, e.g. BEAST are not currently able to run on this type of input
- Lack of understanding of our field from experts outside of aDNA field (aDNA treated as curiosity), makes addressing reviews difficult sometimes, inclusion of experts on specific microbes may help.
- Community standards not an option yet, too much variation between projects and microbes needed, we are still experimenting. Focus instead on reproducibility

# Session Summaries

## Session 3

- Non-mapDamage methods/characteristics are of interest
  - Particularly sedaDNA
  - Control for batch-effects
- Databases: a lot of discomfort in what to build?
- Differences in computing resource size
- Currently not-clear common authentication criteria other than damage
  - But requires nuances

## Session 4

- Ethical aspects of a project need to be taken into account at the start of project design
- Wide variation in sample metadata reporting
  - uniformization through the Dir
- Testing of tools is difficult
  - lack of standardized datasets and outcomes
  - lack of PI support
- Data publication/availability is not as open as it could be
  - project-specific reasons
- Code and protocol sharing are necessary
  - still need encouragements/enforcing
- Outcomes: CAMI challenge? Open letter about data availability?

# Session Summaries

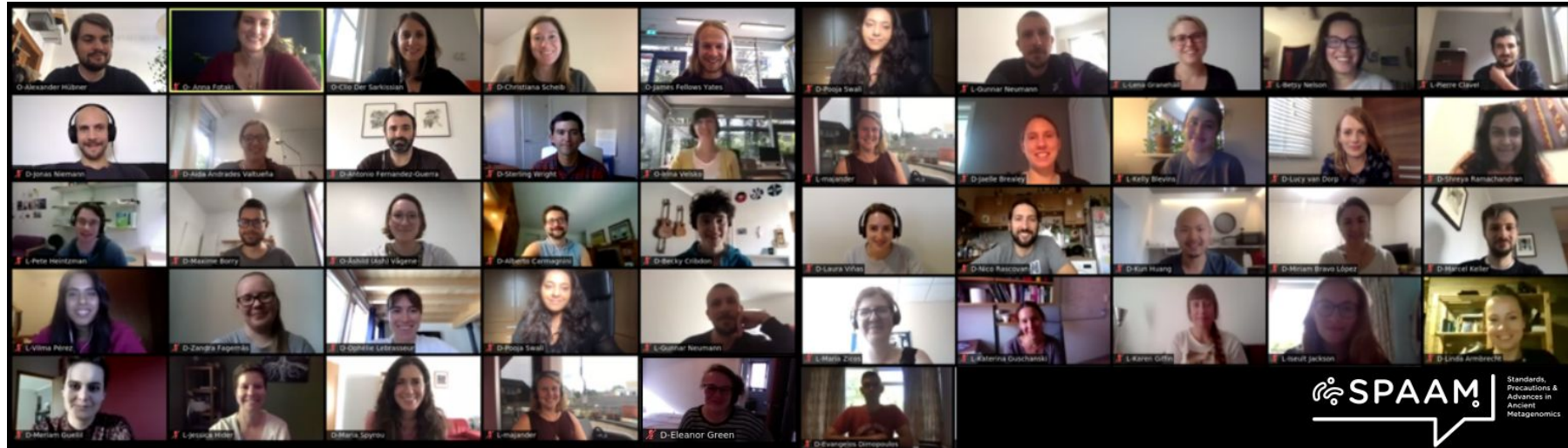
## Session 5

- *De novo* assembly main approach of interest (both application and testing, but latter still needed):
  - challenge is short reads and low coverage
  - damage not really an issue (when enough coverage)?
- A lot of interest in training workshops on assembly techniques and open discussion forums (ie slack) for support in out-of-the box analysis
- sedaDNA people still struggling for ‘established’ approaches on their weird-and-wonderful data



Next Steps

# Can we tweet the group picture?



Anyone missing?

# Outcomes

- Build networks! Talk to each other!

- Formalise - ICSB COSI?

<https://www.iscb.org/iscb-committees/cosi>

- Organically make projects

- Common training?
- Software?
- Working groups?
- Method development?
- Replication?
- Journal clubs?
- Pre-peer review?



<http://gph.is/2lmxD2k>



<http://gph.is/2cKreKv>

- Regular meetings?

- SPAAM3?

- Notes

- Organisers have been taking notes for all sessions
- All attendees who turned up (both tiers) listed
- Will make available with tweet-allowed slides via Google Docs link for one week for additional comments!
- Archived on Zenodo with DOI & put on website

- Opinion Paper

- Current practical challenges in the field
- Based on notes
- Organisers will draft on github privately
- Contributions accepted after (all contributors listed on github can be co-author)!



# Working Committees? Project Proposals? What do you want?



<https://gph.is/g/ZOkR2DY>

# SPAAM2 Social Meeting up? SPAAM3?



<https://gph.is/g/E3e3vOZ>

# Paper procedure?



<http://gph.is/1hekooZ>

# One last thing



<http://gph.is/1FCbg7c>



<http://gph.is/2vyeBXU>