



# Ancient Metagenomic Studies: Considerations for the Wider Scientific Community

Clio Der Sarkissian, Irina M Velsko, Anna K Fotakis, Åshild J Vågene,  
Alexander Hübner, James A Fellows Yates

## ► To cite this version:

Clio Der Sarkissian, Irina M Velsko, Anna K Fotakis, Åshild J Vågene, Alexander Hübner, et al.. Ancient Metagenomic Studies: Considerations for the Wider Scientific Community. mSystems, 2021, 6 (6), 10.1128/msystems.01315-21 . hal-03533810

**HAL Id: hal-03533810**

**<https://hal.science/hal-03533810>**

Submitted on 19 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Ancient Metagenomic Studies: Considerations for the Wider Scientific Community

 Clio Der Sarkissian,<sup>a</sup>  Irina M. Velsko,<sup>b</sup>  Anna K. Fotakis,<sup>c</sup>  Åshild J. Vågene,<sup>d</sup>  Alexander Hübner,<sup>b,e</sup>  James A. Fellows Yates<sup>b,f,g</sup>

<sup>a</sup>Centre for Anthropobiology and Genomics of Toulouse, UMR5288, CNRS, University Toulouse 3 Paul Sabatier, Toulouse, France

<sup>b</sup>Department of Archaeogenetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

<sup>c</sup>Center for Evolutionary Hologenomics, GLOBE Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

<sup>d</sup>Section for Evolutionary Genomics, GLOBE Institute, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

<sup>e</sup>Faculty of Biological Sciences, Friedrich-Schiller University Jena, Jena, Germany

<sup>f</sup>Institut für Vor- und Frühgeschichtliche Archäologie und Provinzialrömische Archäologie, Ludwig Maximilian University Munich, Munich, Germany

<sup>g</sup>Department of Paleobiotechnology, Leibniz Institute for Natural Product Research and Infection Biology Hans Knöll Institute, Jena, Germany

Clio Der Sarkissian, Irina M. Velsko, Anna K. Fotakis, Åshild J. Vågene, Alexander Hübner, and James A. Fellows Yates wrote the manuscript together. All authors contributed equally, and author order was selected by lottery.

**ABSTRACT** Like modern metagenomics, ancient metagenomics is a highly data-rich discipline, with the added challenge that the DNA of interest is degraded and, depending on the sample type, in low abundance. This requires the application of specialized measures during molecular experiments and computational analyses. Furthermore, researchers often work with finite sample sizes, which impedes optimal experimental design and control of confounding factors, and with ethically sensitive samples necessitating the consideration of additional guidelines. In September 2020, early career researchers in the field of ancient metagenomics met (Standards, Precautions & Advances in Ancient Metagenomics 2 [SPAAM2] community meeting) to discuss the state of the field and how to address current challenges. Here, in an effort to bridge the gap between ancient and modern metagenomics, we highlight and reflect upon some common misconceptions, provide a brief overview of the challenges in our field, and point toward useful resources for potential reviewers and newcomers to the field.

**KEYWORDS** ancient metagenomics, ancient DNA, archaeogenomics, paleogenetics, archaeology, microbiome, metagenomics, microbe, authentication, ethics, cultural heritage, peer review

## BACKGROUND

Ancient metagenomics refers to the analysis of the complex DNA content recovered from degraded, nonliving, biological material (e.g., bones, teeth, dental plaque, paleofeces, sediments), primarily via shotgun high-throughput sequencing. Research often focuses on ancient microbes (1, 2) but increasingly also on the simultaneous analysis of numerous (extinct) macroorganisms (see, e.g., references 3 to 5). Research questions in the field are often highly interdisciplinary, spanning the humanities and social and natural sciences, adding new perspectives to our understanding of the past, e.g., characterizing causative candidates of historical epidemics (6), identifying oral microbes in Neanderthals (7), reconstructing postglacial animal and plant successions in North America (8), and integrating detailed social and cultural contexts from archaeology and history (9, 10). Analyzing ancient metagenomes can provide a wide range of exciting and complementary sources of information for modern metagenomic studies, such as millennial-scale insights into the dynamics of metagenomes and the evolution of microbial species through deep-time calibration points for phylogenetic analyses (11–13). However, ancient metagenomics faces particular logistical and molecular challenges in addition to those commonly met when analyzing modern samples.

**Editor** Jack A. Gilbert, University of California San Diego

**Copyright** © 2021 Der Sarkissian et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to James A. Fellows Yates, james\_fellows\_yates@eva.mpg.de.

The authors declare no conflict of interest.

**Published** 21 December 2021

**TABLE 1** Common terms used in ancient metagenomics

Term	Description	Relevant literature
SPAAM	Standards, Precautions, and Advances in Ancient Metagenomics, a community of researchers in ancient metagenomics who run regular discussion meetings and community projects.	<a href="https://spaam-community.github.io">spaam-community.github.io</a>
Ancient	An organism or tissue is often considered to qualify as ancient when it is more than 100 years old. This cutoff, however, is arbitrary.	
Dental calculus	Mineralized dental plaque, also known as dental tartar, contains the petrified remains of an oral biofilm of microorganisms and other micro- and macroremains derived from the oral cavity.	7, 47, 55, 56
Paleofeces	Ancient human or animal feces that remain in an organic or partially organic state. Completely mineralized fecal remains are termed coprolites (however, note that the terms are often used interchangeably in archaeology).	57
Historic specimen	Any specimen that is no longer living and can no longer undergo molecular repair mechanisms. Sometimes used to distinguish natural history museum specimen collections aged less than 200 years old from mostly mineralized tissues found in archaeological excavations older than 200 years.	
Bio-cultural heritage	Biological specimens, tissues, or secondary substrates (e.g., dental calculus and paleofeces) or the habitats that are derived from (human) culture that are of cultural significance to a society or individual.	58
Necrobiome	The (micro)organisms that decompose dead organic materials, mostly referring to dead bodies.	59–61
Postdepositional	Changes or contamination affecting the specimen after deposition in the ground or environmental context not used during life (e.g., burial).	
Degradation	The process of biomolecules being broken up and damaged through a variety of chemical and mechanical processes.	62
Molecular preservation	State of the chemical preservation of the biomolecules, used to evaluate the feasibility of cost-effective hybridization capture, sequencing, and analysis.	
Destructive sampling	Sampling that removes part of the sample or specimen such that the removed/collected part cannot be returned after analysis.	
Ancient DNA	DNA from deceased organisms that has undergone some form of postmortem degradation processes (e.g., fragmentation, deamination damage). Often referred to as aDNA.	63, 64
Fragmentation	Breakage of DNA backbones that normally occurs at sites of base depurination and often results in single-stranded overhangs. Over time, aDNA reads become very short, typically 30–70 base pairs (bp).	65, 66
Misincorporation	Higher frequency of cytosine deamination-derived uracils (C-to-U) that are read as T by nonproofreading polymerases at DNA termini, caused by hydrolysis. This occurs on exposed single-stranded overhangs of fragmented DNA.	65, 66
Damage pattern	Can refer to either nucleotide misincorporation or fragment length distributions or both.	67, 68
Contamination	Ancient and modern DNA not deriving from the original organism or sample of interest (e.g., from the burial environment, museum collection, laboratory environment, curators, researchers).	41, 69
Endogenous DNA	DNA originally derived from the sample or tissue type that does not come from contaminating sources.	
Authentication	Determining whether a given set of DNA molecules is truly ancient and belonging to, or originating from, the sample in question. Normally based on characteristic damage patterns and endogenous content.	
Negative controls	Controls that do not contain samples for analysis (typically water) to act as indicators for the laboratory performance of reagents, or protocols and cross-contamination from the laboratory or samples in the same batch.	41
Radiocarbon dating	A means of annual age determination by measuring the decay of radioactive carbon, which occurs at a consistent rate. Also known as C-14 or carbon dating, with dates typically reported as “before present,” where “present” is usually defined as approximately 1950.	70
FAIR principles	A set of guidelines that focus on sharing data and metadata in a way that is machine readable, easily findable, and reusable to promote reproducibility.	27

These challenges stem from the facts that ancient samples are often rare and precious and that their molecular content is often highly degraded (see Table 1 for definitions of common terms used in the field); the latter makes ancient DNA (aDNA) sensitive to contamination by modern DNA from natural, storage, and experimental environments, including from handling (a selection of review papers can be found in Table 2). To better characterize and mitigate challenges in the field through collaboration, ancient metagenomic researchers organized a workshop for Standards, Precautions, and Advances in Ancient Metagenomics (SPAAM) in 2016 (1). This served as an inspiration to formally establish the SPAAM community in 2020 (<https://spaam-workshop.github.io/>).

