Ancient environmental genomics: An Introduction

Ngoc-Loi Nguyen[1]

[1]Department of Paleoceanography, Institute of Oceanology Polish Academy of Sciences

February 22, 2024

Abstract

Environmental DNA (*e*DNA) obtained from ancient samples such as sediments, ice or water are valuable data sources for a wide range of disciplines in past and present biodiversity and biogeography [1-4]. Within the field of ancient metagenomics, the number of published genetic datasets has risen dramatically in recent years and have become an increasingly powerful tool to investigate wide-ranging topics [5]. However, the ancient environmental metagenomics remains many issues that should be to be addressed relating to ancient DNA (*a*DNA) such as degraded nature, incomplete reference databases, sensitivity to contamination by modern DNA [6-8]. This review aims to provide an overview of the use of ancient metagenomics in large-scale ecological and evolutionary studies of individual taxa and communities of both microbes and eukaryotes and illustrate the limitations, risks, and potentiality of this ancient *e*DNA research via high-throughput sequencing (HTS) technologies. Further, paleogenetic and paleogenomics will provide diverse insights into studying evolution and how the present world came to be.

I. Introduction

Environmental DNA (*e*DNA) obtained from ancient samples such as sediments, ice or water are valuable data sources for a wide range of disciplines in past and present biodiversity and biogeography [1-4]. Within the field of ancient metagenomics, the number of published genetic datasets has risen dramatically in recent years and have become an increasingly powerful tool to investigate wide-ranging topics [5]. However, the ancient environmental metagenomics remains many issues that should be to be addressed relating to ancient DNA (*a*DNA) such as degraded nature, incomplete reference databases, sensitivity to contamination by modern DNA [6-8]. This review aims to provide an overview of the use of ancient metagenomics in large-scale ecological and evolutionary studies of individual taxa and communities of both microbes and eukaryotes and illustrate the limitations, risks, and potentiality of this ancient *e*DNA research via high-throughput sequencing (HTS) technologies. Further, paleogenetic and paleogenomics will provide diverse insights into studying evolution and how the present world came to be.

II. Ancient eDNA and ancient environmental metagenomics

In general, *e*DNA was extracted from ancient samples extremely fragmented and chemically modified depending on the sample types [6]. Typically, the size of ancient *e*DNA fragments is from 70 base pairs (bp) to less than 100 bp long [9] and with ends impacted by cytosine deamination [10]. Only in a few cases, where extraordinary preservation such as Antarctic conditions, for example, 500 bp of *a*DNA were recovered from lake sediment [11], respectively. These conditions generally feature anoxic, cold and dry conditions [6]. In the context of isolating *a*DNA from environmental samples, environmental *a*DNA including sedimentary ancient DNA (or *seda*DNA) is used widely and applies to DNA isolated from sedimentary deposits in lake cores [12-14], marine [15, 16], cave [17-19], ancient forest [20], permafrost [13, 21-23], peat [24], tropical swamp [25]. However, there is potential for many other materials to provide information about the past via *a*DNA analysis as basal ice [20], glacial soil [26], silt-soaked [27]. Analysis of *a*DNA datasets, when

combined with traditional proxy results, appears to complement each other, revealing a greater diversity of species than utilizing the methodologies independently [15, 28, 29]. Therefore, *a*DNA should be considered as a complementary, rather than alternative, approach to assays of more traditional established methods [3, 30].

The metagenomics of ancient environmental DNA can be broadly defined as the study of the total genetic content of samples that have degraded over time from several hundred to hundred-thousand years [5, 31]. Despite an extensive application including studies of genome reconstruction of specific microbial taxa [12, 32], host-associated microbial communities [33, 34], and environmental reconstructions using *sed*aDNA [5, 25, 35], the major source of ancient eDNA has been almost entirely limited to inventorying taxa through time by using DNA metabarcoding approach [15, 16, 36, 37]. Recent advances of next-generation sequencing (NGS), massively parallel or deep sequencing technology, have the potential to radically change this situation, from sequencing of millions of short DNA fragments to generating datasets of genome-scale from extant and extinct species by bioinformatics analyses [12, 13, 32, 37].

