3. THE SEDIMENTARY ANCIENT DNA WORKFLOW

PETER D. HEINTZMAN (peter.d.heintzman@geo.su.se)
Department of Geological Sciences
Stockholm University
Stockholm, Sweden

Centre for Palaeogenetics
Svante Arrhenius väg 20C
Stockholm, Sweden

KEVIN NOTA (kevin_nota@eva.mpg.de)
Department of Evolutionary Genetics
Max Planck Institute for Evolutionary Anthropology
Leipzig, Germany

ALEXANDRA ROUILLARD (alexandra.rouillard@uit.no)
Department of Geosciences
UiT—The Arctic University of Norway
Postboks 6050 Langnes, N-9037
Tromsø, Norway

Centre for GeoGenetics
University of Copenhagen
Øster Voldgade 5-7
1350 Copenhagen K, Denmark

YOURI LAMMERS (youri.lammers@uit.no)
The Arctic University Museum of Norway
UiT - The Arctic University of Norway
Postboks 6050 Langnes, N-9037
Tromsø, Norway

TYLER J. MURCHIE (murchiet@mcmaster.ca)
McMaster Ancient DNA Centre
Department of Anthropology
McMaster University
Hamilton, ON L8S 4L8, Canada

LINDA ARMBRECHT (linda.armbrecht@utas.edu.au)
Institute for Marine and Antarctic Studies (IMAS)
Ecology and Biodiversity Centre
University of Tasmania
Hobart, TAS 7004, Australia

SANDRA GARCÉS PASTOR (sandragarces@ub.edu)
Department of Evolutionary Biology, Ecology and Environmental Sciences
Universitat de Barcelona
Barcelona, Spain

Marine Sciences Institute,
ICM-CSIC
Barcelona, 08003, Spain

BENJAMIN VERNOT (benjamin_vernot@eva.mpg.de)
Department of Evolutionary Genetics
Max Planck Institute for Evolutionary Anthropology
Leipzig, Germany

Keywords : sedimentary ancient DNA, paleogenomics, ancient metagenomics, metabarcoding, contamination

Introduction

Sedimentary ancient DNA (*sed*aDNA) is continuing to revolutionise our understanding of past biological and geological processes by retrieving and analysing the ancient DNA preserved in lake, cave, open terrestrial, midden, permafrozen, and marine environments (Crump, 2021). The study of *sed*aDNA began in the late 1990s (Coolen and Overmann, 1998) with the first reports of extinct animal *sed*aDNA in 2003 (Hofreiter et al., 2003; Willerslev et al., 2003). Since then, it has been shown that *sed*aDNA can be recovered at high resolution from recent (10^1-10^2 year-old) (e.g., Capo et al., 2017) through to deep-time (10^5-10^6 year-old) sediments from a vast diversity of environments (Crump et al., 2021; Zavala et al., 2021; Armbrecht et al., 2022; Kjær et al., 2022). Unlike traditional palaeoenvironmental and palaeoecological proxies, *sed*aDNA is unique in that it is derived from any type of organism that was present in the local area and that may contain population-level information. This latter characteristic means that, unlike any other comparable proxy, *sed*aDNA can be used for evolutionary analyses (Gelabert et al., 2021; Lammers et al., 2021; Pedersen et al., 2021; Vernot et al., 2021).

The advent of next-generation sequencing (NGS) provided massively enlarged, yet economically feasible, dataset sizes and increased analytical sensitivity that has allowed the field to flourish by generating robust datasets in which contamination can be detected and controlled, and hypotheses can be tested. Coupled with ongoing methodological innovations in both molecular data generation and bioinformatics analysis techniques, NGS has driven the exponential growth in *sed*aDNA research over the past two decades.

In this chapter, we present an overview of the state-of-the-art for the *sed*aDNA workflow. We do not present detailed methodologies and descriptions, as these have been published elsewhere (e.g., Armbrecht et al., 2019; Capo et al., 2021 and references therein). For each step in the workflow, from ethical considerations during experimental design to environmental and evolutionary inferences, we instead outline the general rationale for conducting the step, a brief overview of the approach and methods involved, pros and cons, key pitfalls, and how the current state-of-the-art is likely to develop in the near future. Importantly, we highlight that molecular and bioinformatic methods (i.e., steps presented from section ‘DNA extraction’ onwards) are still developing due to the relative infancy of the field.

Ethical considerations

Sources of environmental sediments potentially amenable to *sed*aDNA analysis are ubiquitous and global, thereby providing ample material for a multitude of research questions. Importantly, however, there are ethical, political, and/or legal considerations that must be accounted for during the study design and data collection stages of a project (Bardill et al., 2018). This includes abiding by international agreements on the use and sharing of genetic resources such as the Nagoya protocol (Buck and Hamilton, 2011).

Permits from national/local authorities and/or local landowners are often required to perform the fieldwork necessary to collect sediments. If sediments and/or the extracted DNA are to be shipped internationally, as is often the case in current *sed*aDNA research, then export and/or import permits will be required (e.g., from the United States Department of Agriculture or Australian Government). In many regions of the world, especially with legacies of colonialism, there is a strong need to engage with key stakeholders: the local indigenous communities (Handsley-Davis et al., 2021; Kowal et al., 2023). This is particularly important for *sed*aDNA research, given that questions often revolve around climate, environmental, and land use changes on scales that are of relevance to indigenous peoples.

One of the ethically problematic aspects of *sed*aDNA research that has been most discussed in the literature is the potential for recovering DNA from humans (*Homo sapiens*), of which many metagenomic and animal-targeting assays are both able and prone to detect (see section 'Data generation strategies'). Although most, if not all, human DNA recovered from some sedimentary archives, such as lake, marine, and permafrost, likely represents modern-day contamination, there is potential for authentic human *sed*aDNA to be recovered. In other sedimentary contexts, such as caves and archaeological sites, authenticated human *sed*aDNA sequences that are tens of thousands of years old have been reported (e.g., Braadbaart et al., 2020; Zavala et al., 2021). If there will be the possibility of recovering *sed*aDNA of humans, or other culturally important taxa, then, where applicable, it is crucial to consult with indigenous communities during the consultation stage of experimental design (Alpaslan-Roodenberg et al., 2021; Handsley-Davis et al., 2021; Kowal et al., 2023). A potential solution to dealing with the undesired recovery of human *sed*aDNA is to use a bioinformatic filtering step to remove human, or potentially human, sequences without further analysis. However, we emphasise that the ethics of *sed*aDNA, and particularly human DNA recovered from sediment, are an active area of discussion that is only in the early stages of conversation.

Fieldwork

The general logistics of fieldwork for a *sed*aDNA study are similar to those used for the collection of sediments for other, more traditional proxies. This includes excavations in natural and archaeological settings, horizontal coring from accessible profiles such as permafrost, soil, and cave settings (Fig. 1A), or vertical coring from inaccessible profiles

such as lake, marine and ice systems. The key difference from traditional fieldwork is that steps must be taken to prevent or reduce contamination by human or environmental DNA during sample collection from the sediment profile. Due to the high sensitivity of the *sed*aDNA technique, inferences are particularly prone to sources of contamination, both by modern-day environmental DNA and from sediment cross-contamination that can occur during sample collection and handling.

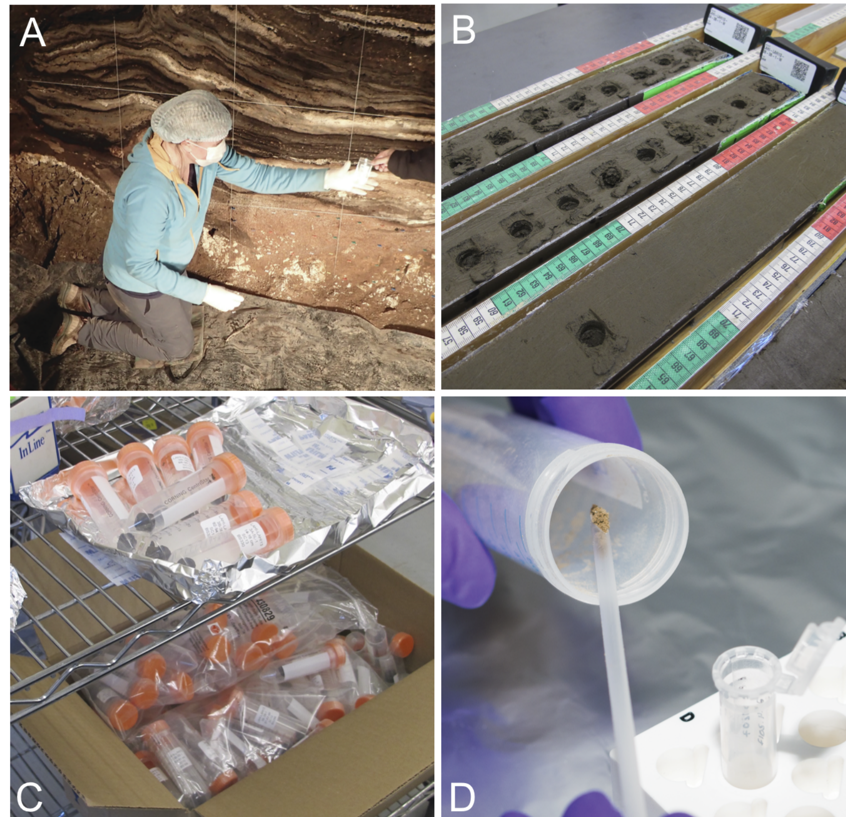

Figure 1. Examples of (A) researchers sampling an accessible cave sediment profile in the field; (B) a freshly subsampled lake core, where the rectangular areas represent discarded sediment surface for access to underlying pristine sediment, and the circular holes correspond to the material that was sampled for DNA analysis. Alternative core sectioning and subsampling strategies can be used for higher resolution analyses; (C) subsamples that were taken in a non-cleanroom space; and (D) a subsample to be used as input for DNA extraction in a cleanroom space. Images are copyright of Dr. Richard G. Roberts (A), Dr. Peter D. Heintzman (B, C), and the Max Planck Institute for Evolutionary Anthropology (D).

Contamination from modern-day environmental DNA, including from the persons performing the fieldwork, is of paramount concern when taking sediments from excavation and accessible profile contexts. Standard ancient DNA precautions can be adopted to reduce this type of contamination, including the use of sterile utensils and storage bags/tubes, and personal protective equipment such as facemask, gloves, and a full-body hazmat suit (Fulton and Shapiro, 2019). Modern-day environmental DNA contamination can be monitored by the use of field controls, whereby sterile tubes, either empty or containing a storage buffer/water, are opened during sediment collection. An alternative approach, especially for inaccessible profiles, is to pre-sterilise coring equipment and/or coat the inside of a coring tube with a known exotic DNA, or synthetic tracer, and monitor for its presence during downstream analyses (Armbrecht et al., 2019; Graham et al., 2016; Pedersen et al., 2016; Rijal et al.,

2021). For highly sensitive applications, such as *sed*aDNA analysis of environmental microorganisms (Table 1), a pipe-based coring system, such as a Livingstone, Nesje (Nesje, 1992), UWITEC (https://www.uwitec.at/en/), or Kasten, is better suited to preventing modern-day microbes from colonising the sediment. All sediment cores or samples should be immediately sealed from the external environment after collection to prevent desiccation and ongoing contamination from modern-day environmental DNA, and to maintain a cold, dark and, if possible, anoxic environment to reduce post-sampling DNA degradation and secondary growth.

Cross-contamination, whereby sediment from a different part of the profile is incorporated into a sample during collection, can be limited by systematically sampling accessible profiles from bottom to top. This reduces the chance of loose sediments displaced during sampling from causing cross-contamination. In inaccessible profiles, cross-contamination can occur if cores are taken in a series of drives, as is standard with the Livingstone coring system, and so results from the tops of core drives should be interpreted cautiously. If possible, taking a single long-core, such as with a Nesje-style or UWITEC system, avoids this type of cross-contamination from inaccessible profiles (e.g., Armbrecht et al., 2019).

In excavations and accessible profiles, it is common practice to avoid both types of contamination in the field by removing the outer layer of exposed sediment, consisting of several mm to cm of the surface depending on context, using sterile implements (see also section ‘Subsampling’). This allows for pristine sediment samples to be taken from inside the profile (ideally still frozen in the case of permafrost), which can be achieved using either a new set of sterile implements or by inserting a sampling tube directly into the sediment profile (e.g. Ardelean et al., 2020; Vernot et al., 2021).

Table 1. Taxon-specific methodological considerations for the *sed*aDNA workflow, with signposts to other chapters.