**TABLE 2** Nonexhaustive list of suggested reviews, laboratory comparisons, dedicated analysis tools, and benchmarking studies relevant for ancient metagenomics

Publication type	Description	Reference
General review	General but comprehensive introduction to microbial archaeology, including ancient metagenomics and pathogen reconstruction and their challenges. Output of the first SPAAM meeting in 2016.	1
	Descriptions of the challenges related to contamination in next-generation sequencing data sets of low-biomass samples, including ancient DNA.	71
	Review of approaches developed for estimating the level of contamination in ancient human DNA data sets, covering aspects also relevant for ancient metagenomics.	72
Laboratory protocol	State-of-the-art summary of ancient pathogen research and what can be learned from such genomes.	73
	Comparison of decontamination and aDNA extraction protocols for ancient dental calculus.	47
	Comparison of decontamination protocols for ancient mammalian dental calculus.	74
	Comparison of decontamination protocols for ancient human dental calculus.	75
	Comparison of different protocols for simultaneous extraction of different ancient biomolecules from dental calculus.	76
	Comparison of DNA extraction methods for paleofeces, including commonly used modern DNA extraction kits.	39
	Comparison of microbial genome enrichment techniques for ancient pathogens.	77
	Development of techniques for aDNA retrieval from cave sediments and mammalian DNA capture techniques.	3
	Extraction method for retrieval of eukaryotic aDNA from marine sediment.	78
	Capture protocol for enrichment of marine eukaryotic aDNA.	79
Computational tool	gargammel: synthetic ancient DNA data set generation with a damage and contamination modeling tool.	54
	MALT: includes a description of the ultrafast BLAST-like metagenomic aligner MALT, which includes adaptations to account for ancient DNA damage. Often used for taxonomic profiling or pathogen detection.	6
	PIA: a taxonomic read biner used to identify the likely host source of typically sedimentary aDNA (sedaDNA) reads, with a focus on accounting for extinct taxa not present in modern databases.	80
	SNP_Evaluation: a tool allowing for evaluating and visualizing confidence in variant calling of low-coverage pathogen genomes.	81
	MEx-IPA: an interactive viewer of output from the HOPS metagenomic authentication pipeline.	82
	cuperdec: an R package for the estimation and visualization of the endogenous taxonomic content of ancient microbiomes.	82
	PyDamage: a tool for separating ancient and modern contigs from <i>de novo</i> assembly of ancient DNA data.	43
Pipeline	Holi: a pipeline for taxonomic profiling of ancient metagenomic reads based on competitive mapping to large databases. Often used to profile ancient environmental samples (e.g., sedaDNA) for reads assigned to animal or plant taxa.	8
	metaBIT: the first pipeline with configurations for high-throughput ancient metagenomic screening with MetaPhlAn and a range of taxonomic profile comparison analyses.	83
	coproID: a pipeline for the prediction of the host organism of ancient fecal material, including taxonomic profiling of the endogenous content of both host and microbial DNAs.	84
	HOPS: a pipeline integrating MALT alignment with postalignment ancient DNA characteristic authentication. Includes damage profiling, fragment length, and visualization of possible contamination.	85
	nf-core/eager: a general ancient DNA genomics pipeline with taxonomic profiling for pathogen screening (and microbiome reconstruction) and authentication components, as well as steps allowing for analysis of microbial genomes.	86
	Comparison of different metagenomic taxonomic classifiers applied to ancient DNA data, with description of effects of aDNA damage and short fragment lengths.	87
	Comparison of the effects of nucleotide-to-nucleotide vs nucleotide-to-protein taxonomic classification for short fragment lengths.	88

The SPAAM2 meeting in September 2020 gathered 63 early career researchers representing 36 institutions from 15 countries with highly diverse and interdisciplinary backgrounds (14). During discussions, it became evident that this young field would greatly benefit from building more bridges and improving communication with metagenomic or microbiology researchers working with modern samples. This was particularly brought to light in reviews of grant applications and publications. In this opinion/hypothesis, we highlight a few common misconceptions and point toward useful resources and field basics aimed at potential reviewers, as well as newcomers to the field. Additionally, we reflect on current research in our field and how we can improve on existing challenges to maximize the potential that ancient metagenomics offers.

### SAMPLING STRATEGIES IN AN ANCIENT METAGENOMIC STUDY

Small sample sizes are a consistent feature of ancient metagenomic studies, meaning that there is often little opportunity for large-scale sampling or elaborate sampling design,

as might be expected by reviewers with a background in modern metagenomics. There are several factors at play, including (i) variability in what was recovered from archaeological excavations and saved from collaborations between curators, archaeologists, and other legal representatives who safeguard the material and (ii) what has sufficient molecular preservation for analysis. These are limiting factors that are, largely, out of the control of researchers working on ancient or historical specimens. In particular, the preservation of the physical samples as well as the molecular preservation are strongly affected by the conditions of both the burial and storage environments. Therefore, deeper sequencing can be financially prohibitive, and increasing sample sizes retroactively from the same archaeological or museum collection or resampling previously accessed samples for replicates is not trivial, if even possible. This is partly due to sample accessibility but also due to “destructive sampling”; samples are often small amounts of material that cannot be regenerated like a microbial culture, nor can they be replaced. Due to the rarity and often small quantity of ancient remains, efforts are in fact made to avoid unnecessary, redundant subsampling by restricting access to these materials (15).

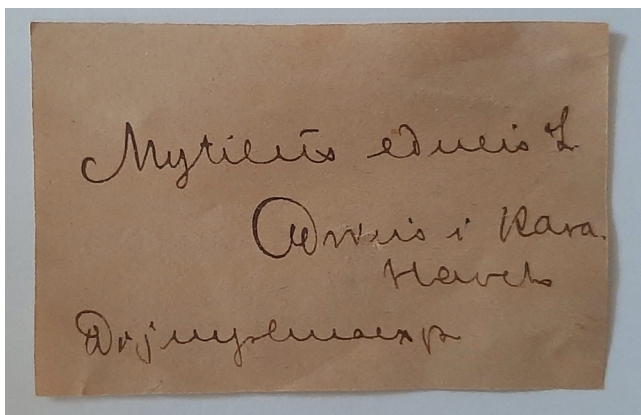
While ancient metagenomic researchers generally aim for high-quality sequencing data sets, reviewers should consider whether results and conclusions are supported by the data as they stand within this context. As more ancient metagenomic papers are published, researchers will be able to increase study sample sizes by incorporating previously published data sets and thus mitigate the effect of small sample size. One such example is the unexpected discovery of early forms of *Yersinia pestis* in only 7 out of 101 Bronze Age individuals, with the disease being argued as “endemic” despite geographic sparsity (11). However, this has since been repeatedly verified and expanded, with an increasing number of genomes recovered from the same and new regions (16–20).

Power analyses are likely to indicate that ancient metagenomic studies are always underpowered, but when performed prior to study design, they may clarify what questions can be answered with the data set at hand and guide the study's focus. Further, involving statisticians during study design is encouraged to determine the best methods to test the validity of their findings for robust verification. The questions being asked of ancient data sets need to be tailored for limited sample sizes, and interpreters of results should not make sweeping or broad claims that cannot be accurately extrapolated beyond the current study. As an example, Velsko et al. (21) collected extensive metadata regarding calculus deposition patterns to perform association tests with the calculus taxonomic profiles. However, it became clear that there were insufficient sample numbers in each category to draw statistically supported conclusions about such associations, and instead, the authors then refocused the paper to address other questions. Finally, to mitigate the bias of small sample sizes, the use of appropriate sample randomization for laboratory processing to avoid batch effects, and the sequencing of extraction and library processing negative control experiments need to be performed.

## ETHICAL CONSIDERATIONS AND CULTURAL SIGNIFICANCE

Ancient metagenomicists are faced with specific limitations and requirements for responsible and ethical research conduct when applying (partially) destructive analyses to archaeological materials. These materials are indeed scarce, irreplaceable, and constituent of our bio-cultural heritage, requiring their historical, cultural, social, and political background to be considered, especially those derived from extremely sensitive contexts, e.g., colonialism, racism, exploitation, or stigmatization. Ethical considerations can also apply to environmental samples, such as sediments, in particular those linked with past or present indigenous communities (22).