III. The problem of environmental ancient DNA

Despite recent methodological strategies for *a*DNA extraction, Polymerase Chain Reaction (PCR) and/or sequencing, the study of *a*DNA could be negatively affected by the applicability and the outcome by several inherent technical issues. Part of the challenge is the fact that ancient samples are often rare and precious materials, such as low DNA quantities, DNA damage, high fragmentation, and contamination with modern sources [6]. In general, the ancient *e*DNA sample processing and analysis should be processed with practical recommendations for ancient DNA research to prevent contamination, reviewed in Capo et al., 2021 [35] for lake sediment cores and Armbrecht et al., 2019 [8] for marine sediment cores.

The current *a*DNA extraction protocols were not very different from the protocols used to obtain DNA from environmental settings including silica-based, alcoholic, and phenol-chloroform protocols [22, 38, 39]. For the molecular analyses, the yield and integrity of the recovered *a*DNA obtained will influence the reliability of subsequent results. Therefore, extraction protocols of *a*DNA should be carefully considered and adapted depending on the physical and chemical properties of sediments, DNA-subtracts interaction, or target organisms [8, 15, 40, 41]. Further, quick, simple and direct DNA extraction procedures are needed for use in regular analysis of *a*DNA.

DNA damage alters the base-pairing properties of individual bases and is vastly over-represented in *a*DNA sequences. This increased rate of polymerase misincorporation errors and therefore sequencing errors by incorporating wrong nucleotides opposite modified bases [42, 43]. During PCR, DNA damages cause blocking primer binding/DNA polymerase progression, preventing the amplification of the templates, or hydrolysis of the phosphodiester bond, resulting in a single-strand break [44-46]. For instance, the majority of errors give by deamination of cytosine to uracil, which pairs up with adenine instead of guanine, leading to thymine to cytosine transitions [45-47]. However, well-characterized degradation features of *a*DNA i.e., damage patterns and high fragmentation, allow us to authenticate 'true' *a*DNA sequences.

IV. How to study ancient metagenomic

The application of several technologies, from PCR and the earlier methods, including Sanger sequencing, to HTS, also known as Next-Generation Sequencing (NGS) [48] for short-read (shotgun) sequencing [49] or long-read sequencing, dramatically started a new revolution in ancient DNA research **(Figure 1).** While

traditional PCR methods could only amplify a small number of specific target sequences, HTS combines amplification and sequencing of up to several billions of individual DNA library templates at a time. DNA/RNA metabarcoding approach is an extension of DNA barcoding, which relies on HTS technologies [36, 50-53]. Furthermore, HTS can sequence shorter DNA fragments - shotgun [37] and event recover whole genome sequences for the study of paleogenomics [12, 54, 55]. These technologies generate large quantities of highly accurate DNA sequences at lower costs than it was possible by using first-generation sequencing technologies.

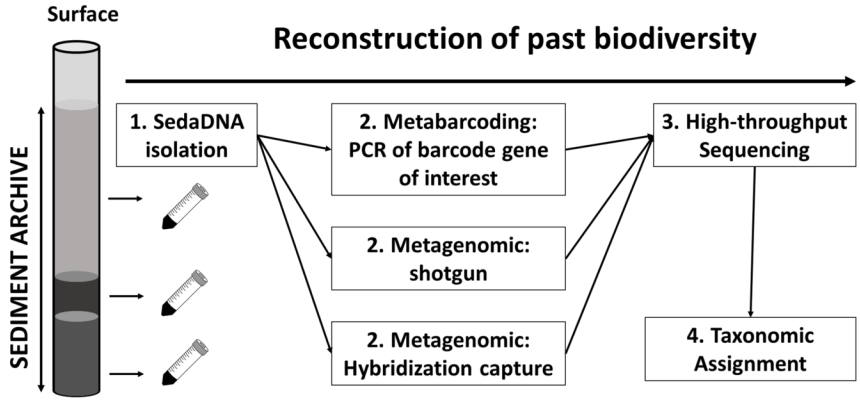

Figure 1. Conceptual workflow of ancient metagenomic approach applied to DNA preserved in environmental archives (e.g., marine and freshwater sediment cores) to reconstruct the past diversity.