	Fresh cores	**Contamination**	**Suitable for**		**Further**
Group	**recommended**	**potential**	**Amplicon**	**Metagenomics******	**reading**
Prokaryotes	yes	high	no**	yes	Chapters 4,5
Microbial eukaryotes	yes	high	no**	yes	Chapters 6,7
Fungi	yes	high	yes***	yes	Chapter 6
Wildlife*	no	low	yes***	yes	Chapters 8,9,10,11
Domesticated/farmed taxa*	no	intermediate	yes***	yes	Chapters 10,11
Human	yes	high	no	yes	Chapter 11

* Wildlife and domesticated/farmed taxa include both animal and plant taxa.
** Amplicon methods may be appropriate for taxa that are unable to have colonised the sediment after collection or be present in the modern sampling or laboratory environment.
*** Data are prone to contamination and so need to be interpreted carefully. See also **Chapter 11**.
**** The feasibility of metagenomics may be limited by the availability of reference genome data for certain taxonomic groups.

Sediment storage

The storage conditions for collected sediments depends on a variety of factors, including logistic availability and the intended downstream analyses. For intact sediment cores taken from lake or marine non-frozen contexts, refrigerated storage is preferred because freezing can adversely affect stratigraphic integrity as well as being impractical given the total sediment mass. Cores for *sed*aDNA analysis are therefore usually stored in cold rooms at 4 ˚C, with studies reporting successful *sed*aDNA retrieval from cores stored for >20 years at this temperature (e.g., Seeber et al., 2022). For some applications, freezing is necessary to maintain sediment integrity. Such occasions include, for example, the preservation of permafrost sediment plugs and when sediments are subsampled for *sed*aDNA analysis immediately after coring (e.g., on board research vessels, Armbrecht et al., 2019). However, in such circumstances, complementary archive cores are often taken to allow for future follow-up study (Selway et al., 2022). If cores are frozen, it is important to keep freeze-thaw cycles to a minimum as these can negatively impact the integrity of *sed*aDNA molecules by allowing for microbial-induced degradation (e.g., case study 2 in Capo et al. 2021) or through physical shearing during ice crystal formation (Shao et al., 2012; Ross et al., 1990). However, the degree to which freeze-thaw meaningfully degrades *sed*aDNA remains somewhat unresolved as others have found negligible impacts from repeated freeze/thaw cycles (Safarikova et al., 2021).

Evidence for the detrimental effects of X-rays on *sed*aDNA within sediments, such as those conducted during transport at airports and customs or on the intact cores for multi-proxy analyses of the sediment (ITRAX or μCT scanning), is currently lacking. To our knowledge, retrieval of *sed*aDNA has been successful following X-rays during transport, but should likely be minimised where possible. The impacts of X-rays on sub-fossil bone ancient DNA during μCT scanning have been found to be particularly damaging in wet and/or frozen materials (Immel et al., 2016). We therefore recommend limiting resolution and total exposure to X-rays prior to *sed*aDNA subsampling (Immel et al., 2016).

Sediment cores and samples are often stored in specialist facilities, such as the LacCore facility in Minnesota (Schnurrenberger et al., 2001), the Permafrost ArChives Science (PACS) Laboratory at the University of Alberta, core repositories of the International Ocean Discovery Program (IODP, https://www.iodp.org/resources/core-repositories), or various institutional archives. These archival facilities not only allow for the long-term preservation of these sediments, but provide material to allow for follow-up and reproducible research following the Findability, Accessibility, Interoperability, and Reusability (FAIR) principles (Wilkinson et al., 2016).

Sedimentary ancient DNA has been successfully recovered from sediments collected years, or even decades, prior to analysis (e.g., Massilani et al., 2022; Rijal et al., 2021; Wang et al., 2021), thereby demonstrating the value of utilising existing collections for *sed*aDNA research. However, caution should be taken when using such sediments for the study of microbes and fungi, due to the potential for secondary growth of modern-day contaminants on, and potentially within, the sediments (Selway et al., 2022). In some instances, these modern organisms can overwhelm the *sed*aDNA signal below meaningful detection levels (see also section ‘Identification and validation: Validation’). Where possible, sediment cores

should remain sealed with plastic caps and duct-tape until immediately prior to *sed*aDNA subsampling. This is to prevent microbial and fungal growth by maintaining either an air-tight or anoxic environment within the core.

Subsampling

The goal of subsampling is to remove a small portion of sediment core or sample, either to be used directly in a DNA extraction (~50 mg to 1 g) or to provide a larger stock of subsample suitable for multiple extractions (~10 to 50 g) (Fig. 1B&1C).

Subsampling is the first step in the workflow to be conducted ideally in a specialised clean lab, and should always follow standard anti-contamination precautions (Epp et al., 2019; Fulton and Shapiro, 2019). Recommended measures include bleach and ultraviolet (UV) treatment of work surfaces and equipment, use of sterile utensils and personal protective equipment (see also section ‘Sediment storage’), working within a HEPA-filtered positive-pressured atmosphere, and the one-way movement of personnel and equipment from the clean-room environment to laboratories that use high-copy and/or amplification-based molecular biology procedures (see section 'Data generation strategies: Metagenomics'). If access to a clean lab is not possible for subsampling, then the work environment should be made as clean as is practical and located in an area without active ventilation and where no previous molecular biology procedures have been conducted. Subsampling negative controls should be used to monitor for modern-day environmental DNA contamination in a similar manner as described in the section ‘Sediment storage’, particularly when subsampling in non-cleanroom environments.

A detailed step-by-step guide for subsampling frozen and non-frozen sediment cores can be found in Epp et al. (2019). Prior to and immediately after entry into the cleanroom facility, the external surfaces of sediment sample bags/tubes and core tubes should be decontaminated with a bleach solution. We emphasise that, whilst UV may be used to decontaminate the surface of a subsample, chemical decontamination, such as the use of bleach, should not be used as *sed*aDNA within the subsample will not be protected from chemically-induced degradation. An alternative approach to remove surface contamination from blocks of sediment, such as unthawed permafrost or sediment cores, is to sequentially remove one or two exterior layers so that a previously undisturbed sector of the sediment is subsampled similarly to that described in section ‘Sediment storage’ (Fig. 1B; Graham et al., 2016; Pedersen et al., 2016; Parducci et al., 2017). The implements used to collect subsamples will depend on the sediment characteristics, such as hardness and consistency, but, in all cases, sediment in direct contact with tube edges should be avoided. If a tracer was applied during coring, then subsamples should be taken from the exterior of the core to confirm the presence of the tracer. Surface cores, such as those obtained by gravity coring devices, often contain unconsolidated and/or watery sediment that may prove challenging to subsample cleanly, but protocols are available (e.g., Pawlowski et al., 2022). Once taken, it is standard practice to store subsamples frozen at -20 ˚C in an ancient DNA laboratory in preparation for DNA extraction.

DNA extraction

The isolation and purification of *sed*aDNA from a sediment subsample is known as DNA extraction. This should always be performed in cleanroom conditions and include negative controls consisting of all steps and reagents used in the extraction protocol with no sediment input (Fig. 2). Sediment input masses vary between protocols, but can be as low as 50 mg and as high as 15 g (e.g., Ficetola et al. 2018; Massilani et al., 2022, Fig. 1D).

DNA extraction involves three major steps: disintegration, inhibitor removal, and purification (e.g., Epp et al., 2019). The disintegration step frees DNA molecules that are either mineral-bound (extracellular) or still associated with organic material in tissues or small clumps of cells (intracellular), using physical, chemical, and/or enzymatic reagents to break up the sediment and its organic components. Inhibitor removal steps use chemical or physical filtering to remove complex organic molecules, such as cellular debris, humic acids and polysaccharides, that can inhibit the downstream enzymatic molecular biology procedures outlined in section 'Data generation strategies'. Inhibitor removal can be non-trivial, as humic acids, for example, are biochemically similar to DNA for extraction purposes. This means that aggressive inhibitor removal can also remove *sed*aDNA molecules, whereas too lenient inhibitor removal results in the co-extraction of inhibitors (e.g., Murchie et al., 2021). Consequently, inhibitor removal may be omitted from the extraction of DNA from some sediment types, such as those with a low organic content where inhibitor content is expected to be low. Purification of the *sed*aDNA involves the removal of remaining cellular debris and reagents from previous steps. This can be achieved through binding the *sed*aDNA to silica, which is then washed and the purified DNA released (e.g., Epp et al., 2019), or through the use of organic phase separation (e.g., Wang et al., 2021). The resulting *sed*aDNA extract should be stored frozen at -20 ºC in the clean laboratory prior to downstream molecular biology procedures.

The DNA extraction procedure is critical, as suboptimal performance in any of the three steps listed above will adversely impact the quantity and quality of the recovered *sed*aDNA. Given the wide variety and complexity of sediment types, both in terms of geochemical and organic composition, the precise *sed*aDNA extraction protocol to use should be chosen carefully and may require testing and optimisation (e.g., Sand et al., 2023). In such a scenario, we suggest testing a candidate protocol on a small, but representative, subset of the sediment subsamples to assess DNA extraction efficiency prior to major dataset production. Although optimised *sed*aDNA extraction protocols exist for lake (Epp et al., 2019; Capo et al., 2021; Rijal et al., 2021), cave (Slon et al., 2017), marine (Armbrecht et al., 2020), and permafrost (Murchie et al., 2021) sediment archives, we note that these have often been optimised for sediments from specific localities and so may not be broadly applicable. The testing and further optimisation of DNA extraction protocols is therefore an active area of ongoing methodological development research within the *sed*aDNA community, especially with regard to improving the efficiency of both inhibitor removal from organic-rich sediments and the release of *sed*aDNA from minerogenic substrates, such as smectite clay, that tightly bind DNA molecules (Jelavić et al., 2023).

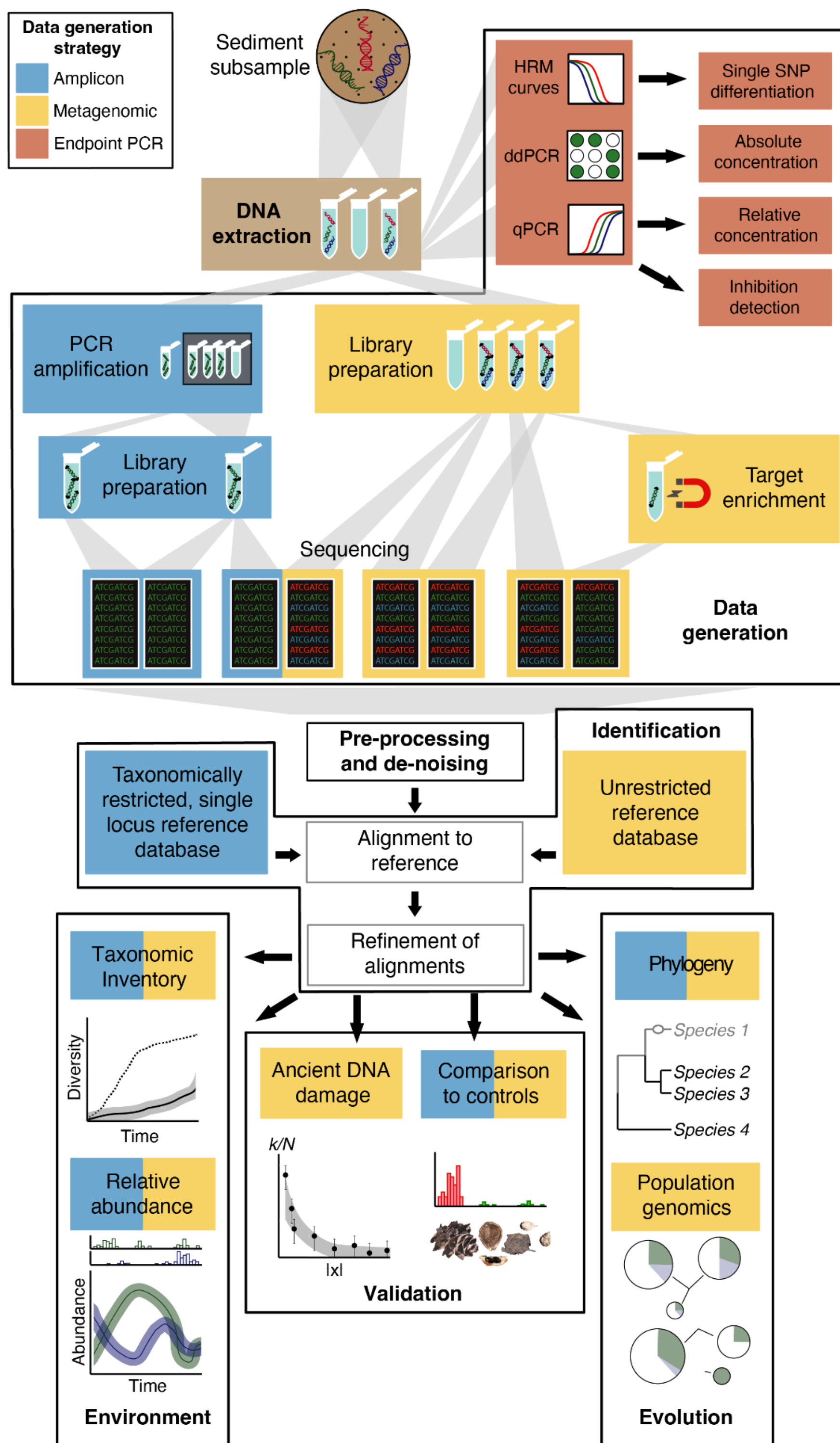

Figure 2. A schematic of the complete *sed*aDNA workflow after subsampling. The main steps are listed in bold. The figure is modified from Williams et al. (2023), with illustrations of validation, environmental, and evolutionary analyses inspired by Kjær et al. (2022), Rijal et al. (2021), and Wang et al. (2021). Note that library preparation involves an indexing PCR amplification step. Acronyms are defined in the text. Source credit for macrofossil images: Dr. Alexandra Rouillard.