Currently, there is no internationally agreed-upon framework that specifically oversees genetic work on ancient specimens, and the responsibility of ensuring ethical ancient metagenomic research admittedly falls upon us and collection/material curators. Agreements such as the Convention on Biological Diversity (<https://www.cbd.int/>) and the subsequent Nagoya protocol (89) exist for genetic and biological material; however, these remain ambiguous with regard to ancient material. There is currently no precedent for intellectual property rights



**FIG 1** Historical label for a museum specimen at the Zoological Museum of Copenhagen, Denmark. Museum documentation enabled curators to decipher and translate the Danish inscription into "*Mytilus edulis* L. Drift Ice in the Kara Sea. The Dijnphna Expedition." This documentation, however, leaves scarce contextual information useful for downstream analyses and interpretations, such as precise collection date and geographical coordinates, environmental conditions of the find, and postcollection processing. (Photo courtesy of Clio Der Sarkissian.)

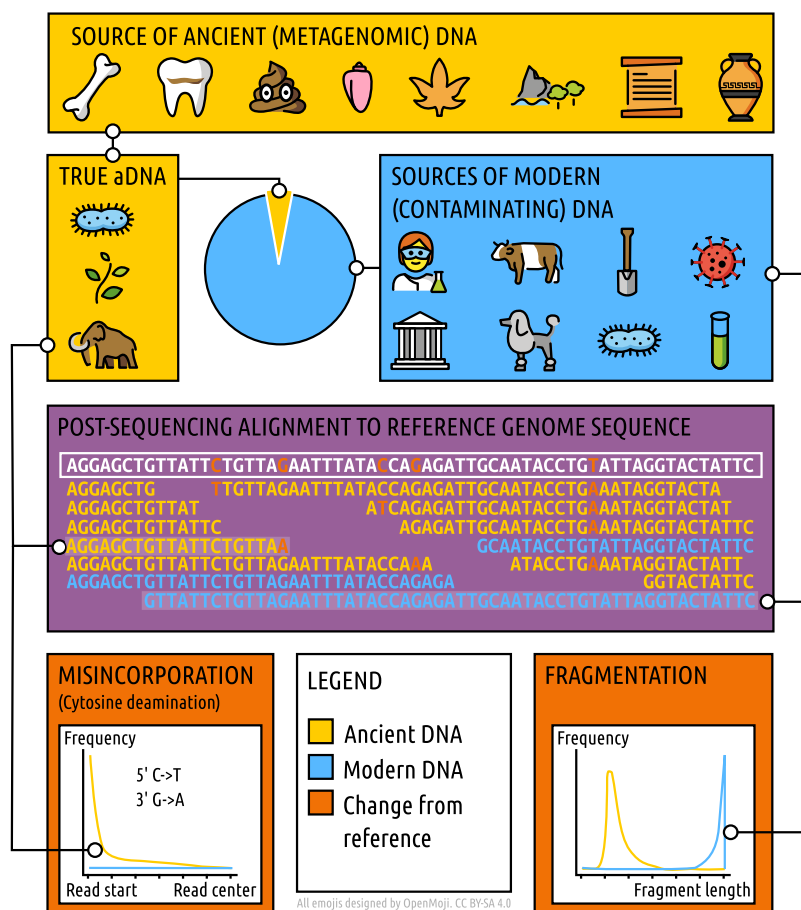
generated from ancient metagenomic studies. However, it is something of which the community is aware (14, 23, 24) and needs to be carefully considered in the future as studies move in this direction. A major role in the control of ethical practices relating to aDNA studies can also be played by other actors, such as reviewers and panels considering grant applications and project officers of funding agencies, as well as journal editors and reviewers of scientific articles. Such individuals may check for proof of ethical clearance for all samples included in studies, e.g., a letter of intent with stakeholders, decisions from relevant ethical review boards (approval numbers), or authorizations/permits to analyze or export from museums or collections.

For aDNA researchers, several action-based recommendations have been made to improve ethical practices (22, 25). Access to samples should be conditioned on an agreement after discussions with stakeholders from the scientific (e.g., archaeologists, anthropologists, collection curators) and/or civil (e.g., local communities, indigenous peoples, descendants) community. Agreements can include aspects such as the perimeter of the project (e.g., objectives, team composition, host institution, methods with regard to the state of the art, and funding), the active participation of local researchers and community members, the conditions for sample storage or repatriation upon study completion, and the strategy to communicate progress and results both throughout the project and when releasing final articles and data sets. Existing documentation for such agreements should be provided upon submission of grant proposals or scientific articles. Ethical procedures should be shared within and outside our community so as to encourage widespread ethical practices and transparency in ancient metagenomics. Additionally, researchers working in ancient metagenomics should be encouraged, when offered, to take or develop ethics courses.

### ESSENTIAL METADATA AND WHERE TO FIND IT

Ancient DNA researchers are also confronted with the erosion of the information necessary to rigorously analyze and interpret ancient metagenomic data in their full temporal, biological, social, and cultural contexts. Metadata from long-dead individuals is often compiled indirectly from the curated documents at hand for past collections and excavations or from reconstructions based on previous and ongoing anthropological, archaeological, and historical analyses. These metadata can often be incomplete, inconsistent, difficult to access, and sometimes not trivially standardized (26), especially for samples opportunistically collected in the past (see Fig. 1 as an example). Therefore, reviewers should pay attention to the representation of such important but limited metadata in manuscripts and alert researchers who fail to provide information in an accessible manner, i.e., in machine-readable file formats following the fair, accessible, interoperable, and reproducible (FAIR) principles (27). While





**FIG 2** Main characteristics and challenges of ancient metagenomic DNA. The original “endogenous” aDNA (i.e., the original DNA in the sample) of a wide range of sample types often makes up only a small fraction of the total genetic content of a sample. Contaminating environmental DNA from close relatives will often reduce confidence in taxonomic profiling and variant calling due to alignment of reads with variants from organisms other than the true ancient organism under investigation. True aDNA has a greater frequency of characteristic C-to-T deamination miscoding lesions at 5′ molecule termini (and a corresponding G-to-A mutation on the complementary 3′ strand of double-stranded libraries). True aDNA fragment distributions typically peak at short lengths, often around 30 to 70 bp, compared to the long lengths of modern DNA.

aDNA researchers have been celebrated for being at the forefront of raw data sharing (28), this is useful only if the corresponding metadata is made just as freely available.

To increase scientific yield from precious samples, it is important that ancient metagenomicists aim to maximize documentation prior to destruction (e.g., photography, photogrammetry, three-dimensional computed tomography [3D-CT] scanning, X-ray fluorescence [XRF] sediment scanning), to digitize and confirm information in historical museum records from other sources, or to combine complementary analyses (e.g., macrofossils, proteins, and stable isotopes, such as radiocarbon for dating) in their studies, project planning, and grant applications. The field should also come together to develop solutions to increase the linkability of these different types of data, allowing efficient incorporation into interdisciplinary analysis, but also standardizing reporting formats. Initiatives such as AncientMetagenomeDir, established by the SPAAM community (29), and developing context-specific “Minimum Information about any (x) Sequence” (MlxS) checklists (30), can act as templates for how this might be performed. By doing so, these initiatives will allow researchers to more efficiently compile relevant data for downstream analyses, as has been initiated in modern human microbiome research (31).

### COMPATIBILITY OF MOLECULAR LABORATORY METHODS

In a wet lab, the processing of ancient samples is challenging due to the degraded nature of aDNA molecules (Fig. 2). Standardization of optimized laboratory protocols

has allowed for the high throughput of many of the steps for successful sequencing of aDNA (3, 32–37). However, other methods that have gained popularity for modern sample types, such as Hi-C or long-read sequencing, require pure, high-quality, and long DNA molecules and therefore cannot be readily, if at all, applied to ancient samples.

Moreover, the features of aDNA mean that molecular tools, such as standardized DNA extraction kits that have been applied to most modern legacy data sets, e.g., the Human Microbiome Project (38), cannot be readily utilized for archaeological samples. While these commercial kits are optimized for the nuances of various modern sample types, even those optimized for low-biomass samples favor the retention of longer molecules and typically remove short fragments below 70 bp, making them inappropriate when targeting aDNA (Fig. 2; Table 1) (39). However, there is a growing number of studies comparing the suitabilities of different laboratory methods specifically for ancient metagenomics (Table 2).