Data generation strategies

At present, there are three main strategies for the generation of *sed*aDNA data from DNA extracts: endpoint polymerase chain reaction (PCR), amplicon, and metagenomics (shotgun, target enrichment), which we outline below (Fig. 2). Regardless of the strategy applied, but especially for amplicon approaches, it is recommended to use replicates at each step to verify the consistency of detections (Ficetola et al., 2015). Both the amplicon and metagenomics approaches involve the sequencing of DNA nucleotides, which are read using high-throughput sequencing platforms to generate 'sequence read' data. Both approaches also involve the conversion of *sed*aDNA into a sequencing library. Library preparation involves the addition of sequencing adapters and indexes, both consisting of an artificial DNA sequence, to the ends of *sed*aDNA molecules. The addition of adapter sequences is required for *sed*aDNA molecules to be 'read' by the DNA sequencing platform. Indexes consist of unique, short sequence motifs that allow libraries from different samples or replicates to be differentiated when combined on a single sequencing run (see section 'Data generation strategies: Sequencing'). Library negative controls should be generated in parallel with libraries from DNA extracts.

Endpoint PCR

High-throughput sequencing may not be required for some research questions, such as quantifying template abundance, quantifying inhibition, detecting the presence of a single taxon, or discriminating between closely related genotypes in a *sed*aDNA extract. Instead, taxon- or gene function-specific detection assays, or assays that can discriminate between a small number of targets, might be more appropriate. Such methods are not frequently used for detecting higher organisms (Nota et al., 2022), but are relatively common for quantifying microorganisms (e.g., Pal et al., 2015; Pilon et al., 2019; Picard et al., 2022; Nwosu et al., 2023).

Endpoint PCR methods rely on designing taxon- or functional gene-specific primers (short pre-designed sequences used to initiate PCR) with high specificity during amplification. Using quantitative PCR (qPCR) or droplet digital PCR (ddPCR), it is possible to estimate the original number of template molecules present. The ddPCR approach has the advantage over qPCR in that standard curves are not required for quantification and might improve template amplification in inhibited samples (Capo et al., 2021). When the aim is to detect a single taxon or to distinguish a small number of taxa, qPCR has been used together with (high resolution) melting curve analysis (HRM). Melting curve analysis takes advantage of denaturing DNA from double to single-stranded molecules, which is a function of molecule length and GC content. For performing melting curve analysis, enough variation between amplicons is required to create distinct melting peaks. It has been shown that even single nucleotide polymorphisms (SNPs) can be genotyped using melting curve analysis and specific high-resolution instruments (Liew et al., 2004). For *sed*aDNA applications, melting curves have been used to assess the specificity of qPCR amplicons (Savichtcheva et al., 2011; Domaizon et al., 2013; Nwosu et al., 2023) and for genotyping (Nota et al., 2022).

For inhibition quantification, endpoint PCR can be used by spiking a *sed*aDNA extract into a PCR designed to amplify an unrelated template, such as a synthetic control or lambda phage (Warinner et al., 2014; Murchie et al., 2021). Assays with *sed*aDNA extract inhibition will exhibit an increased cycle threshold (Ct) value, due to impeded exponential amplification during qPCR. By comparing *sed*aDNA extracts in dilution series to positive controls (i.e., with no *sed*aDNA extract spike), relative inhibition levels can be determined between *sed*aDNA extracts.

Caution is required using endpoint PCR approaches for template quantification and SNP detection due to the potential amplification of non-target DNA from unrelated organisms. The specificity of endpoint PCR is generally tested on DNA from present-day environments whereby results can be validated by independent observations. This is not possible for *sed*aDNA. It is therefore strongly advised to verify that the signal is coming from the expected target by sequencing a subset of the positive PCR amplifications. The validation of qPCR results may be facilitated with melting curves, although different sequences can give rise to the same melting curve profile.

Amplicon

Amplicon methods involve the high-throughput sequencing of PCR amplicons that are either generated using a generic primer set, which amplify templates from a broad spectrum of taxa or functional genes (metabarcoding PCR), or PCRs that simultaneously use multiple primer sets (multiplex PCR). Primer sequences are often extended with unique, short motifs (tags) that allow for amplicons to be assignable to their original PCR in experiments where multiple PCR products are pooled together prior to library preparation (Binladen et al., 2007).

Metabarcoding PCR is used to simultaneously amplify *sed*aDNA templates from broad groups such as vascular plants or vertebrates, and thereby assess the diversity and composition of taxa represented within the *sed*aDNA extract (Fig. 2; Capo et al., 2021). This robust, relatively cheap, and widely used technique is based on the barcoding principle whereby a pair of generic primers bind to separate parts of the DNA template, which are conserved across the group of interest, and flank a variable region of the template, the ‘barcode’ region, that allows for discrimination of species or genera (Nichols et al., 2019; Taberlet et al., 2018b). As metabarcoding requires intact *sed*aDNA fragments that include the primer binding sites, there is a trade-off between using a barcode that is short enough to be compatible with degraded ancient DNA (typically <150 bp) yet long enough to be taxonomically informative (e.g. Taberlet et al., 2007). Metabarcoding is a highly sensitive technique, which makes it ideal for detecting template molecules that occur at ultra-low concentrations in a *sed*aDNA extract. However, it is also prone to amplification biases (Ziesemer et al., 2015; Nichols et al., 2018; Capo et al., 2021) and is sensitive to contamination. Human contamination can be particularly problematic for animal metabarcoding assays, and so blocking primers, which compete with metabarcoding primers during binding to human template molecules, have been developed to limit the amplification of contaminating human DNA (e.g., Boessenkool et al., 2012). Contamination sensitivity coupled with the fact that PCR amplicons cannot be authenticated using deamination-based damage profiles (see also section ‘Identification and validation: Validation’) means that

metabarcoding assays should not target bacteria or human *sed*aDNA and caution should be applied to studies of plants, fungi, and animals (Table 1, see also **Chapter 11**).

Multiplex PCR uses multiple primer sets within a single PCR reaction to target multiple templates simultaneously. This greatly increases the amount of information obtained and eliminates the need to run multiple reactions, which deplete *sed*aDNA extracts and increase costs (Taberlet et al., 2018a). Multiplex PCR has been applied to ancient DNA to improve the taxonomic resolution of key taxa that are otherwise difficult to resolve with a single primer set (Côté et al., 2016). Based on applications to modern environmental DNA, multiplex PCR could be applied to *sed*aDNA to study cryptic species (Brosseau et al., 2019), simultaneously target multiple broad taxonomic groups of interest (Weber et al., 2023), or recover population-wide genotyping data (Andres et al., 2021). However, the increased complexity of multiplex PCR can cause problems and therefore requires careful consideration during study design (Ficetola and Taberlet, 2023). First, all primers need to have compatible annealing temperatures and amplify templates of comparable size, in order to ensure equal amplification. Secondly, the primers included have to be selected so as to avoid potential dimer formation during the PCR (Côté et al., 2016; Taberlet et al., 2018a). These issues can be mitigated by dividing the primer sets into several multiplex PCRs with their own specific annealing temperatures, lengths, or lack of hybridising partner primers, though at increased cost (Côté et al., 2016; Brosseau et al., 2019). Another concern is that of different template concentrations when multiplexing metabarcode primers that target multiple taxonomic groups, which can result in uneven amplicon concentrations after PCR. Adjustments can be made to the primer concentrations to offset these issues, but optimisation can require considerable time (Taberlet et al., 2018a). Lastly, the complexity multiplex PCR data may be challenging to bioinformatically analyse (Ficetola and Taberlet, 2023).

Metagenomics

Metagenomics-based data generation approaches involve the conversion of total *sed*aDNA in an extract into a sequencing library followed by either direct sequencing (shotgun strategy) or by a targeted enrichment step to enrich the library for *sed*aDNA molecules of interest prior to sequencing. In both cases, entire *sed*aDNA molecules are sequenced, allowing both for molecules too short for amplicon approaches to be analysed and for the assessment of DNA damage profiles (see section ‘Identification and validation’).

Shotgun metagenomics is the most conceptually simple form of sequence data generation, whereby any of the *sed*aDNA molecules present in a DNA extract can potentially be sequenced. As *sed*aDNA molecules can exist in either single- or double-stranded form (i.e., a half or complete double helix), shotgun library preparation procedures can either target the double-stranded fraction (e.g., Meyer and Kircher, 2010) or both the double- and single-stranded DNA molecular components (e.g., Gansauge et al., 2017). Traditionally, double-stranded DNA library preparation has been preferred due to simplicity and lower cost (e.g., Graham et al., 2016; Pedersen et al., 2016; Wang et al., 2021), but methods that also target the single-stranded DNA component are now economically competitive (Kapp et al., 2021) and are being used in an increasing number of studies (e.g., Schulte et al., 2021; Courtin et al., 2022). Once the *sed*aDNA extract has been converted into a library, it is

directly sequenced without any form of enrichment. The resulting data therefore represent, relative to other methods mentioned in this section, an unbiased snapshot of the *sed*aDNA extract. Given that total *sed*aDNA is generally dominated by bacteria and fungi (or are unidentifiable), shotgun metagenomics data are often overwhelmed by sequences from these taxa (e.g., Pedersen et al., 2016; Lammers et al., 2021) (see also section 'Identification and validation: Identification'). If the aim is to study plants and animals, which are typically present in very low proportions of shotgun metagenomics data, then additional analyses beyond the detection of taxa and assessment of DNA damage may not be feasible or economically justifiable. For other questions, such as those requiring population-level data, then a targeted enrichment approach may be more appropriate.

Targeted enrichment, also referred to as targeted or hybridization capture, provides a middle ground for *sed*aDNA sequence data generation by targeting only a subset of the library molecules that align with the research question, akin to metabarcoding, but by capturing entire *sed*aDNA molecules, akin to shotgun metagenomics. This approach therefore allows for a greater proportion of *sed*aDNA of interest to be analysed and to be assessed via DNA damage analyses. Targeted enrichment uses pre-designed DNA or RNA 'bait' molecules, also known as probes, to capture complementary molecules in a shotgun metagenomics library. Uncaptured 'off-target' molecules are then removed, thereby enriching the library for *sed*aDNA molecules of interest (e.g., Armbrecht et al., 2021; Murchie et al., 2021; Soares, 2019; Vernot et al., 2021). A weakness of the target enrichment approach is that it requires *a priori* knowledge of the reference sequences required to design baits, although, if uncontaminated specimens are available, probes can also be generated directly from modern DNA (e.g., Maricic et al., 2010). However, experimental parameters (i.e., hybridisation temperature) can be modified to reduce the specificity of the capture, which can allow for the enrichment of more distantly related desired sequences but also increase the likelihood of capturing undesirable off-target molecules. Targeted enrichment has been used in *sed*aDNA studies to capture barcode loci (Murchie et al., 2021), functional genic loci (Armbrecht et al., 2021), mitochondrial and chloroplast genomes (Slon et al., 2017; Schulte et al., 2021; Kjær et al., 2022), and genome-wide data (Vernot et al., 2021) from either single species or whole plant and animal communities. The application of target enrichment to generate genome-wide data is an area of anticipated growth that will allow for the investigation of evolutionary and population-level questions using *sed*aDNA (see section 'Identification and validation: Validation').

Sequencing

The final step of sequence data generation is the high-throughput sequencing of amplicon and metagenomics libraries, in which the library molecules are 'read' to produce sequencing reads that are used for downstream bioinformatics analyses. At present, *sed*aDNA sequencing is performed almost exclusively on short-read *Illumina* next-generation sequencers due to low cost and high-throughput, producing $>10^7$-10^9 reads in a single sequencing run. However, we note that emerging short-read sequencers, such as the *PacBio Onso* platform that has a lower error rate than *Illumina*, may become dominant in the future (Eisenstein, 2023). Illumina sequencing can be performed either in single-end mode, whereby library molecules

are read from one end only, or in paired-end mode, in which molecules are read from both ends. Single-end sequencing may be more economical whereas pair-end sequencing has a lower error rate due to parts of the library molecule being read twice.