Furthermore, while there are continuous improvements in molecular methods for limiting destruction and maximizing the output of molecular data from archaeological samples, these are often buried within large publications. As a field, we should therefore utilize resources such as protocols.io (<https://www.protocols.io/>; see, e.g., reference 40) or publish standalone method papers that would increase the visibility (and usability) of these improvements. Last, but not least, the low concentration of endogenous genetic material of our samples requires us to monitor lab contamination to confirm the validity of the results. Therefore, aside from having negative controls at each step, it is also more than necessary to publish these negative controls alongside the sequenced samples. This is in order to allow for a thorough evaluation of the results and improve reproducibility, as taxa present in the negative controls are sometimes removed from downstream sample data analyses.

## VALIDATION AND AUTHENTICATION OF ANCIENT DNA

A primary objective of any study of ancient metagenomics is to distinguish sequencing reads deriving from true aDNA molecules originally “endogenous” to the sample against a complex environmental background containing reads originating from both ancient and modern sources. We and others in the field (personal correspondence) have experienced reviews that outright dismiss ancient metagenomic studies as a novelty not to be taken seriously, despite the painstaking computational validation efforts included in the majority of studies.

As a first step, well-characterized degradation features of aDNA, i.e., damage patterns and high fragmentation, allow us to confidently identify true aDNA (Fig. 2) when it is combined with more general genomic reconstruction metrics, such as reasonable depth and breadth coverages. However, in cases of proposed exceptional preservation and in the absence of detectable or significant damage patterns, it is paramount that alternative lines of evidence for aDNA authenticity are combined. Other evidence can come from ancestral positioning and/or branch shortening of an ancient microbial genome in a phylogenetic tree, detection of established-species-specific genes or extrachromosomal genetic elements, radiocarbon dating of the material in question if the microbe is not associated with the environment (e.g., a host-restricted pathogen), or detection of microbial taxa reflecting the expected sample type (e.g., paleofeces contain predominantly gut taxa, instead of soil). Therefore, identifying sections in grants or publications that describe how aDNA will be, or has been, authenticated should be a critical part of any reviewer’s checklist. Several reviews of validation techniques and examples of newly developed methods that may be useful resources for reviewers have been published, and some of these are summarized in Table 2.

While the above aDNA validation methods are well established in paleogenomics, the highly complex nature of ancient metagenomic data may require additional authentication steps. For example, postdepositional contaminating DNA from the environment can also display damage, as it is subjected to the same degradational processes as the DNA of interest (41). Take taxa such as *Clostridium botulinum*, whose toxin causes deadly illnesses in humans but is a widely existing member of the soil/depositional environment



and necrobiome (42–44). Therefore, insufficient understanding of the biological context of different taxa may result in falsely attributing the source of an organism of interest (45).

Finding solutions to such problems remains a primary goal for the field over the coming years. These might be, for example, ensuring that studies include environmental controls, such as associated skeletal material (for example bone for dental calculus) (see, e.g., references 46 to 48) or soil (see, e.g., reference 6), when available. In addition, for studies targeting specific microbial species, more routinely developing collaborative projects and publications that include experts on modern strains would be highly beneficial.

## CONSIDERATIONS FOR DATA ANALYSIS

Implementing modern bioinformatics methods requires an increasing number of assumptions to be met regarding the underlying data, such as the expected sequencing error distribution, which is often violated by aDNA data. Therefore, these improvements, as hugely beneficial as they might be for modern data analysis, often cannot be directly applied to aDNA.

One prominent example is the calling of single nucleotide variants (SNVs) from high-throughput sequencing data. In the last decade, there have been large improvements in correctly identifying SNVs by determining haplotypes by local reassembly (GATK HaplotypeCaller [49]) or by the exact evaluation of haplotypes rather than read alignments (FreeBayes [50]). Neither can be directly applied to aDNA data due to the DNA's increased frequency of C-to-T substitutions caused by postmortem cytosine deamination and its, on average, short read length, which interferes with the correct haplotype inference. This often forces us to stick to older methods, even deprecated ones, such as GATK UnifiedGenotyper. Therefore, it is important that reviewers consider whether a method is actually applicable to aDNA data prior to suggesting its usage. Additionally, it should be taken into account that DNA preservation, and thus the available amount of unique endogenous DNA molecules, might further limit the possibility of performing analyses, such as strain analysis, that typically require high sequencing depth.

However, these limitations should not be used as a scapegoat to justify scientific lagging in our field, as methods tailored for aDNA analysis do exist. For example, damage-aware variant callers (e.g., snpAD [51], AntCaller [52], and ARIADNA [53]) have been developed, primarily for human population genomics, but could also potentially be tested for metagenomic contexts. Hence, it is very important that we do spend the appropriate amount of time and energy on validating our analyses to show that they do not suffer from biases introduced by unfit methods and on developing tailored software when necessary (examples of existing benchmarking studies, software, and pipelines can be found in Table 2). To test the suitability of a method, simulations of ancient metagenomic sequencing data with programs such as gargammel (54) can provide the required framework for such validations. Reviewers should thus consider requesting such validation to ensure the high quality of the analysis, while we should prioritize these types of analyses, improving quality over quantity.

## CONCLUSIONS

The analysis of complex DNA content via shotgun high-throughput sequencing, also known as metagenomics, has recently been used on a wide range of source materials in numerous studies. Although very similar, analyses of complex ancient DNA content face some additional challenges which might not be immediately apparent and have led to a number of misunderstandings in recent peer review processes. The majority of these are related to either the samples included in the study, such as the preservation of their DNA and their limited availability, or the methods used in the laboratory or during analysis, which are often unusual from the point of view of a reviewer but are necessary when considering the characteristics of aDNA. Despite its particularities, analyzing aDNA samples provides a rich and valuable resource of complementary information that will also help the field of modern metagenomics to advance in the future.

We hope that this opinion/hypothesis provides a useful resource for future reviewers to be able to thoroughly evaluate ancient metagenomic studies, with respect to both their ethical conduct and the scientific quality of the presented research, while improving awareness of the challenges that go hand in hand with such samples. The list of topics discussed here is, by far, not exhaustive, and we hope that through the Standards, Precautions, and Advances in Ancient Metagenomics (SPAAM) community, we can further extend it in the future and will be able to establish standards in our field that ensure the reproducibility and the interoperability of future work. One of the strengths of the SPAAM community is the diversity of its members, and this is highlighted by the high degree of interdisciplinarity of their research, which spans from microbiology to evolutionary biology and from archaeology to history. We therefore cordially invite modern metagenomicists to participate in SPAAM (<https://spaam-community.github.io/>) to help us improve the communication between our closely related fields and increase the scientific yield of ancient metagenomic studies in the future.

## ACKNOWLEDGMENTS

We thank Jaelle C. Brealey, Becky Cribdon, Katerina Guschanski, Ophélie Lebrasseur, Aleksandar D. Kostic, Frank Maixner, Mohamed S. Sarhan, Laura S. Weyrich, Marsha Wibowo, and Sterling L. Wright for providing reviews and summaries of common review critiques. We also thank Peter D. Heintzman, Nikolay Oskolkov, Adam Koziol, Jessica Hider, Katerina Guschanski, Ophélie Lebrasseur, Anan Ibrahim, Miriam J. Bravo-Lopez, Eleanor J. Green, Siobhan Mor, Aida Andrades Valtueña, Christina Warinner, and members of the Department of Paleobiotechnology at the Leibniz Institute for Natural Product Research and Infections Biology Hans Knöll Institute for comments on the manuscript. Furthermore, we thank the attendees of SPAAM2 and the SPAAM community. We also thank Tom Gilbert, Ludovic Orlando, and Christina Warinner for further comments on the manuscript and their support for early career researchers.

We declare that we have no competing interests.

I.M.V. was supported by the Werner Siemens-Stiftung (Paleobiotechnology, awarded to Pierre Stallforth, Hans-Knöll Institute, and Christina Warinner, Max Planck Institute for Evolutionary Anthropology). A.K.F. was supported by the Danish National Research Foundation award (CEH-DNRF143 awarded to M. Thomas P. Gilbert, University of Copenhagen, Denmark). Å.J.V. is supported by Carlsbergfondet Semper Ardens grant CF18-1109 (to M. Thomas P. Gilbert, University of Copenhagen, Denmark). A.H. is funded by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) under Germany's Excellence Strategy EXC 2051, project ID 390713860. J.A.F.Y. was supported by grant ERC-2015-StG 678901-FoodTransforms (to Philipp W. Stockhammer, Ludwig Maximilian University, Germany). J.A.F.Y., I.M.V., and A.H. are supported by the Max Planck Society. C.D.S. is supported by the Centre National de la Recherche Scientifique, France.

## REFERENCES

- Warinner C, Herbig A, Mann A, Fellows Yates JA, Weiß CL, Burbano HA, Orlando L, Krause J. 2017. A robust framework for microbial archaeology. *Annu Rev Genomics Hum Genet* 18:321–356. <https://doi.org/10.1146/annurev-genom-091416-035526>.
- Spyrou MA, Bos KI, Herbig A, Krause J. 2019. Ancient pathogen genomics as an emerging tool for infectious disease research. *Nat Rev Genet* 20: 323–340. <https://doi.org/10.1038/s41576-019-0119-1>.
- Slon V, Hopfe C, Weiß CL, Mafessoni F, de la Rasilla M, Lalueza-Fox C, Rosas A, Soressi M, Knul MV, Miller R, Stewart JR, Derevianko AP, Jacobs Z, Li B, Roberts RG, Shunkov MV, de Lumley H, Perrenoud C, Gušić I, Kučan Ž, Rudan P, Aximu-Petri A, Essel E, Nagel S, Nickel B, Schmidt A, Prüfer K, Kelso J, Burbano HA, Pääbo S, Meyer M. 2017. Neandertal and Denisovan DNA from Pleistocene sediments. *Science* 356:605–608. <https://doi.org/10.1126/science.aam9695>.
- Gelabert P, Sawyer S, Bergström A, Margaryan A, Collin TC, Meshveliani T, Belfer-Cohen A, Lordkipanidze D, Jakeli N, Matskevich Z, Bar-Oz G, Fernandes DM, Cheronet O, Özdoğan KT, Oberreiter V, Feeney RNM, Stahlschmidt MC, Skoglund P, Pinhasi R. 2021. Genome-scale sequencing and analysis of human, wolf, and bison DNA from 25,000-year-old sediment. *Curr Biol* 31:3564–3574.e9. <https://doi.org/10.1016/j.cub.2021.06.023>.
- Zavala EI, Jacobs Z, Vernot B, Shunkov MV, Kozlikin MB, Derevianko AP, Essel E, de Filippo C, Nagel S, Richter J, Romagné F, Schmidt A, Li B, O'Gorman K, Slon V, Kelso J, Pääbo S, Roberts RG, Meyer M. 2021. Pleistocene sediment DNA reveals hominin and faunal turnovers at Denisova Cave. *Nature* 595:399–403. <https://doi.org/10.1038/s41586-021-03675-0>.
- Vägene AJ, Herbig A, Campana MG, Robles García NM, Warinner C, Sabin S, Spyrou MA, Andrades Valtueña A, Huson D, Tuross N, Bos KI, Krause J. 2018. *Salmonella enterica* genomes from victims of a major sixteenth-century epidemic in Mexico. *Nat Ecol Evol* 2:520–528. <https://doi.org/10.1038/s41559-017-0446-6>.
- Weyrich LS, Duchene S, Soubrier J, Arriola L, Llamas B, Breen J, Morris AG, Alt KW, Caramelli D, Dresely V, Farrell M, Farrer AG, Francken M, Gully N, Haak W, Hardy K, Harvati K, Held P, Holmes EC, Kaidonis J, Lalueza-Fox C, de la Rasilla M, Rosas A, Semal P, Soltysiak A, Townsend G, Usai D, Wahl J, Huson DH, Dobney K, Cooper A. 2017. Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus. *Nature* 544:357–361. <https://doi.org/10.1038/nature21674>.