Due to the high throughput of DNA sequencers, it is often desirable to multiplex libraries, whereby multiple samples are sequenced on the same run. DNA reads from individual libraries are then identified based on their library index (see introduction to this section) and demultiplexed during pre-processing (see section 'Data pre-processing and denoising'). Index-hopping, whereby a read is assigned the wrong index, will result in a form of cross-contamination between libraries, whereas tag-jumping is the within-library equivalent for amplicon experiments (Rodriguez-Martinez et al., 2023). These artefacts can be avoided by using a unique dual-indexing strategy, in which each library molecule has two indexes (Kircher et al., 2012), the use of unique dual-tags in amplicon experiments (Binladen et al., 2007), unique molecular identifiers in library preparation (MacConaill et al., 2018), and/or by not including genomic DNA libraries on the same sequencing runs as *sed*aDNA libraries (Graham et al., 2016). If unique dual-tagging/indexing is used, then the incidences of library index-hopping and amplicon tag-jumps can be checked using the prevalence of unexpected tag and/or index combinations (Zinger et al., 2019; Rodriguez-Martinez et al., 2023)

Lastly, it is crucial that negative controls from each step in the data generation pipeline (e.g., sampling, DNA extraction, PCR amplification or metagenomic library preparation) are sequenced to monitor for contamination, and, where possible, trace its origins in the workflow.

Data pre-processing and denoising

The first bioinformatics step in processing *sed*aDNA sequence data is to filter out sequences and reads that are either experimental artefacts or have insufficient information content to be confidently identified and authenticated. Pre-processing consists of demultiplexing, or 'separating', of libraries together with the removal of library adapter/primer sequences and, if applicable, merging of overlapping paired-end reads. While library demultiplexing is often performed by the sequencing platform's on-board software, read merging and adapter/primer trimming are achieved using tools such as fastp (Chen et al., 2018), AdapterRemoval2 (Schubert et al. 2016), or SeqPrep (https://github.com/jstjohn/SeqPrep). Denoising then removes reads that are of an incorrect length, have low information content, and/or are duplicated. Pre-processing steps are generally identical for both amplicon and metagenomics experiments, whereas denoising procedures diverge.

In an amplicon experiment, denoising involves the removal of reads that are likely to represent artefacts or off-target amplification. This includes reads that lack expected primer and tag sequences and/or are longer or shorter than the expected amplicon length (including primers and tags). Duplicated amplicon reads are then collapsed to a single representative read, often whilst also accounting for errors introduced by PCR or sequencing. After denoising, amplicons may be kept as amplicon sequence variants (ASVs) or collapsed by a pairwise-identity threshold (e.g., 95%) into molecular operational taxonomic units (MOTUs)

prior to identification. The short lengths of *sed*aDNA amplicons generally result in reduced taxonomic precision with the MOTU approach, and so ASVs are more commonly used for identification. Pipelines for denoising of amplicon data include ObiTools (Boyer et al., 2016), Anacapa (Curd et al., 2019), and LULU (Frøslev et al., 2017).

Denoising of metagenomics data is more complex, as there is no *a priori* expectation regarding DNA fragment length and sequence complexity. Denoising procedures therefore focus on the removal of reads with low information content, such as those shorter than ~25-35 bp (Graham et al., 2016; de Filippo et al., 2018; Wang et al., 2021) and/or exhibiting low sequence complexity such as tandem repeats (e.g., Graham et al., 2016; Pedersen et al., 2016). Read duplicates can either be removed during denoising or after alignment (see section 'Identification and validation'). As metagenomics analysis of *sed*aDNA is a recent development, the denoising of these data is currently not standardised between studies.

While there are well-established tools for pre-processing and the denoising of amplicon data, the denoising of metagenomics data is likely to be further developed and optimised to allow for standardisation between data sets collected by different research groups. In addition, iterative testing of settings may be required to determine optimum filtering thresholds that are likely to be data set specific.

Identification and validation

Identification

After data denoising, retained sequences are taxonomically assigned by comparison to a reference database, followed by refinement of these identifications based on the application of filtering thresholds and *a priori* knowledge (Fig. 2).

Taxonomic assignment can be achieved either by aligning whole reads (mapping) or comparing unique equal-size motifs (k-mers) within reads to a reference database. K-mer based assignments are rapid and scalable but can result in high proportions of false positives due to reduced specificity from the minimum k-mer-size limit. Mapping-based approaches have the opposite characteristics. For metagenomics data, widely used software for k-mer based analyses includes KRAKEN2 (Wood et al., 2019), whereas BWA (Li and Durbin, 2009), Bowtie2 (Langmead and Salzberg, 2012), and BLAST (Altschul et al., 1990) are used for mapping with Bowtie2 as the basis for the Holi pipeline (Pedersen et al., 2016). The relatively small size of amplicon datasets means that mapping approaches with software such as ObiTools (Boyer et al., 2016) and Anacapa (Curd et al., 2019) are commonly used.

Reference databases can consist of barcodes (e.g. BOLD (Ratnasingham and Hebert, 2007), SILVA (https://www.arb-silva.de/), mitochondrial or chloroplast genomes (e.g. NCBI RefSeq (O'Leary et al., 2016), PhyloNorway for Arctic plants (Wang et al., 2021), single nuclear genomes, multiple genomes (e.g. NCBI RefSeq, the Genome Taxonomy Database (GTDB; https://gtdb.ecogenomic.org/) for Bacteria and Archaea (Parks et al., 2022)), or global nucleotide databases (e.g. EMBL, NCBI). High quality, complete, and well curated reference databases are essential for accurate taxonomic assignment. It is therefore important to have a representative reference database containing not only taxa expected from the region,

but also common contaminants and taxa closely related to those that are expected (e.g., Graham et al., 2016; Kjær et al., 2022). Database incompleteness can lead to both false negatives, due to the absence of a reference close enough to the *sed*aDNA sequence in question, or false positives, due to the presence of a close relative to the *sed*aDNA sequence that is not the correct taxon. Although global nucleotide databases are often the most complete in terms of taxonomic coverage, they are minimally curated with errors occurring in taxonomic labelling and/or sequence quality within the database. Furthermore, global databases are generally over-represented by model and human-related organisms, such as animal husbandry, crop, and disease taxa, and so caution should be applied when identifying *sed*aDNA, especially if from natural environments, to avoid false positives. On the other hand, barcode databases for amplicon data are generally more straightforward, complete, and well-curated since the genomic region is constrained. Reference databases should therefore be selected carefully with consideration of the research question and data generation strategy used. We highlight that, at present, genomic reference databases are sparse for many groups of non-model organisms, such as invertebrates, fungi, and microbes.

The taxonomic resolution of a *sed*aDNA sequence identification is dependent on its length and whether it originates from a genomic region with diagnostic motifs for assignment at high taxonomic resolution such as the species or population level. For this reason, amplicon and target enrichment experiments are specifically designed to enrich taxonomically-informative genomic regions. On the other hand, taxonomic classification is usually less specific with shotgun metagenomics data, as these data originate from any part of the genome including from large genomic tracts that are shared between organisms. This reduced specificity may be exacerbated if reference genomes are lacking for the target taxa. Regardless of data generation strategy, there will be sequences, especially those that are short (~30-60 bp), that can only be identified to the genus level or higher. These identifications are achieved using lowest common ancestor (LCA) algorithms, which are automatically applied in the amplicon mapping and k-mer softwares listed above, but specialised software exists for other pipelines [e.g., MEGAN6 (Huson et al., 2016), ngsLCA (Wang et al., 2022), PIA (Cribdon et al., 2020)]. LCA algorithms assign the sequences to a taxonomic level that all share the same sequence at a given genetic region. It is therefore important that the used alignment software, such as Bowtie2, outputs all possible alignments and not only the selected best hit.

Setting a threshold can increase the confidence of detections by reducing false positives caused by errors and artefacts. These ‘above the detection limit’ thresholds, or cut-offs, can be based on read counts assigned to a certain taxon, the presence of taxa in the negative controls, or whether the detection is replicated. The latter is particularly applicable to amplicon experiments, where sporadic false-positive detections can be commonplace. However, threshold values can be rather arbitrary and experiment-specific, and so there is a need for better standardization and more objective approaches to setting cut-off values for *sed*aDNA-based detections.

The curation of identifications after bioinformatics-based assignments may be helpful in cases where there is an obvious or well characterised error in the assignment. For example, when using a global database, an assignment may be to an exotic taxon as the expected locally-occurring and closely-related taxon is absent from the database. While this step may

be necessary at present, increased database completeness coupled with incorporating biogeographic data into assignment algorithms should reduce the need to curate data in future.

Validation

A critical stage of bioinformatic processing and interpretation is data set validation using *sed*aDNA authentication evidence (Fig. 2), comparison to independent proxies, and scrutiny of potential taphonomic impacts. For characterising little understood, difficult, or complex study systems, such as abyssal plain and/or deep-time sediments, then simulated or synthetic datasets can be used to test assignment specificity.

Given its propensity for results to include contamination, evidence for the existence of authentic *sed*aDNA molecules should always be presented as part of a study (Smith et al., 2015; Weiß et al., 2015; Kistler et al., 2015). For metagenomics datasets, authentication evidence includes assessment of DNA fragment lengths, which should average <100 bp, and the existence of ancient DNA damage as exhibited by the deamination-induced excess of cytosine-to-thymine misincorporations at the ends of molecules (Dabney et al., 2013). Tools now exist to explore metagenomic *sed*aDNA damage profiles across multiple taxa simultaneously (Michelsen et al., 2022; Everett and Cribdon, 2023). A basal or divergent phylogenetic placement, as compared to a well-sampled modern-day reference panel, may also provide supporting evidence for *sed*aDNA being ancient, but these analyses require genome-wide data and relatively large quantities of reads for a given taxon (Pedersen et al., 2021; Kjær et al., 2022). In both amplicon and metagenomics experiments, negative controls are used to detect contaminant taxa and therefore test the veracity of a sediment sample-derived detection. We note that some taxa, such as domesticated plants and animals, are often sporadically detected in negative controls due to contamination, but may be reproducibly detected in samples, and so results should therefore be interpreted cautiously (see also **Chapter 11**). We highlight that, unless the negative controls were also impacted, cross-contamination cannot necessarily be detected using these authentication approaches.

An independent source of validation is to compare occurrences of taxa in *sed*aDNA and other proxy data sets from the same sediment profile. For example, macrofossils, pollen and coprophilous spore records can be used to cross-validate plant and mammalian *sed*aDNA datasets, respectively (Graham et al., 2016; Clarke et al., 2020; ter Schure et al., 2021; van Vugt et al., 2022; Kjær et al., 2022; Garcés-Pastor et al., 2022). However, it is important to account for proxy-specific taphonomic factors when comparing multiple proxies in this manner. For instance, plant pollen provides a regional signal dominated by wind-dispersed taxa, whereas *sed*aDNA provides a local signal of both wind and pollinator-dispersed plant taxa. The two proxies therefore give different yet complementary patterns (Parducci et al., 2015; Clarke et al., 2020).

Sedimentary ancient DNA data sets are also prone to sources of error, including taphonomy-induced taxon dropout and leaching in some environments that can impact environmental and evolutionary inferences (see section ‘Environmental and evolutionary analysis’; Sand et al., 2023). If DNA from a subset of taxa overwhelms the *sed*aDNA signal, also known as ‘swamping’, due to, for example, a taxonomically heterogenous distribution of

*sed*aDNA within the sediment sample, then this could lead to drop out of the underrepresented taxa resulting in false negatives (Alsos et al., 2020). The vertical leaching of *sed*aDNA between sediment layers, through the movement of porewater, is a serious concern, as this could lead to erroneous interpretations of the timing of events inferred from the *sed*aDNA record (Haile et al., 2007). Leaching can be monitored by comparison to independent proxies and/or assessing the rapidity of turnovers, as a clear rapid turnover would not be apparent from a severely leached sediment profile. Although leaching may be an issue for sediments that allow the movement of porewater, the growing body of literature suggests leaching does not occur in waterlogged sediments (e.g., Graham et al., 2016; Sjögren et al., 2017).

Environmental and evolutionary analysis

Environmental analysis

The environmental analysis of *sed*aDNA data follows the same principles as for other more traditional palaeoecological and archaeological proxies, such as pollen and spores, whereby detections, interpreted as either presence/absence or relative abundance data, are mapped by taxonomy and time. This allows for analysis of co-correlating taxa (communities), periods of community change (turnovers), within- and between-species biogeography (Schulte et al., 2022), and times of first or last occurrence (e.g. Graham et al., 2016; Ficetola et al., 2018; Alsos et al., 2022), as well as the calculation of biodiversity indices (e.g., Chen and Ficetola, 2020) and the identification of independent drivers of changes such as climate or human impact (e.g. Clarke et al., 2019; ter Schure et al., 2021; Garcés-Pastor et al., 2022; Barouillet et al., 2023) However, it is important to account for false negatives in *sed*aDNA-derived environmental datasets (Chen and Ficetola, 2020) and for heterogeneity when combining datasets from multiple sites (Rijal et al., 2021). The key difference over traditional proxies is the taxonomic precision of *sed*aDNA identifications, which can be to the genus, species, and even population level (e.g., Garcés-Pastor et al., 2022) coupled with the potential for scalable and simultaneous analysis of multiple ecosystem components (e.g., Wang et al., 2021; Courtin et al., 2022).