8. Pedersen MW, Ruter A, Schweger C, Friebe H, Staff RA, Kjeldsen KK, Mendoza MLZ, Beaudoin AB, Zutter C, Larsen NK, Potter BA, Nielsen R, Rainville RA, Orlando L, Meltzer DJ, Kjaer KH, Willerslev E. 2016. Postglacial viability and colonization in North America's ice-free corridor. *Nature* 537: 45–49. <https://doi.org/10.1038/nature19085>.
9. Ardelean CF, Becerra-Valdivia L, Pedersen MW, Schwenninger J-L, Oviatt CG, Macías-Quintero JI, Arroyo-Cabrales J, Sikora M, Ocampo-Díaz YZE, Rubio-Cisneros II, Watling JG, de Medeiros VB, De Oliveira PE, Barba-Pingarón L, Ortiz-Butrón A, Blancas-Vázquez J, Rivera-González I, Solís-Rosales C, Rodríguez-Ceja M, Gandy DA, Navarro-Gutiérrez Z, De La Rosa-Díaz JJ, Huerta-Arellano V, Marroquín-Fernández MB, Martínez-Riojas LM, López-Jiménez A, Higham T, Willerslev E. 2020. Evidence of human occupation in Mexico around the Last Glacial Maximum. *Nature* 584:87–92. <https://doi.org/10.1038/s41586-020-2509-0>.
10. Guellil M, Kersten O, Namouchi A, Luciani S, Marota I, Arcini CA, Iregren E, Lindemann RA, Warfvinge G, Bakanidze L, Bitadze L, Rubini M, Zai P, Zai M, Neri D, Stenseth NC, Bramanti B. 2020. A genomic and historical synthesis of plague in 18th century Eurasia. *Proc Natl Acad Sci U S A* 117: 28328–28335. <https://doi.org/10.1073/pnas.2009677117>.
11. Rasmussen S, Allentoft ME, Nielsen K, Orlando L, Sikora M, Sjögren K-G, Pedersen AG, Schubert M, Van Dam A, Kapel CMO, Nielsen HB, Brunak S, Avetisyan P, Epimakhov A, Khalyapin MV, Gnuni A, Kriiska A, Lasak I, Metspalu M, Moiseyev V, Gromov A, Pokutta D, Saag L, Varul L, Yepiskoposyan L, Sicheritz-Pontén T, Foley RA, Lahr MM, Nielsen R, Kristiansen K, Willerslev E. 2015. Early divergent strains of *Yersinia pestis* in Eurasia 5,000 years ago. *Cell* 163:571–582. <https://doi.org/10.1016/j.cell.2015.10.009>.
12. Duggan AT, Perdomo MF, Piombino-Mascoli D, Marciniak S, Poinar D, Emery MV, Buchmann JP, Duchêne S, Jankauskas R, Humphreys M, Golding GB, Southon J, Devault A, Rouillard J-M, Sahl JW, Dutour O, Hedman K, Sajantila A, Smith GL, Holmes EC, Poinar HN. 2016. 17th century variola virus reveals the recent history of smallpox. *Curr Biol* 26: 3407–3412. <https://doi.org/10.1016/j.cub.2016.10.061>.
13. Wibowo MC, Yang Z, Borry M, Hübner A, Huang KD, Tierney BT, Zimmerman S, Barajas-Olmos F, Contreras-Cubas C, García-Ortiz H, Martínez-Hernández A, Luber JM, Kirstahler P, Blohm T, Smiley FE, Arnold R, Ballal SA, Pamp SJ, Russ J, Maixner F, Rota-Stabelli O, Segata N, Reinhard K, Orozco L, Warinner C, Snow M, LeBlanc S, Kostic AD. 2021. Reconstruction of ancient microbial genomes from the human gut. *Nature* 594:234–239. <https://doi.org/10.1038/s41586-021-03532-0>.
14. SPAAM Community. 2020. SPAAM-community/spaam-event-notes: SPAAM2 notes. Zenodo <https://doi.org/10.5281/zenodo.4109609>.
15. Pálsdóttir AH, Bläuer A, Rannamäe E, Boessenkool S, Hallsson JH. 2019. Not a limitless resource: ethics and guidelines for destructive sampling of archaeofaunal remains. *R Soc Open Sci* 6:191059. <https://doi.org/10.1098/rsos.191059>.
16. Andrades Valtueña A, Mitnik A, Key FM, Haak W, Allmäe R, Belinskij A, Daubaras M, Feldman M, Jankauskas R, Janković I, Massy K, Novak M, Pfrangl S, Reinhold S, Šlaus M, Spyrou MA, Szécsényi-Nagy A, Törv M, Hansen S, Bos KI, Stockhammer PW, Herbig A, Krause J. 2017. The Stone Age plague and its persistence in Eurasia. *Curr Biol* 27:3683–3691.e8. <https://doi.org/10.1016/j.cub.2017.10.025>.
17. Spyrou MA, Tukhbatova RI, Wang C-C, Valtueña AA, Lankapalli AK, Kondrashin VV, Tsybin VA, Khokhlov A, Kühnert D, Herbig A, Bos KI, Krause J. 2018. Analysis of 3800-year-old *Yersinia pestis* genomes suggests Bronze Age origin for bubonic plague. *Nat Commun* 9:2234. <https://doi.org/10.1038/s41467-018-04550-9>.
18. Rascovan N, Sjögren K-G, Kristiansen K, Nielsen R, Willerslev E, Desnues C, Rasmussen S. 2019. Emergence and spread of basal lineages of *Yersinia pestis* during the Neolithic decline. *Cell* 176:295–305.e10. <https://doi.org/10.1016/j.cell.2018.11.005>.
19. Yu H, Spyrou MA, Karapetian M, Shnaider S, Radzevičiūtė R, Nägele K, Neumann GU, Penske S, Zech J, Lucas M, LeRoux P, Roberts P, Pavlenok G, Buzhilova A, Posth C, Jeong C, Krause J. 2020. Paleolithic to Bronze Age Siberians reveal connections with first Americans and across Eurasia. *Cell* 181:1232–1245.e20. <https://doi.org/10.1016/j.cell.2020.04.037>.
20. Susat J, Lübke H, Immel A, Brinker U, Macäne A, Meadows J, Steer B, Tholey A, Zagorska I, Gerhards G, Schmölcke U, Kalniņš M, Franke A, Petersone-Gordina E, Teßman B, Törv M, Schreiber S, Andree C, Bērziņš V, Nebel A, Krause-Kyora B. 2021. A 5,000-year-old hunter-gatherer already plagued by *Yersinia pestis*. *Cell Rep* 35:109278. <https://doi.org/10.1016/j.celrep.2021.109278>.
21. Velsko IM, Fellows Yates JA, Aron F, Hagan RW, Frantz LAF, Loe L, Martinez JBR, Chaves E, Gosden C, Larson G, Warinner C. 2019. Microbial differences between dental plaque and historic dental calculus are related to oral biofilm maturation stage. *Microbiome* 7:102. <https://doi.org/10.1186/s40168-019-0717-3>.
22. Handsley-Davis M, Kowal E, Russell L, Weyrich LS. 2021. Researchers using environmental DNA must engage ethically with Indigenous communities. *Nat Ecol Evol* 5:146–148. <https://doi.org/10.1038/s41559-020-01351-6>.
23. Curry A. 2021. Ancient poop reveals extinction in gut bacteria. *Science* 372: 671. <https://doi.org/10.1126/science.372.6543.671>.
24. SPAAM Community. 2021. SPAAM-community/spaam-event-notes: SPAAM3 notes. Zenodo <https://doi.org/10.5281/zenodo.5724028>.
25. Wagner JK, Colwell C, Claw KG, Stone AC, Bolnick DA, Hawks J, Brothers KB, Garrison NA. 2020. Fostering responsible research on ancient DNA. *Am J Hum Genet* 107:183–195. <https://doi.org/10.1016/j.ajhg.2020.06.017>.
26. Graham CH, Ferrier S, Huettman F, Moritz C, Peterson AT. 2004. New developments in museum-based informatics and applications in biodiversity analysis. *Trends Ecol Evol* 19:497–503. <https://doi.org/10.1016/j.tree.2004.07.006>.
27. Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3:160018. <https://doi.org/10.1038/sdata.2016.18>.
28. Anagnostou P, Capocasa M, Milia N, Sanna E, Battaglia C, Luzi D, Destro Bisol G. 2015. When data sharing gets close to 100%: what human paleogenetics can teach the open science movement. *PLoS One* 10:e0121409. <https://doi.org/10.1371/journal.pone.0121409>.
29. Fellows Yates JA, Andrades Valtueña A, Vågane ÅJ, Cribdon B, Velsko IM, Borry M, Bravo-Lopez MJ, Fernandez-Guerra A, Green EJ, Ramachandran SL, Heintzman PD, Spyrou MA, Hübner A, Gancz AS, Hider J, Allshouse AF, Zaro V, Warinner C. 2021. Community-curated and standardised meta-data of published ancient metagenomic samples with AncientMetagenomeDir. *Sci Data* 8:31. <https://doi.org/10.1038/s41597-021-00816-y>.
30. Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, Gilbert JA, Karsch-Mizrachi I, Johnston A, Cochrane G, Vaughan R, Hunter C, Park J, Morrison N, Rocca-Serra P, Sterk P, Arumugam M, Bailey M, Baumgartner L, Birren BW, Blaser MJ, Bonazzi V, Booth T, Bork P, Bushman FD, Buttigieg PL, Chain PSG, Charlson E, Costello EK, Huot-Creasy H, Dawyndt P, DeSantis T, Fierer N, Fuhrman JA, Gallery RE, Gevers D, Gibbs RA, San Gil I, Gonzalez A, Gordon JI, Guralnick R, Hankeln W, Highlander S, Hugenholtz P, Jansson J, Kau AL, Kelley ST, Kennedy J, Knights D, Koren O, et al. 2011. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MInX) specifications. *Nat Biotechnol* 29:415–420. <https://doi.org/10.1038/nbt.1823>.
31. Mirzayi C, Renson A, Zohra F, Elsafoury S, Geistlinger L, Kasselmann LJ, Eckenrode K, van de Wijgert J, Loughman A, Marques FZ, MacIntyre DA, Arumugam M, Azhar R, Beghini F, Bergstrom K, Bhatt A, Bisanz JE, Braun J, Bravo HC, Buck GA, Bushman F, Casero D, Clarke G, Collado MC, Cotter PD, Cryan JF, Demmer RT, Devkota S, Elinav E, Escobar JS, Fettweis J, Finn RD, Fodor AA, Forslund S, Franke A, Furlanello C, Gilbert J, Grice E, Haibe-Kains B, Handley S, Herd P, Holmes S, Jacobs JP, Karstens L, Knight R, Knights D, Koren O, Kwon DS, Langille M, Lindsay B, McGovern D, McHardy AC, McWeeney S, Mueller NT, Genomic Standards Consortium, et al. 2021. Reporting guidelines for human microbiome research: the STORMS checklist. *Nat Med* 27:1885–1888. <https://doi.org/10.1038/s41591-021-01552-x>.
32. Dabney J, Knapp M, Glocke I, Gansauge M-T, Weihmann A, Nickel B, Valdiosera C, García N, Pääbo S, Arsuaga J-L, Meyer M. 2013. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci U S A* 110: 15758–15763. <https://doi.org/10.1073/pnas.1314445110>.
33. Carøe C, Gopalakrishnan S, Vinner L, Mak SST, Sinding MHS, Samaniego JA, Wales N, Sicheritz-Pontén T, Gilbert MTP. 2018. Single-tube library preparation for degraded DNA. *Methods Ecol Evol* 9:410–419. <https://doi.org/10.1111/2041-210X.12871>.
34. Rohland N, Glocke I, Aximu-Petri A, Meyer M. 2018. Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. *Nat Protoc* 13:2447–2461. <https://doi.org/10.1038/s41596-018-0050-5>.
35. Gansauge M-T, Aximu-Petri A, Nagel S, Meyer M. 2020. Manual and automated preparation of single-stranded DNA libraries for the sequencing of DNA from ancient biological remains and other sources of highly degraded DNA. *Nat Protoc* 15:2279–2300. <https://doi.org/10.1038/s41596-020-0338-0>.