The use of *sed*aDNA abundance data, based on relative read counts assigned to taxa, to infer either relative biomass or relative abundance of individuals, both key measures for many environmental analyses, is controversial and poorly understood in palaeoecological and archaeological systems. This is because multiple biological, taphonomic, and experimental factors can theoretically cause severe deviation between *sed*aDNA abundance and the true past relative biomass or individual abundance of taxa (see also **Chapter 2**; Capo et al., 2021). Controlled experiments (e.g., Alsos et al., 2018) coupled with a greater diversity of palaeoecological and comparative ecological datasets will be needed to understand whether, and if so in what contexts, *sed*aDNA abundance is indicative of past relative biomass or relative individual abundance.

A new and exciting application of *sed*aDNA for environmental inference is using trait analysis to examine the development of environmental, structural, and functional changes

across communities through time. This is possible due to the species discriminating power of *sed*aDNA, but has so far only been applied to plants (Alsos et al., 2022; Revéret et al., 2023) and microorganisms (Keck et al., 2020). We anticipate that the use of ecological traits will be expanded across multiple taxonomic groups thereby allowing for integrated multi-trophic trait analysis to unveil palaeo-ecosystem reconstructions in unprecedented detail.

Evolutionary analysis

Amongst sediment-derived proxies, *sed*aDNA is unique in its ability to simultaneously track evolutionary change in multiple taxa and components of ecosystems through time (e.g., Dussex et al., 2021; Gelabert et al., 2021; Capo et al., 2022). This includes reconstructing mitochondrial and chloroplast genomes for phylogenetics, detecting population turnovers, characterising population histories, tracking functional responses to past environmental stressors (Poulain et al., 2015), and reconstructing genome-wide data for dating the age of *sed*aDNA (Fig. 2).

As *sed*aDNA is derived from entire genomes of multiple taxa, the analysis of reconstructed environmental palaeogenomes or genome-wide datasets can allow for serially-sampled evolutionary analyses of multiple taxa simultaneously. Such analyses are possible if data cover multiple alleles, which include datasets generated by shotgun metagenomics and certain target enrichment or multiplex PCR assays. The major challenge of using these data for evolutionary analyses is that *sed*aDNA is a complex mixture derived from multiple individuals and taxa. This precludes the possibility of some evolutionary analyses, such as those requiring estimates of individual heterozygosity (e.g., inbreeding). However, population structure, turnover, and admixture analyses can be used (e.g., Gelabert et al., 2021; Pedersen et al., 2021; Vernot et al., 2021), as the multi-individual composition of *sed*aDNA (the "environmental genome") is effectively a population sample in a similar manner to a single genome derived from an individual organism (Pedersen et al., 2021). For the reconstruction of mitochondrial and chloroplast genomes (Seersholm et al., 2016; Schulte et al., 2021), which often consist of multiple haplotypes within a population and can be sequenced to high coverage, it is possible to tease apart haplotypes within a single *sed*aDNA extract based on relative allele frequencies. Haplotypes can then be placed into phylogenies (Lammers et al., 2021; Murchie et al., 2022). By comparing the haplotypic and population structure composition of *sed*aDNA extracts across a sediment profile i.e., through time, it is possible to detect population turnovers (e.g., Vernot et al., 2021).

Given *sed*aDNA can be preserved over evolutionarily relevant timescales, molecular tip-dating approaches may be applied to reconstructed haplotypes to provide an independent age estimate of the sediment layer (e.g., Shapiro et al., 2011). Tip dating leverages the fact that ancient haplotypes are missing evolutionary changes that would have evolved during the time since the *sed*aDNA was deposited. By using a reference panel of well-dated comparative sequences and assuming a molecular clock, it is possible to use this missing evolution to estimate an age for the haplotype and therefore the sediment layer. This has been applied to Early Pleistocene sediments from northern Greenland, which showed remarkable agreement between the tip-dated ages of both birch (*Betula*) and mastodon (*Mammut*) haplotypes (Kjær et al., 2022). Molecular clock dating has also been applied to reconstructed mammal

mitochondrial genomes from Late Pleistocene Eurasian cave sediments, including those derived from human (*Homo sapiens*), hominin (*Homo*), wolf (*Canis*), bison (*Bison*), and red deer (*Cervus*) (Gelabert et al., 2021; Vernot et al., 2021; Essel et al., 2023). However, tip dating methods are reliant on the accurate recovery of a single haplotype for each age estimate.

For any evolutionary analyses, it is crucial that *sed*aDNA derived from other taxa is not mis-characterized as originating from the taxon under investigation. To avoid this, loci that are easily differentiated between related taxa can be targeted (Vernot et al., 2021) or competitive mapping approaches can be applied, whereby reads are simultaneously mapped against multiple related reference genomes to determine the most appropriate identity (e.g., Pedersen et al., 2021; Kjær et al., 2022). The ability to conduct population-level evolutionary analyses is dependent on available reference datasets for the taxa under investigation.

The application of evolutionary analyses to *sed*aDNA data is a recent development, with the first mitochondrial and genome-wide reconstructions only published in 2016 and 2021, respectively (Seersholm et al., 2016; Vernot et al., 2021). However, the recovery of sufficient reads for the taxa of interest remains a bottleneck for genomic investigations in metagenomic studies, although target enrichment methods show promising results for specific taxa (e.g., Vernot et al., 2021). Consequently, we anticipate that there will be significant methodological innovation to fully leverage *sed*aDNA for evolutionary analyses in the near future, in particular for the integrated analysis of multiple taxa (Dussex et al., 2021).

Data and script availability

The volume of *sed*aDNA datasets is growing exponentially both in terms of counts of publications (Capo et al., 2021; Von Eggers et al., 2022) and dataset size (e.g., >700 samples in Zavala et al., 2021). Regardless of data generation strategy, these datasets have huge potential for re-analysis, for example due to newly available reference databases exponentially expanding from increased global genome sequencing efforts (Lewin et al., 2018), improved or extended analytical methods (Williams et al., 2023), and/or basic reproducibility. It is therefore crucial that researchers continue to make their raw sequence data publicly and freely available, as is also common in the general ancient metagenomics field (Fellows Yates et al., 2021). Going forward as bioinformatic methods continue to develop, it is increasingly important that scripts, tools, and pipelines are also made publicly available, which will enable continued rapid innovation and discovery in the field (e.g., Cribdon et al., 2020; Lammers et al., 2021; Michelsen et al., 2022). International efforts are also underway to integrate *sed*aDNA data into existing public repositories of palaeoecological and archaeological proxies (e.g., Neotoma Paleoecology Database; Williams et al., 2023).

Future directions

Development and optimisation of the molecular biology and bioinformatics steps of the *sed*aDNA workflow is an ongoing endeavour by the research community. Based on current trends, there is the distinct possibility that target enrichment will supersede the metabarcoding approach, and that shotgun metagenomics may eventually supersede target enrichment. This is due to the ongoing decline in sequencing costs, the increasingly routine automation of laboratory steps (Gansauge et al., 2020), coupled with the accelerating availability of reference genomes (Lewin et al., 2018), although new efficient computational pipelines for *sed*aDNA sequence identification will be required to leverage these increasingly large data sets. In addition to higher throughput in the laboratory, robotisation may contribute to the standardisation of sediment collection and sampling steps and further minimise contamination risks in the field. We speculate that these advances will lead to new applications for *sed*aDNA data, such as the detection of single alleles to infer population history at a site through time (Dussex et al., 2021; Vernot et al., 2021) or using *sed*aDNA damage profiles as a routine relative dating method within a sedimentary sequence.

Summary

Sedimentary ancient DNA is a relatively new and highly promising approach for deriving palaeoenvironmental and evolutionary information from sediments. However, the workflow used to collect, generate, and analyze *sed*aDNA data is complex and experiencing ongoing rapid innovations. In this chapter, we have provided an overview of the *sed*aDNA workflow currently employed in palaeoenvironmental research, by outlining the rationale, general approach, pros and cons, pitfalls, and future perspectives for each workflow step.

Acknowledgements

We thank Eric Capo, Frances Pick, Gentile Francesco Ficetola, Yucheng Wang, and Mikkel Winther Pedersen for comments that greatly improved earlier drafts of this chapter. P.D.H. was supported by the Knut and Alice Wallenberg Foundation (KAW 2021.0048 and KAW 2022.0033). A.R. was supported by the Research Council of Norway (KLIMAFORSK, 294929). L.A. was supported by an Australian Research Council (ARC) Discovery Early Career Researcher (DECRA) Fellowship (DE210100929). S.G.P was supported by the Beatriu de Pinós Programme (BP-2021-00131).

Author contributions

P.D.H., T.M., K.N., B.V., and L.A. designed the chapter. P.D.H., K.N., A.R., T.M., S.G.P and Y.L. drafted the chapter. All authors revised the text.

References

Alpaslan-Roodenberg S, Anthony D, Babiker H, Bánffy E, Booth T, Capone P, Deshpande-Mukherjee A, Eisenmann S, Fehren-Schmitz L, Frachetti M, Fujita R, Frieman CJ, Fu Q, Gibbon V, Haak W, Hajdinjak M, Hofmann KP, Holguin B, Inomata T, Kanzawa-Kiriyama H, Keegan W, Kelso J, Krause J, Kumaresan G, Kusimba C, Kusimba S, Lalueza-Fox C, Llamas B, MacEachern S, Mallick S, Matsumura H, Morales-Arce AY, Matuzeviciute GM, Mushrif-Tripathy V, Nakatsuka N, Nores R, Ogola C, Okumura M, Patterson N, Pinhasi R, Prasad SPR, Prendergast ME, Punzo JL, Reich D, Sawafuji R, Sawchuk E, Schiffels S, Sedig J, Shnaider S, Sirak K, Skoglund P, Slon V, Snow M, Soressi M, Spriggs M, Stockhammer PW, Szécsényi-Nagy A, Thangaraj K, Tiesler V, Tobler R, Wang C-C, Warinner C, Yasawardene S, Zahir M (2021) Ethics of DNA research on human remains: five globally applicable guidelines. Nature 599(7883):41–46. https://doi.org/10.1038/s41586-021-04008-x

Alsos IG, Lammers Y, Yoccoz NG, Jørgensen T, Sjögren P, Gielly L, Edwards ME (2018) Plant DNA metabarcoding of lake sediments: How does it represent the contemporary vegetation. PLOS ONE 13(4):e0195403. https://doi.org/10.1371/journal.pone.0195403

Alsos IG, Sjögren P, Brown AG, Gielly L, Merkel MKF, Paus A, Lammers Y, Edwards ME, Alm T, Leng M, Goslar T, Langdon CT, Bakke J, van der Bilt WGM (2020) Last Glacial Maximum environmental conditions at Andøya, northern Norway; evidence for a northern ice-edge ecological "hotspot." Quaternary Science Reviews 239:106364. https://doi.org/10.1016/j.quascirev.2020.106364

Alsos IG, Rijal DP, Ehrich D, Karger DN, Yoccoz NG, Heintzman PD, Brown AG, Lammers Y, Pellissier L, Alm T, Bråthen KA, Coissac E, Merkel MKF, Alberti A, Denoeud F, Bakke J, PHYLONORWAY CONSORTIUM (2022) Postglacial species arrival and diversity buildup of northern ecosystems took millennia. Science Advances 8(39):eabo7434. https://doi.org/10.1126/sciadv.abo7434

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215(3):403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Andres KJ, Sethi SA, Lodge DM, Andrés J (2021) Nuclear eDNA estimates population allele frequencies and abundance in experimental mesocosms and field samples. Mol Ecol 30(3):685–697. https://doi.org/10.1111/mec.15765

Ardelean CF, Becerra-Valdivia L, Pedersen MW, Schwenninger JL, 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-Gutierrez 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(7819):87–92. https://doi.org/10.1038/s41586-020-2509-0

Armbrecht LH, Coolen MJL, Lejzerowicz F, George SC, Negandhi K, Suzuki Y, Young J, Foster NR, Armand LK, Cooper A, Ostrowski M, Focardi A, Stat M, Moreau JW, Weyrich LS (2019) Ancient DNA from marine sediments: Precautions and considerations for seafloor coring, sample handling and data generation. Earth-Science Reviews 196:102887. https://doi.org/10.1016/j.earscirev.2019.102887

Armbrecht 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. Molecular Ecology Resources 20(4):906–919. https://doi.org/10.1111/1755-0998.13162

Armbrecht 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

Armbrecht L, Weber ME, Raymo ME, Peck VL, Williams T, Warnock J, Kato Y, Hernández-Almeida I, Hoem F, Reilly B, Hemming S, Bailey I, Martos YM, Gutjahr M, Percuoco V, Allen C, Brachfeld S, Cardillo FG, Du Z, Fauth G, Fogwill C, Garcia M, Glüder A, Guitard M, Hwang J-H, Iizuka M, Kenlee B, O'Connell S, Pérez LF, Ronge TA, Seki O, Tauxe L, Tripathi S, Zheng X (2022) Ancient marine sediment DNA reveals diatom transition in Antarctica. Nat Commun 13(1):5787. https://doi.org/10.1038/s41467-022-33494-4

Bardill J, Bader AC, Garrison NA, Bolnick DA, Raff JA, Walker A, Malhi RS, THE SUMMER INTERNSHIP FOR INDIGENOUS PEOPLES IN GENOMICS (SING) CONSORTIUM (2018) Advancing the ethics of paleogenomics. Science 360(6387):384–385. https://doi.org/10.1126/science.aaq1131

Barouillet C, Monchamp M-E, Bertilsson S, Brasell K, Domaizon I, Epp LS, Ibrahim A, Mejbel H, Nwosu EC, Pearman JK, Picard M, Thomson-Laing G, Tsugeki N, Von Eggers J, Gregory-Eaves I, Pick FR, Wood SA, Capo E (2023) Investigating the effects of anthropogenic stressors on lake biota using sedimentary DNA. Freshwater Biology n/a(n/a). https://doi.org/10.1111/fwb.14027

Binladen J, Gilbert MTP, Bollback JP, Panitz F, Bendixen C, Nielsen R, Willerslev E (2007) The Use of Coded PCR Primers Enables High-Throughput Sequencing of Multiple Homolog Amplification Products by 454 Parallel Sequencing. PLOS ONE 2(2):e197. https://doi.org/10.1371/journal.pone.0000197

Boessenkool S, Epp LS, Haile J, Bellemain E, Edwards M, Coissac E, Willerslev E, Brochmann C (2012) Blocking human contaminant DNA during PCR allows amplification of rare mammal species from sedimentary ancient DNA. Mol Ecol 21(8):1806–1815. https://doi.org/10.1111/j.1365-294X.2011.05306.x

Boyer F, Mercier C, Bonin A, Le Bras Y, Taberlet P, Coissac E (2016) obitools: a unix-inspired software package for DNA metabarcoding. Mol Ecol Resour 16(1):176–182. https://doi.org/10.1111/1755-0998.12428

Braadbaart F, Reidsma FH, Roebroeks W, Chiotti L, Slon V, Meyer M, Théry-Parisot I, van Hoesel A, Nierop KGJ, Kaal J, van Os B, Marquer L (2020) Heating histories and taphonomy of ancient fireplaces: A multi-proxy case study from the Upper Palaeolithic sequence of Abri Pataud (Les Eyzies-de-Tayac, France). https://doi.org/10.13039/501100004189

Brosseau L, Udom C, Sukkanon C, Chareonviriyaphap T, Bangs MJ, Saeung A, Manguin S (2019) A multiplex PCR assay for the identification of five species of the Anopheles barbirostris complex in Thailand. Parasites & Vectors 12(1):223. https://doi.org/10.1186/s13071-019-3494-8

Buck M, Hamilton C (2011) The 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. Review of European Community & International Environmental Law 20(1):47–61. https://doi.org/10.1111/j.1467-9388.2011.00703.x

Capo E, Debroas D, Arnaud F, Perga M-E, Chardon C, Domaizon I (2017) Tracking a century of changes in microbial eukaryotic diversity in lakes driven by nutrient enrichment and climate warming. Environmental Microbiology 19(7):2873–2892. https://doi.org/10.1111/1462-2920.13815

Capo E, Giguet-Covex C, Rouillard A, Nota K, Heintzman PD, Vuillemin A, Ariztegui D, Arnaud F, Belle S, Bertilsson S, Bigler C, Bindler R, Brown AG, Clarke CL, Crump SE, Debroas D, Englund G, Ficetola GF, Garner RE, Gauthier J, Gregory-Eaves I, Heinecke L, Herzschuh U, Ibrahim A, Kisand V, Kjær KH, Lammers Y, Littlefair J, Messager E, Monchamp M-E, Olajos F, Orsi W, Pedersen MW, Rijal DP, Rydberg J, Spanbauer T, Stoof-Leichsenring KR, Taberlet P, Talas L, Thomas C, Walsh DA, Wang Y, Willerslev E, van Woerkom A, Zimmermann HH, Coolen MJL, Epp LS, Domaizon I, G. Alsos I, Parducci L (2021) Lake Sedimentary DNA Research on Past Terrestrial and Aquatic Biodiversity: Overview and Recommendations. Quaternary 4(1):6. https://doi.org/10.3390/quat4010006

Capo E, Monchamp M-E, Coolen MJL, Domaizon I, Armbrecht L, Bertilsson S (2022) Environmental paleomicrobiology: using DNA preserved in aquatic sediments to its full potential. Environmental Microbiology 24(5):2201–2209. https://doi.org/10.1111/1462-2920.15913

Chen W, Ficetola GF (2020) Numerical methods for sedimentary-ancient-DNA-based study on past biodiversity and ecosystem functioning. Environmental DNA 2(2):115–129. https://doi.org/10.1002/edn3.79

Clarke CL, Edwards ME, Gielly L, Ehrich D, Hughes PDM, Morozova LM, Haflidason H, Mangerud J, Svendsen JI, Alsos IG (2019) Persistence of arctic-alpine flora during 24,000 years of environmental change in the Polar Urals. Sci Rep 9(1):19613. https://doi.org/10.1038/s41598-019-55989-9

Clarke CL, Alsos IG, Edwards ME, Paus A, Gielly L, Haflidason H, Mangerud J, Regnéll C, Hughes PDM, Svendsen JI, Bjune AE (2020) A 24,000-year ancient DNA and pollen record from the Polar Urals reveals temporal dynamics of arctic and boreal plant communities. Quaternary Science Reviews 247:106564. https://doi.org/10.1016/j.quascirev.2020.106564

Coolen MJL, Overmann J (1998) Analysis of Subfossil Molecular Remains of Purple Sulfur Bacteria in a Lake Sediment. Applied and Environmental Microbiology 64(11):4513–4521. https://doi.org/10.1128/AEM.64.11.4513-4521.1998

Côté NML, Daligault J, Pruvost M, Bennett EA, Gorgé O, Guimaraes S, Capelli N, Bailly ML, Geigl E-M, Grange T (2016) A New High-Throughput Approach to Genotype Ancient Human Gastrointestinal Parasites. PLOS ONE 11(1):e0146230. https://doi.org/10.1371/journal.pone.0146230

Courtin J, Perfumo A, Andreev AA, Opel T, Stoof-Leichsenring KR, Edwards ME, Murton JB, Herzschuh U (2022) Pleistocene glacial and interglacial ecosystems inferred from ancient DNA analyses of permafrost sediments from Batagay megaslump, East Siberia. Environmental DNA 4(6):1265–1283. https://doi.org/10.1002/edn3.336

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. Frontiers in Ecology and Evolution 8

Crump SE (2021) Sedimentary ancient DNA as a tool in paleoecology. Nat Rev Earth Environ 2(4):229–229. https://doi.org/10.1038/s43017-021-00158-8

Crump SE, Fréchette B, Power M, Cutler S, de Wet G, Raynolds MK, Raberg JH, Briner JP, Thomas EK, Sepúlveda J, Shapiro B, Bunce M, Miller GH (2021) Ancient plant DNA reveals High Arctic greening during the Last Interglacial. Proceedings of the National Academy of Sciences 118(13):e2019069118. https://doi.org/10.1073/pnas.2019069118

Curd EE, Gold Z, Kandlikar GS, Gomer J, Ogden M, O'Connell T, Pipes L, Schweizer TM, Rabichow L, Lin M, Shi B, Barber PH, Kraft N, Wayne R, Meyer RS (2019) Anacapa Toolkit: An environmental DNA toolkit for processing multilocus metabarcode datasets. Methods in Ecology and Evolution 10(9):1469–1475. https://doi.org/10.1111/2041-210X.13214

Dabney J, Meyer M, Pääbo S (2013) Ancient DNA damage. Cold Spring Harb Perspect Biol 5(7):a012567. https://doi.org/10.1101/cshperspect.a012567

de Filippo C, Meyer M, Prüfer K (2018) Quantifying and reducing spurious alignments for the analysis of ultra-short ancient DNA sequences. BMC Biology 16(1):121. https://doi.org/10.1186/s12915-018-0581-9

Domaizon I, Savichtcheva O, Debroas D, Arnaud F, Villar C, Pignol C, Alric B, Perga ME (2013) DNA from lake sediments reveals the long-term dynamics and diversity of *Synechococcus* assemblages. Biogeosciences 10(6):3817–3838. https://doi.org/10.5194/bg-10-3817-2013

Dussex N, Bergfeldt N, de Anca Prado V, Dehasque M, Díez-del-Molino D, Ersmark E, Kanellidou F, Larsson P, Lemež Š, Lord E, Mármol-Sánchez E, Meleg IN, Måsviken J, Naidoo T, Studerus J, Vicente M, von Seth J, Götherström A, Dalén L, Heintzman PD (2021) Integrating multi-taxon palaeogenomes and sedimentary ancient DNA to study past ecosystem dynamics. Proceedings of the Royal Society B: Biological Sciences 288(1957):20211252. https://doi.org/10.1098/rspb.2021.1252

Eisenstein M (2023) Illumina faces short-read rivals. Nat Biotechnol 41(1):3–5. https://doi.org/10.1038/s41587-022-01632-4

Epp LS, Zimmermann HH, Stoof-Leichsenring KR (2019) Sampling and extraction of ancient DNA from sediments. In: Ancient DNA: Methods and protocols. Springer, pp 31–44

Essel E, Zavala EI, Schulz-Kornas E, Kozlikin MB, Fewlass H, Vernot B, Shunkov MV, Derevianko AP, Douka K, Barnes I, Soulier M-C, Schmidt A, Szymanski M, Tsanova T, Sirakov N, Endarova E, McPherron SP, Hublin J-J, Kelso J, Pääbo S, Hajdinjak M, Soressi M, Meyer M (2023) Ancient human DNA recovered from a Palaeolithic pendant. Nature :1–5. https://doi.org/10.1038/s41586-023-06035-2

Everett R, Cribdon B (2023) MetaDamage tool: Examining post-mortem damage in sedaDNA on a metagenomic scale. Frontiers in Ecology and Evolution 10 https://doi.org/10.3389/fevo.2022.888421

Fellows Yates JA, Andrades Valtueña A, Vågene 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 metadata of published ancient metagenomic samples with AncientMetagenomeDir. Sci Data 8(1):31. https://doi.org/10.1038/s41597-021-00816-y

Ficetola GF, Pansu J, Bonin A, Coissac E, Giguet-Covex C, De Barba M, Gielly L, Lopes CM, Boyer F, Pompanon F, Rayé G, Taberlet P (2015) Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. Molecular Ecology Resources 15(3):543–556. https://doi.org/10.1111/1755-0998.12338

Ficetola GF, Poulenard J, Sabatier P, Messager E, Gielly L, Leloup A, Etienne D, Bakke J, Malet E, Fanget B, Støren E, Reyss J-L, Taberlet P, Arnaud F (2018) DNA from lake sediments reveals long-term ecosystem changes after a biological invasion. Science Advances 4(5):eaar4292. https://doi.org/10.1126/sciadv.aar4292

Ficetola GF, Taberlet P (2023) Towards exhaustive community ecology via DNA metabarcoding. Mol Ecol. https://doi.org/10.1111/mec.16881

Frøslev TG, Kjøller R, Bruun HH, Ejrnæs R, Brunbjerg AK, Pietroni C, Hansen AJ (2017) Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. Nat Commun 8(1):1188. https://doi.org/10.1038/s41467-017-01312-x

Fulton TL, Shapiro B (2019) Setting Up an Ancient DNA Laboratory. Methods Mol Biol 1963:1–13. https://doi.org/10.1007/978-1-4939-9176-1_1

Gansauge M-T, Gerber T, Glocke I, Korlevic P, Lippik L, Nagel S, Riehl LM, Schmidt A, Meyer M (2017) Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. Nucleic Acids Res 45(10):e79. https://doi.org/10.1093/nar/gkx033