36. Suchan T, Kusliy MA, Khan N, Chauvey L, Tonasso-Calvière L, Schiavinato S, Southon J, Keller M, Kitagawa K, Krause J, Bessudnov AN, Bessudnov AA, Graphodatsky AS, Valenzuela-Lamas S, Wilczyński J, Pospula S, Tunia K, Nowak M, Moskal-delHoyo M, Tishkin AA, Pryor AJE, Outram AK, Orlando L. 28 September 2021. Performance and automation of ancient DNA capture with RNA hyRAD probes. *Mol Ecol Resour* <https://doi.org/10.1111/1755-0998.13518>.
37. Kapp JD, Green RE, Shapiro B. 2021. A fast and efficient single-stranded genomic library preparation method optimized for ancient DNA. *J Hered* 112:241–249. <https://doi.org/10.1093/jhered/esab012>.
38. Human Microbiome Project Consortium. 2012. Structure, function and diversity of the healthy human microbiome. *Nature* 486:207–214. <https://doi.org/10.1038/nature11234>.
39. Hagan RW, Hofman CA, Hübner A, Reinhard K, Schnorr S, Lewis CM, Jr, Sankaranarayanan K, Warinner CG. 2020. Comparison of extraction methods for recovering ancient microbial DNA from paleofeces. *Am J Phys Anthropol* 171:275–284. <https://doi.org/10.1002/ajpa.23978>.
40. Fellows Yates JA, Franziska A, Neumann GU, Velsko I, Skourtanioti E, Orfanou E, Fagernäs Z, Stahl R, Valtuena AA, Warinner C, Haak W, Brandt G. 2021. A-Z of ancient DNA protocols for shotgun Illumina next generation sequencing. *protocols* <https://doi.org/10.17504/protocols.io.bvt9n6r6>.
41. Llamas B, Valverde G, Fehren-Schmitz L, Weyrich LS, Cooper A, Haak W. 2017. From the field to the laboratory: controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era. *Sci Technol Archaeol Res* 3:1–14. <https://doi.org/10.1080/20548923.2016.1258824>.
42. Espelund M, Klaveness D. 2014. Botulism outbreaks in natural environments—an update. *Front Microbiol* 5:287. <https://doi.org/10.3389/fmicb.2014.00287>.
43. Borry M, Hübner A, Rohrlach AB, Warinner C. 2021. PyDamage: automated ancient damage identification and estimation for contigs in ancient DNA de novo assembly. *PeerJ* 9:e11845. <https://doi.org/10.7717/peerj.11845>.
44. Harrison L, Kooienga E, Speights C, Tomberlin J, Lashley M, Barton B, Jordan H. 2020. Microbial succession from a subsequent secondary death event following mass mortality. *BMC Microbiol* 20:309. <https://doi.org/10.1186/s12866-020-01969-3>.
45. Campana MG, Robles García N, Rühli FJ, Tuross N. 2014. False positives complicate ancient pathogen identifications using high-throughput shotgun sequencing. *BMC Res Notes* 7:111. <https://doi.org/10.1186/1756-0500-7-111>.
46. Schuenemann VJ, Singh P, Mendum TA, Krause-Kyora B, Jäger G, Bos KI, Herbig A, Economou C, Benjak A, Busso P, Nebel A, Boldsen JL, Kjellström A, Wu H, Stewart GR, Taylor GM, Bauer P, Lee OY-C, Wu HHT, Minnikin DE, Besra GS, Tucker K, Roffey S, Sow SO, Cole ST, Nieselt K, Krause J. 2013. Genome-wide comparison of medieval and modern *Mycobacterium leprae*. *Science* 341:179–183. <https://doi.org/10.1126/science.1238286>.
47. Warinner C, Rodrigues JFM, Vyas R, Trachsel C, Shved N, Grossmann J, Radini A, Hancock Y, Tito RY, Fiddymont S, Speller C, Hendy J, Charlton S, Luder HU, Salazar-García DC, Eppler E, Seiler R, Hansen LH, Castruita JAS, Barkow-Oesterreicher S, Teoh KY, Kelstrup CD, Olsen JV, Nanni P, Kawai T, Willerslev E, von Mering C, Lewis CM, Jr, Collins MJ, Gilbert MTP, Rühli F, Cappellini E. 2014. Pathogens and host immunity in the ancient human oral cavity. *Nat Genet* 46:336–344. <https://doi.org/10.1038/ng.2906>.
48. Mann AE, Sabin S, Ziesemer K, Vågane ÅJ, Schroeder H, Ozga AT, Sankaranarayanan K, Hofman CA, Fellows Yates JA, Salazar-García DC, Frohlich B, Aldenderfer M, Hoogland M, Read C, Milner GR, Stone AC, Lewis CM, Jr, Krause J, Hofman C, Bos KI, Warinner C. 2018. Differential preservation of endogenous human and microbial DNA in dental calculus and dentin. *Sci Rep* 8:9822. <https://doi.org/10.1038/s41598-018-28091-9>.
49. Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, Kling DE, Gauthier LD, Levy-Moonshine A, Roazen D, Shakir K, Thibault J, Chandran S, Whelan C, Lek M, Gabriel S, Daly MJ, Neale B, MacArthur DG, Banks E. 2018. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv* <https://doi.org/10.1101/201178>.
50. Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. *arXiv* <https://arxiv.org/abs/1207.3907>.
51. Prüfer K. 2018. snpAD: an ancient DNA genotype caller. *Bioinformatics* 34: 4165–4171. <https://doi.org/10.1093/bioinformatics/bty507>.
52. Zhou B, Wen S, Wang L, Jin L, Li H, Zhang H. 2017. AntCaller: an accurate variant caller incorporating ancient DNA damage. *Mol Genet Genomics* 292:1419–1430. <https://doi.org/10.1007/s00438-017-1358-5>.
53. Kawash JK, Smith SD, Karaiskos S, Grigoriev A. 2018. ARIADNA: machine learning method for ancient DNA variant discovery. *DNA Res* 25:619–627. <https://doi.org/10.1093/dnares/dsy029>.
54. Renaud G, Hanghøj K, Willerslev E, Orlando L. 2017. gargammel: a sequence simulator for ancient DNA. *Bioinformatics* 33:577–579. <https://doi.org/10.1093/bioinformatics/btw670>.
55. White DJ. 1997. Dental calculus: recent insights into occurrence, formation, prevention, removal and oral health effects of supragingival and subgingival deposits. *Eur J Oral Sci* 105:508–522. <https://doi.org/10.1111/j.1600-0722.1997.tb00238.x>.
56. Lieverse AR. 1999. Diet and the aetiology of dental calculus. *Int J Osteoarchaeol* 9:219–232. [https://doi.org/10.1002/\(SICI\)1099-1212\(199907/08\)9:4<219::AID-OA475>3.0.CO;2-V](https://doi.org/10.1002/(SICI)1099-1212(199907/08)9:4<219::AID-OA475>3.0.CO;2-V).
57. Shillito L-M, Blong JC, Green EJ, van Asperen EN. 2020. The what, how and why of archaeological coprolite analysis. *Earth Sci Rev* 207:103196. <https://doi.org/10.1016/j.earscirev.2020.103196>.
58. UNESCO. 2008. Links between biological and cultural diversity: report of the international workshop. UNESCO, Paris, France. <https://unesdoc.unesco.org/ark:/48223/pf0000159255>.
59. Benbow ME, Lewis AJ, Tomberlin JK, Pechal JL. 2013. Seasonal necrophagous insect community assembly during vertebrate carrion decomposition. *J Med Entomol* 50:440–450. <https://doi.org/10.1603/me12194>.
60. Metcalf JL, Carter DO, Knight R. 2016. Microbiology of death. *Curr Biol* 26: R561–R563. <https://doi.org/10.1016/j.cub.2016.03.042>.
61. Metcalf JL, Xu ZZ, Bouslimani A, Dorrestein P, Carter DO, Knight R. 2017. Microbiome tools for forensic science. *Trends Biotechnol* 35:814–823. <https://doi.org/10.1016/j.tibtech.2017.03.006>.
62. Dabney J, Meyer M, Pääbo S. 2013. Ancient DNA damage. *Cold Spring Harb Perspect Biol* 5:a012567. <https://doi.org/10.1101/cshperspect.a012567>.
63. Pääbo S, Poinar H, Serre D, Jaenicke-Despres V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, Hofreiter M. 2004. Genetic analyses from ancient DNA. *Annu Rev Genet* 38:645–679. <https://doi.org/10.1146/annurev.genet.37.110801.143214>.
64. Orlando L, Allaby R, Skoglund P, Sarkissian CD, Stockhammer PW, Ávila-Arcos MC, Fu Q, Krause J, Willerslev E, Stone AC, Warinner C. 2021. Ancient DNA analysis. *Nat Rev Methods Primers* 1:15. <https://doi.org/10.1038/s43586-020-00011-0>.
65. Briggs AW, Stenzel U, Johnson PLF, Green RE, Kelso J, Prüfer K, Meyer M, Krause J, Ronan MT, Lachmann M, Pääbo S. 2007. Patterns of damage in genomic DNA sequences from a Neandertal. *Proc Natl Acad Sci U S A* 104: 14616–14621. <https://doi.org/10.1073/pnas.0704665104>.
66. Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S. 2012. Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* 7:e34131. <https://doi.org/10.1371/journal.pone.0034131>.
67. Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29:1682–1684. <https://doi.org/10.1093/bioinformatics/btt193>.
68. Neukamm J, Peltzer A, Nieselt K. 2021. DamageProfiler: fast damage pattern calculation for ancient DNA. *Bioinformatics* 37:3652–3653. <https://doi.org/10.1093/bioinformatics/btab190>.
69. Haile J, Holdaway R, Oliver K, Bunce M, Gilbert MTP, Nielsen R, Munch K, Ho SYW, Shapiro B, Willerslev E. 2007. Ancient DNA chronology within sediment deposits: are paleobiological reconstructions possible and is DNA leaching a factor? *Mol Biol Evol* 24:982–989. <https://doi.org/10.1093/molbev/msm016>.
70. Libby WF. 1955. Radiocarbon dating. University of Chicago Press, Chicago, IL.
71. Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R, Weyrich LS. 2019. Contamination in low microbial biomass microbiome studies: issues and recommendations. *Trends Microbiol* 27:105–117. <https://doi.org/10.1016/j.tim.2018.11.003>.
72. Peyrégne S, Prüfer K. 2020. Present-day DNA contamination in ancient DNA datasets. *Bioessays* 42:e2000081. <https://doi.org/10.1002/bies.202000081>.
73. Duchêne S, Ho SYW, Carmichael AG, Holmes EC, Poinar H. 2020. The recovery, interpretation and use of ancient pathogen genomes. *Curr Biol* 30:R1215–R1231. <https://doi.org/10.1016/j.cub.2020.08.081>.
74. Breailey JC, Leitão HG, van der Valk T, Xu W, Bougiouri K, Dalén L, Guschanski K. 2020. Dental calculus as a tool to study the evolution of the mammalian oral microbiome. *Mol Biol Evol* 37:3003–3022. <https://doi.org/10.1093/molbev/msaa135>.
75. Farrer AG, Wright SL, Skelly E, Eisenhofer R, Dobney K, Weyrich LS. 2021. Effectiveness of decontamination protocols when analyzing ancient DNA preserved in dental calculus. *Sci Rep* 11:7456. <https://doi.org/10.1038/s41598-021-86100-w>.
76. Fagernäs Z, García-Collado MI, Hendy J, Hofman CA, Speller C, Velsko I, Warinner C. 2020. A unified protocol for simultaneous extraction of DNA and proteins from archaeological dental calculus. *J Archaeol Sci* 118: 105135. <https://doi.org/10.1016/j.jas.2020.105135>.
77. Furtwängler A, Neukamm J, Böhme L, Reiter E, Vollstedt M, Arora N, Singh P, Cole ST, Knauf S, Calvignac-Spencer S, Krause-Kyora B, Krause J, Schuenemann VJ, Herbig A. 2020. Comparison of target enrichment strategies for ancient