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(8):2279–2300. https://doi.org/10.1038/s41596-020-0338-0
Garcés-Pastor S, Coissac E, Lavergne S, Schwörer C, Theurillat J-P, Heintzman PD, Wangensteen OS, Tinner W, Rey F, Heer M, Rutzer A, Walsh K, Lammers Y, Brown AG, Goslar T, Rijal DP, Karger DN, Pellissier L, Heiri O, Alsos IG (2022) High resolution ancient sedimentary DNA shows that alpine plant diversity is associated with human land use and climate change. Nat Commun 13(1):6559. https://doi.org/10.1038/s41467-022-34010-4
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(16):3564-3574.e9. https://doi.org/10.1016/j.cub.2021.06.023
Graham RW, Belmecheri S, Choy K, Culleton BJ, Davies LJ, Froese D, Heintzman PD, Hritz C, Kapp JD, Newsom LA, Rawcliffe R, Saulnier-Talbot É, Shapiro B, Wang Y, Williams JW, Wooller MJ (2016) Timing and causes of mid-Holocene mammoth extinction on St. Paul Island, Alaska. Proceedings of the National Academy of Sciences 113(33):9310–9314. https://doi.org/10.1073/pnas.1604903113
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? Molecular Biology and Evolution 24(4):982–989. https://doi.org/10.1093/molbev/msm016
Handsley-Davis M, Kowal E, Russell L, Weyrich LS (2021) Researchers using environmental DNA must engage ethically with Indigenous communities. https://doi.org/10.1038/s41559-020-01351-6
Hofreiter M, Mead JI, Martin P, Poinar HN (2003) Molecular caving. Curr Biol 13(18):R693-695. https://doi.org/10.1016/j.cub.2003.08.039
Huson DH, Auch AF, Qi J, Schuster SC (2007) MEGAN analysis of metagenomic data. Genome Res 17(3):377–386. https://doi.org/10.1101/gr.5969107
Immel A, Le Cabec A, Bonazzi M, Herbig A, Temming H, Schuenemann VJ, Bos KI, Langbein F, Harvati K, Bridault A, Pion G, Julien M-A, Krotova O, Conard NJ, Münzel SC, Drucker DG, Viola B, Hublin J-J, Tafforeau P, Krause J (2016) Effect of X-ray irradiation on ancient DNA in sub-fossil bones – Guidelines for safe X-ray imaging. Sci Rep 6(1):32969. https://doi.org/10.1038/srep32969
Jelavic S, Lanson B, Findling N, Lanson M (2023) A route for long-term DNA preservation through nanoconfinement in smectites
Kapp JD, Green RE, Shapiro B (2021) A Fast and Efficient Single-stranded Genomic Library Preparation Method Optimized for Ancient DNA. J Hered 112(3):241–249. https://doi.org/10.1093/jhered/esab012
Keck F, Millet L, Debroas D, Etienne D, Galop D, Rius D, Domaizon I (2020) Assessing the response of micro-eukaryotic diversity to the Great Acceleration using lake sedimentary DNA. Nat Commun 11(1):3831. https://doi.org/10.1038/s41467-020-17682-8
Kircher M, Sawyer S, Meyer M (2012) Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Res 40(1):e3. https://doi.org/10.1093/nar/gkr771
Kistler L, Smith O, Ware R, Momber G, Bates R, Garwood P, Fitch S, Pallen M, Gaffney V, Allaby RG (2015) Thermal age, cytosine deamination and the veracity of 8,000 year old wheat DNA from sediments. 032060
Kjær KH, Winther Pedersen M, De Sanctis B, De Cahsan B, Korneliussen TS, Michelsen CS, Sand KK, Jelavić S, Ruter AH, Schmidt AMA, Kjeldsen KK, Tesakov AS, Snowball I, Gosse JC, Alsos IG, Wang Y, Dockter C, Rasmussen M, Jørgensen ME, Skadhauge B, Prohaska A, Kristensen JÅ, Bjerager M, Allentoft ME, Coissac E, Rouillard A, Simakova A, Fernandez-Guerra A, Bowler C, Macias-Fauria M, Vinner L, Welch JJ, Hidy AJ, Sikora M, Collins MJ, Durbin R, Larsen NK, Willerslev E (2022) A 2-million-year-old ecosystem in Greenland uncovered by environmental DNA. Nature 612(7939):283–291. https://doi.org/10.1038/s41586-022-05453-y
Kowal E, Weyrich LS, Argüelles JM, Bader AC, Colwell C, Cortez AD, Davis JL, Figueiro G, Fox K, Malhi RS, Matisoo-Smith E, Nayak A, Nelson EA, Nicholas G, Nieves-Colón MA, Russell L, Ulm S, Vergara-Silva F, Villanea FA, Wagner JK, Yracheta JM, Tsosie KS (2023) Community partnerships are fundamental to ethical ancient DNA research. Human Genetics and Genomics Advances 4(2):100161. https://doi.org/10.1016/j.xhgg.2022.100161
Lammers Y, Heintzman PD, Alsos IG (2021) Environmental palaeogenomic reconstruction of an Ice Age algal population. Commun Biol 4(1):1–11. https://doi.org/10.1038/s42003-021-01710-4
Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9(4):357–359. https://doi.org/10.1038/nmeth.1923

Lewin HA, Robinson GE, Kress WJ, Baker WJ, Coddington J, Crandall KA, Durbin R, Edwards SV, Forest F, Gilbert MTP, Goldstein MM, Grigoriev IV, Hackett KJ, Haussler D, Jarvis ED, Johnson WE, Patrinos A, Richards S, Castilla-Rubio JC, van Sluys M-A, Soltis PS, Xu X, Yang H, Zhang G (2018) Earth BioGenome Project: Sequencing life for the future of life. Proceedings of the National Academy of Sciences 115(17):4325–4333. https://doi.org/10.1073/pnas.1720115115

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14):1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Liew M, Pryor R, Palais R, Meadows C, Erali M, Lyon E, Wittwer C (2004) Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. Clin Chem 50(7):1156–1164. https://doi.org/10.1373/clinchem.2004.032136

MacConaill LE, Burns RT, Nag A, Coleman HA, Slevin MK, Giorda K, Light M, Lai K, Jarosz M, McNeill MS, Ducar MD, Meyerson M, Thorner AR (2018) Unique, dual-indexed sequencing adapters with UMIs effectively eliminate index cross-talk and significantly improve sensitivity of massively parallel sequencing. BMC Genomics 19(1):30. https://doi.org/10.1186/s12864-017-4428-5

Maricic T, Whitten M, Pääbo S (2010) Multiplexed DNA Sequence Capture of Mitochondrial Genomes Using PCR Products. PLOS ONE 5(11):e14004. https://doi.org/10.1371/journal.pone.0014004

Massilani D, Morley MW, Mentzer SM, Aldeias V, Vernot B, Miller C, Stahlschmidt M, Kozlikin MB, Shunkov MV, Derevianko AP, Conard NJ, Wurz S, Henshilwood CS, Vasquez J, Essel E, Nagel S, Richter J, Nickel B, Roberts RG, Pääbo S, Slon V, Goldberg P, Meyer M (2022) Microstratigraphic preservation of ancient faunal and hominin DNA in Pleistocene cave sediments. Proceedings of the National Academy of Sciences 119(1):e2113666118. https://doi.org/10.1073/pnas.2113666118

Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc 2010(6):pdb.prot5448. https://doi.org/10.1101/pdb.prot5448

Michelsen C, Pedersen MW, Fernandez-Guerra A, Zhao L, Petersen TC, Korneliussen TS (2022) metaDMG – A Fast and Accurate Ancient DNA Damage Toolkit for Metagenomic Data. 2022.12.06.519264

Murchie TJ, Kuch M, Duggan AT, Ledger ML, Roche K, Klunk J, Karpinski E, Hackenberger D, Sadoway T, MacPhee R, Froese D, Poinar H (2021) Optimizing extraction and targeted capture of ancient environmental DNA for reconstructing past environments using the PalaeoChip Arctic-1.0 bait-set. Quaternary Research 99:305–328. https://doi.org/10.1017/qua.2020.59

Murchie TJ, Karpinski E, Eaton K, Duggan AT, Baleka S, Zazula G, MacPhee RDE, Froese D, Poinar HN (2022) Pleistocene mitogenomes reconstructed from the environmental DNA of permafrost sediments. Current Biology 32(4):851-860.e7. https://doi.org/10.1016/j.cub.2021.12.023

Nesje, A. (1992). A Piston Corer for Lacustrine and Marine Sediments. *Arctic and Alpine Research*, *24*(3), 257–259. https://doi.org/10.1080/00040851.1992.12002956

Nichols RV, Vollmers C, Newsom LA, Wang Y, Heintzman PD, Leighton M, Green RE, Shapiro B (2018) Minimizing polymerase biases in metabarcoding. Molecular Ecology Resources 18(5):927–939. https://doi.org/10.1111/1755-0998.12895

Nichols RV, Curd E, Heintzman PD, Shapiro B (2019) Targeted Amplification and Sequencing of Ancient Environmental and Sedimentary DNA. Methods Mol Biol 1963:149–161. https://doi.org/10.1007/978-1-4939-9176-1_16

Nota K, Klaminder J, Milesi P, Bindler R, Nobile A, van Steijn T, Bertilsson S, Svensson B, Hirota SK, Matsuo A, Gunnarsson U, Seppä H, Väliranta MM, Wohlfarth B, Suyama Y, Parducci L (2022) Norway spruce postglacial recolonization of Fennoscandia. Nat Commun 13(1):1333. https://doi.org/10.1038/s41467-022-28976-4

Nwosu EC, Brauer A, Monchamp M-E, Pinkerneil S, Bartholomäus A, Theuerkauf M, Schmidt J-P, Stoof-Leichsenring KR, Wietelmann T, Kaiser J, Wagner D, Liebner S (2023) Early human impact on lake cyanobacteria revealed by a Holocene record of sedimentary ancient DNA. Commun Biol 6(1):1–12. https://doi.org/10.1038/s42003-023-04430-z

O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, Astashyn A, Badretdin A, Bao Y, Blinkova O, Brover V, Chetvernin V, Choi J, Cox E, Ermolaeva O, Farrell CM, Goldfarb T, Gupta T, Haft D, Hatcher E, Hlavina W, Joardar VS, Kodali VK, Li W, Maglott D, Masterson P, McGarvey KM, Murphy MR, O'Neill K, Pujar S, Rangwala SH, Rausch D, Riddick LD, Schoch C, Shkeda A, Storz SS, Sun H, Thibaud-Nissen F, Tolstoy I, Tully RE, Vatsan AR, Wallin C, Webb D, Wu W, Landrum MJ, Kimchi A, Tatusova T, DiCuccio M, Kitts P, Murphy TD, Pruitt KD (2016) Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 44(D1):D733-745. https://doi.org/10.1093/nar/gkv1189

Pal S, Gregory-Eaves I, Pick FR (2015) Temporal trends in cyanobacteria revealed through DNA and pigment analyses of temperate lake sediment cores. J Paleolimnol 54(1):87–101. https://doi.org/10.1007/s10933-015-9839-1

Parducci L, Väliranta M, Salonen JS, Ronkainen T, Matetovici I, Fontana SL, Eskola T, Sarala P, Suyama Y (2015) Proxy comparison in ancient peat sediments: pollen, macrofossil and plant DNA. Philos Trans R Soc Lond B Biol Sci 370(1660):20130382. https://doi.org/10.1098/rstb.2013.0382

Parducci L, Bennett KD, Ficetola GF, Alsos IG, Suyama Y, Wood JR, Pedersen MW (2017) Ancient plant DNA in lake sediments. New Phytologist 214(3):924–942. https://doi.org/10.1111/nph.14470

Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil P-A, Hugenholtz P (2022) GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. Nucleic Acids Research 50(D1):D785–D794. https://doi.org/10.1093/nar/gkab776

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, Kjær KH, Willerslev E (2016) Postglacial viability and colonization in North America's ice-free corridor. Nature 537(7618):45–49. https://doi.org/10.1038/nature19085

Pedersen MW, De Sanctis B, Saremi NF, Sikora M, Puckett EE, Gu Z, Moon KL, Kapp JD, Vinner L, Vardanyan Z, Ardelean CF, Arroyo-Cabrales J, Cahill JA, Heintzman PD, Zazula G, MacPhee RDE, Shapiro B, Durbin R, Willerslev E (2021) Environmental genomics of Late Pleistocene black bears and giant short-faced bears. Current Biology 31(12):2728-2736.e8. https://doi.org/10.1016/j.cub.2021.04.027

Picard M, Pochon X, Atalah J, Pearman JK, Rees A, Howarth JD, Moy CM, Vandergoes MJ, Hawes I, Khan S, Wood SA (2022) Using metabarcoding and droplet digital PCR to investigate drivers of historical shifts in cyanobacteria from six contrasting lakes. Sci Rep 12(1):12810. https://doi.org/10.1038/s41598-022-14216-8

Pilon S, Zastepa A, Taranu ZE, Gregory-Eaves I, Racine M, Blais JM, Poulain AJ, Pick FR (2019) Contrasting histories of microcystin-producing cyanobacteria in two temperate lakes as inferred from quantitative sediment DNA analyses. Lake and Reservoir Management 35(1):102–117. https://doi.org/10.1080/10402381.2018.1549625

Poulain AJ, Aris-Brosou S, Blais JM, Brazeau M, Keller W (Bill), Paterson AM (2015) Microbial DNA records historical delivery of anthropogenic mercury. ISME J 9(12):2541–2550. https://doi.org/10.1038/ismej.2015.86

Ratnasingham S, Hebert PDN (2007) bold: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes 7(3):355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Revéret A, Rijal DP, Heintzman PD, Brown AG, Stoof-Leichsenring KR, Alsos IG (2023) Environmental DNA of aquatic macrophytes: the potential for reconstructing past and present vegetation and environments. 2023.03.27.533457

Rijal DP, Heintzman PD, Lammers Y, Yoccoz NG, Lorberau KE, Pitelkova I, Goslar T, Murguzur FJA, Salonen JS, Helmens KF, Bakke J, Edwards ME, Alm T, Bråthen KA, Brown AG, Alsos IG (2021) Sedimentary ancient DNA shows terrestrial plant richness continuously increased over the Holocene in northern Fennoscandia. Science Advances 7(31):eabf9557. https://doi.org/10.1126/sciadv.abf9557

Rodriguez-Martinez S, Klaminder J, Morlock MA, Dalén L, Huang DY-T (2023) The topological nature of tag jumping in environmental DNA metabarcoding studies. Molecular Ecology Resources 23(3):621–631. https://doi.org/10.1111/1755-0998.13745

Ross KS, Haites NE, Kelly KF (1990) Repeated freezing and thawing of peripheral blood and DNA in suspension: effects on DNA yield and integrity. Journal of Medical Genetics 27(9):569–570. https://doi.org/10.1136/jmg.27.9.569

Safarikova M, Kubena AA, Frankova V, Zima T, Kalousova M (2021) The Effects of Different Storage Conditions and Repeated Freeze/Thaw Cycles on the Concentration, Purity and Integrity of Genomic DNA. Folia Biol (Praha) 67(1):10–15

Sand KK, Jelavić S, Kjær KH, Prohaska A (2023) Importance of eDNA taphonomy and provenance for robust ecological inference: insights from interfacial geochemistry. 2023.01.24.525431

Savichtcheva O, Debroas D, Kurmayer R, Villar C, Jenny JP, Arnaud F, Perga ME, Domaizon I (2011) Quantitative PCR Enumeration of Total/Toxic Planktothrix rubescens and Total Cyanobacteria in Preserved DNA Isolated from Lake Sediments. Applied and Environmental Microbiology 77(24):8744–8753. https://doi.org/10.1128/AEM.06106-11

Schnurrenberger DW, Kelts KR, Johnson TC, Shane LCK, Ito E (2001) National lacustrine core repository (LacCore). Journal of Paleolimnology 25(1):123–127. https://doi.org/10.1023/A:1008171027125

Schubert M, Lindgreen S, Orlando L (2016) AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Research Notes 9(1):88. https://doi.org/10.1186/s13104-016-1900-2

Schulte L, Bernhardt N, Stoof-Leichsenring K, Zimmermann HH, Pestryakova LA, Epp LS, Herzschuh U (2021) Hybridization capture of larch (Larix Mill.) chloroplast genomes from sedimentary ancient DNA reveals past changes of Siberian forest. Molecular Ecology Resources 21(3):801–815. https://doi.org/10.1111/1755-0998.13311

Schulte L, Meucci S, Stoof-Leichsenring KR, Heitkam T, Schmidt N, von Hippel B, Andreev AA, Diekmann B, Biskaborn BK, Wagner B, Melles M, Pestryakova LA, Alsos IG, Clarke C, Krutovsky KV, Herzschuh U (2022) Larix species range dynamics in Siberia since the Last Glacial captured from sedimentary ancient DNA. Commun Biol 5(1):1–11. https://doi.org/10.1038/s42003-022-03455-0

Seeber PA, von Hippel B, Kauserud H, Löber U, Stoof-Leichsenring KR, Herzschuh U, Epp LS (2022) Evaluation of lake sedimentary ancient DNA metabarcoding to assess fungal biodiversity in Arctic paleoecosystems. Environmental DNA 4(5):1150–1163. https://doi.org/10.1002/edn3.315

Seersholm FV, Pedersen MW, Søe MJ, Shokry H, Mak SST, Ruter A, Raghavan M, Fitzhugh W, Kjær KH, Willerslev E, Meldgaard M, Kapel CMO, Hansen AJ (2016) DNA evidence of bowhead whale exploitation by Greenlandic Paleo-Inuit 4,000 years ago. Nat Commun 7(1):13389. https://doi.org/10.1038/ncomms13389

Selway CA, Armbrecht L, Thornalley D (2022) An Outlook for the Acquisition of Marine Sedimentary Ancient DNA (sedaDNA) From North Atlantic Ocean Archive Material. Paleoceanography and Paleoclimatology 37(5):e2021PA004372. https://doi.org/10.1029/2021PA004372

Shao W, Khin S, Kopp WC (2012) Characterization of effect of repeated freeze and thaw cycles on stability of genomic DNA using pulsed field gel electrophoresis. Biopreserv Biobank 10(1):4–11. https://doi.org/10.1089/bio.2011.0016

Shapiro B, Ho SYW, Drummond AJ, Suchard MA, Pybus OG, Rambaut A (2011) A Bayesian phylogenetic method to estimate unknown sequence ages. Mol Biol Evol 28(2):879–887. https://doi.org/10.1093/molbev/msq262

Sjögren P, Edwards ME, Gielly L, Langdon CT, Croudace IW, Merkel MKF, Fonville T, Alsos IG (2017) Lake sedimentary DNA accurately records 20th Century introductions of exotic conifers in Scotland. New Phytologist 213(2):929–941. https://doi.org/10.1111/nph.14199

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, Kućan 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(6338):605–608. https://doi.org/10.1126/science.aam9695

Smith O, Momber G, Bates R, Garwood P, Fitch S, Pallen M, Gaffney V, Allaby RG (2015) Archaeology. Sedimentary DNA from a submerged site reveals wheat in the British Isles 8000 years ago. Science 347(6225):998–1001. https://doi.org/10.1126/science.1261278

Soares AER (2019) Hybridization Capture of Ancient DNA Using RNA Baits. In: Shapiro B, Barlow A, Heintzman PD, Hofreiter M, Paijmans JLA, Soares AER (eds) Ancient DNA: Methods and Protocols. Springer, New York, NY, pp 121–128

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G, Brochmann C, Willerslev E (2007) Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acids Res 35(3):e14. https://doi.org/10.1093/nar/gkl938

Taberlet P, Bonin A, Zinger L, Coissac E (2018a). DNA amplification and multiplexing. In P Taberlet, A Bonin, L Zinger, E Coissac (Eds.), *Environmental DNA: For Biodiversity Research and Monitoring*. Oxford University Press. https://doi.org/10.1093/oso/9780198767220.003.0006

Taberlet P, Bonin A, Zinger L, Coissac E (2018b). DNA metabarcode choice and design. In P Taberlet, A Bonin, L Zinger, E Coissac (Eds.), *Environmental DNA: For Biodiversity Research and Monitoring* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780198767220.003.0002

ter Schure ATM, Bajard M, Loftsgarden K, Høeg HI, Ballo E, Bakke J, Støren EWN, Iversen F, Kool A, Brysting AK, Krüger K, Boessenkool S (2021) Anthropogenic and environmental drivers of vegetation change in southeastern Norway during the Holocene. Quaternary Science Reviews 270:107175. https://doi.org/10.1016/j.quascirev.2021.107175

van Vugt L, Garcés-Pastor S, Gobet E, Brechbühl S, Knetge A, Lammers Y, Stengele K, Alsos IG, Tinner W, Schwörer C (2022) Pollen, macrofossils and sedaDNA reveal climate and land use impacts on Holocene mountain vegetation of the Lepontine Alps, Italy. Quaternary Science Reviews 296:107749. https://doi.org/10.1016/j.quascirev.2022.107749

Vernot B, Zavala EI, Gómez-Olivencia A, Jacobs Z, Slon V, Mafessoni F, Romagné F, Pearson A, Petr M, Sala N, Pablos A, Aranbur A, De Castro JMB, Carbonell E, Li B, Krajcarz MT, Krivoshapkin AI,

Kolobova KA, Kozlikin MB, Shunkov MV, Derevianko AP, Viola B, Grote S, Essel E, Herraéz DL, Nagel S, Nickel B, Richter J, Schmidt A, Peter B, Kelso J, Roberts RG, Arsuaga JL, Meyer M (2021) Unearthing Neanderthal population history using nuclear and mitochondrial DNA from cave sediments. Science 372(6542). https://doi.org/10.1126/science.abf1667

Von Eggers J, Monchamp M-E, Capo E, Giguet-Covex C, Spanbauer T, Heintzman PD (2022) Inventory of ancient environmental DNA from sedimentary archives: locations, methods, and target taxa. Zenodo. https://doi.org/10.5281/ZENODO.6847522

Wang Y, Pedersen MW, Alsos IG, De Sanctis B, Racimo F, Prohaska A, Coissac E, Owens HL, Merkel MKF, Fernandez-Guerra A, Rouillard A, Lammers Y, Alberti A, Denoeud F, Money D, Ruter AH, McColl H, Larsen NK, Cherezova AA, Edwards ME, Fedorov GB, Haile J, Orlando L, Vinner L, Korneliussen TS, Beilman DW, Bjørk AA, Cao J, Dockter C, Esdale J, Gusarova G, Kjeldsen KK, Mangerud J, Rasic JT, Skadhauge B, Svendsen JI, Tikhonov A, Wincker P, Xing Y, Zhang Y, Froese DG, Rahbek C, Bravo DN, Holden PB, Edwards NR, Durbin R, Meltzer DJ, Kjær KH, Möller P, Willerslev E (2021) Late Quaternary dynamics of Arctic biota from ancient environmental genomics. Nature 600(7887):86–92. https://doi.org/10.1038/s41586-021-04016-x

Wang Y, Korneliussen TS, Holman LE, Manica A, Pedersen MW (2022) ngsLCA—A toolkit for fast and flexible lowest common ancestor inference and taxonomic profiling of metagenomic data. Methods in Ecology and Evolution 13(12):2699–2708. https://doi.org/10.1111/2041-210X.14006

Warinner C, Rodrigues JFM, Vyas R, Trachsel C, Shved N, Grossmann J, Radini A, Hancock Y, Tito RY, Fiddyment 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, Collins MJ, Gilbert MTP, Rühli F, Cappellini E (2014) Pathogens and host immunity in the ancient human oral cavity. Nat Genet 46(4):336–344. https://doi.org/10.1038/ng.2906

Weber S, Junk I, Brink L, Wörner M, Künzel S, Veith M, Teubner D, Klein R, Paulus M, Krehenwinkel H (2023) Molecular diet analysis in mussels and other metazoan filter feeders and an assessment of their utility as natural eDNA samplers. Molecular Ecology Resources 23(2):471–485. https://doi.org/10.1111/1755-0998.13710

Weiß CL, Dannemann M, Prüfer K, Burbano HA (2015) Contesting the presence of wheat in the British Isles 8,000 years ago by assessing ancient DNA authenticity from low-coverage data. Elife 4:e10005. https://doi.org/10.7554/eLife.10005

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, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3(1):160018. https://doi.org/10.1038/sdata.2016.18

Willerslev E, Hansen AJ, Binladen J, Brand TB, Gilbert MTP, Shapiro B, Bunce M, Wiuf C, Gilichinsky DA, Cooper A (2003) Diverse plant and animal genetic records from Holocene and Pleistocene sediments. Science 300(5620):791–795. https://doi.org/10.1126/science.1084114

Williams JW, Spanbauer TL, Heintzman PD, Blois J, Capo E, Goring SJ, Monchamp M-E, Parducci L, Eggers JMV, Alsos IG, Bowler C, Coolen MJL, Cullen N, Crump S, Epp LS, Fernandez-Guerra A, Grimm E, Herzschuh U, Mereghetti A, Meyer RS, Nota K, Pedersen MW, Pérez V, Shapiro B, Stoof-Leichsenring KR, Wood J (2023) Strengthening global-change science by integrating aeDNA with paleoecoinformatics. Trends in Ecology & Evolution 0(0). https://doi.org/10.1016/j.tree.2023.04.016

Wood DE, Lu J, Langmead B (2019) Improved metagenomic analysis with Kraken 2. Genome Biology 20(1):257. https://doi.org/10.1186/s13059-019-1891-0

Zavala EI, Jacobs Z, Vernot B, Shunkov MV, Kozlikin MB, Derevianko AP, Essel E, de Fillipo 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(7867):399–403. https://doi.org/10.1038/s41586-021-03675-0

Ziesemer KA, Mann AE, Sankaranarayanan K, Schroeder H, Ozga AT, Brandt BW, Zaura E, Waters-Rist A, Hoogland M, Salazar-García DC, Aldenderfer M, Speller C, Hendy J, Weston DA, MacDonald SJ, Thomas GH, Collins MJ, Lewis CM, Hofman C, Warinner C (2015) Intrinsic challenges in ancient microbiome reconstruction using 16S rRNA gene amplification. Sci Rep 5:16498. https://doi.org/10.1038/srep16498

Zinger L, Bonin A, Alsos IG, Bálint M, Bik H, Boyer F, Chariton AA, Creer S, Coissac E, Deagle BE, De Barba M, Dickie IA, Dumbrell AJ, Ficetola GF, Fierer N, Fumagalli L, Gilbert MTP, Jarman S, Jumpponen A, Kauserud H, Orlando L, Pansu J, Pawlowski J, Tedersoo L, Thomsen PF, Willerslev E, Taberlet P (2019) DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. Molecular Ecology 28(8):1857–1862. https://doi.org/10.1111/mec.15060