- pathogen DNA. *Biotechniques* 69:455–459. <https://doi.org/10.2144/btn-2020-0100>.
78. Armbrrecht L, Herrando-Pérez S, Eisenhofer R, Hallegraeff GM, Bolch CJS, Cooper A. 2020. An optimized method for the extraction of ancient eukaryote DNA from marine sediments. *Mol Ecol Resour* 20:906–919. <https://doi.org/10.1111/1755-0998.13162>.
  79. Armbrrecht L, Hallegraeff G, Bolch CJS, Woodward C, Cooper A. 2021. Hybridisation capture allows DNA damage analysis of ancient marine eukaryotes. *Sci Rep* 11:3220. <https://doi.org/10.1038/s41598-021-82578-6>.
  80. Cribdon B, Ware R, Smith O, Gaffney V, Allaby RG. 2020. PIA: more accurate taxonomic assignment of metagenomic data demonstrated on sedaDNA from the North Sea. *Front Ecol Evol* 8:84. <https://doi.org/10.3389/fevo.2020.00084>.
  81. Keller M, Spyrou MA, Scheib CL, Neumann GU, Kröpelin A, Haas-Gebhard B, Paffgen B, Haberstroh J, Ribera I, Lacombe A, Raynaud C, Cessford C, Durand R, Stadler P, Nägele K, Bates JS, Trautmann B, Inskip SA, Peters J, Robb JE, Kivisild T, Castex D, McCormick M, Bos KI, Harbeck M, Herbig A, Krause J. 2019. Ancient *Yersinia pestis* genomes from across Western Europe reveal early diversification during the first pandemic (541–750). *Proc Natl Acad Sci U S A* 116:12363–12372. <https://doi.org/10.1073/pnas.1820447116>.
  82. Fellows Yates JA, Velsko IM, Aron F, Posth C, Hofman CA, Austin RM, Parker CE, Mann AE, Nägele K, Arthur KW, Arthur JW, Bauer CC, Crevecoeur I, Cupillard C, Curtis MC, Dalén L, Bonilla MD-Z, Carlos Díez Fernández-Lomana J, Drucker DG, Escrivá EE, Francken M, Gibbon VE, González Morales MR, Mateu AG, Harvati K, Henry AG, Humphrey L, Menéndez M, Mihailović D, Peresani M, Moroder SR, Roksandic M, Rougier H, Sázlová S, Stock JT, Straus LG, Svoboda J, Teßmann B, Walker MJ, Power RC, Lewis CM, Sankaranarayanan K, Guschanski K, Wrangham RW, Dewhurst FE, Salazar-García DC, Krause J, Herbig A, Warinner C. 2021. The evolution and changing ecology of the African hominid oral microbiome. *Proc Natl Acad Sci U S A* 118:e2021655118. <https://doi.org/10.1073/pnas.2021655118>.
  83. Louvel G, Der Sarkissian C, Hanghøj K, Orlando L. 2016. metaBIT, an integrative and automated metagenomic pipeline for analysing microbial profiles from high-throughput sequencing shotgun data. *Mol Ecol Resour* 16:1415–1427. <https://doi.org/10.1111/1755-0998.12546>.
  84. Borry M, Cordova B, Perri A, Wibowo M, Honap TP, Ko J, Yu J, Britton K, Girdland-Flink L, Power RC, Stuijts I, Salazar-García DC, Hofman C, Hagan R, Kagoné TS, Meda N, Carabin H, Jacobson D, Reinhard K, Lewis C, Kostic A, Jeong C, Herbig A, Hübner A, Warinner C. 2020. CoproID predicts the source of coprolites and paleofeces using microbiome composition and host DNA content. *PeerJ* 8:e9001. <https://doi.org/10.7717/peerj.9001>.
  85. Hübner R, Key FM, Warinner C, Bos KI, Krause J, Herbig A. 2019. HOPS: automated detection and authentication of pathogen DNA in archaeological remains. *Genome Biol* 20:280. <https://doi.org/10.1186/s13059-019-1903-0>.
  86. Fellows Yates JA, Lamnidis TC, Borry M, Andrades Valtueña A, Fagernäs Z, Clayton S, Garcia MU, Neukamm J, Peltzer A. 2021. Reproducible, portable, and efficient ancient genome reconstruction with nf-core/eager. *PeerJ* 9:e10947. <https://doi.org/10.7717/peerj.10947>.
  87. Velsko IM, Frantz LAF, Herbig A, Larson G, Warinner C. 2018. Selection of appropriate metagenome taxonomic classifiers for ancient microbiome research. *mSystems* 3:e00080-18. <https://doi.org/10.1128/mSystems.00080-18>.
  88. Eisenhofer R, Weyrich LS. 2019. Assessing alignment-based taxonomic classification of ancient microbial DNA. *PeerJ* 7:e6594. <https://doi.org/10.7717/peerj.6594>.
  89. Secretariat of the Convention on Biological Diversity. 2011. Nagoya Protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their utilization to the Convention on Biological Diversity: text and annex. Secretariat of the Convention on Biological Diversity. <https://www.cbd.int/abs/>.