In brief in **Figure 1**, two main approaches to the study of *a*DNA are metabarcoding, the taxonomic identification of the community via analysis of short DNA sequences of one or a few genes, and metagenomics, the analysis of total DNA of the community via whole-genome sequencing. For workflow of the wet laboratory, total DNA is initially isolated from the sample, for example, sediment cores. Next, the DNA metabarcoding standard steps include PCR amplification, library preparation, and sequencing followed by bioinformatic analyses. Depending on the targeted organisms, the specific primers are used to amplify DNA fragments, e.g., the mitochondrial COI region [56], foraminiferal 37f hypervariable region [57-59], and the internal transcribed spacer (ITS) region [60]. For distinguishing samples during bioinformatic processing, specific tags or indexes are added using ligation or other PCR-round. After quantification and normalization steps, the final library is then sequenced on one of the various available sequencing platforms, e.g., Illumina, Ion Torrent, PacBio, or Oxford Nanopore. In contrast, after collecting suitable samples under the guideline of *a*DNA research, the wet lab workflow for (shotgun) metagenomics can be roughly divided into three steps: DNA extraction, library preparation, and sequencing, without PCR.

4.1. Metabarcoding and its limitations

To date, most paleoecological *a*DNA investigations have employed the widely used DNA metabarcoding method, usually, with a focus on a particular organismal group [61]. DNA metabarcoding represents a molecular approach to contemporary taxonomy and identification, e.g., plant [50, 62-65], fungi [60], foraminifera [57, 58], metazoan [56, 66, 67]. The PCR-metabarcoding approach uses primer pairs to target and maximize portions of the hypervariable regions of the phylogenetic marker genes. Amplicons from separate samples are then given molecular barcodes, pooled together, and sequenced by amplicon-based HTS approaches. Fragments of *a*DNA are analyzed with a bioinformatics pipeline and identified from environmental archives, by comparison, them against sequences of reference database taken from modern reference organisms [29, 36, 68, 69].

However, metabarcoding which is applied to environmental *a*DNA is complicated by its natural degradation. The PCR-based approach for sequencing can generate incorrect sequence data from *a*DNA for several reasons. The total amplified sequence count is likely to reflect the original abundance of different DNA sequences in the sample. Damages of *a*DNA could inhibit DNA polymerase progression or prevent primers from binding to templates during PCR. The *a*DNA fragments are extremely short and low-yields, while preferential random amplification is longer or requires abundant DNA molecules. As a result, a lot of PCR cycles are needed, and false-positive findings are more frequent, and heavily biased towards well-preserved or more abundant sequences, possibly from present-day DNA contamination during the first few cycles [37, 70]. It can be induced predictably biased in multi-template PCR and significantly distort the final output. To solve this problem, PCRs can be repeated independently and increase the total number of replicates for each sample as well as using negative controls should be applied [71]. This approach makes short and rare sequences more likely to be identified than if only one replicate were used since they are likely to be missed in a single PCR but should be expected in one or more of the repeat PCRs. Further, based on using genetic markers in molecular studies of previous paleo-microbiome research, the length of taxonomic marker genes is a major cause of differential amplification resulting in a taxonomic bias in ancient reconstructions [72].

4.2. Shotgun sequencing and Whole Genome Sequencing

Shotgun sequencing is the untargeted (shotgun) sequencing of all genetic material (metagenomics) present in a sample, which has the potential to look for population genomic variation from multi-taxon mixtures and independent of DNA fragment size [36, 72]. Compared to metabarcoding, the shotgun approach is less subject to bias introduced by laboratory processing, ever-reducing sequencing costs. Generally, shotgun sequencing randomly breaks DNA sequences of the entire chromosome or entire genome into many small fragments and reassembles the sequences by computers via observing the overlapping sequences or regions. The shotgun approach can detect this genomic variation of the population by utilizing extensive intraspecific genomic reference datasets [73, 74] or assembling *de novo* genomes [75, 76]. Furthermore, the whole-genome shotgun (WGS) method entails sequencing many overlapping DNA fragments in parallel and then using a computer to assemble the small fragments into larger contigs and, eventually, chromosomes within a short period. NGS has also been used to obtain RNA and pathogen genome sequences from ancient plant remains [77]. The adoption of NGS technologies significantly expanded the range of *a*DNA studies possible, enabling the analysis of full chloroplast [54, 78], and mitochondrial and nuclear genomes [79, 80] from ancient samples. For instance, chloroplast and mitochondrial genomes of single-celled microalgae (*Nannochloropsis limnetica*) were successfully reconstructed from 20 000-year-old lake sediments [12].

Shotgun sequencing is a faster method and cheaper to carry out compared with traditional sequencing. Usefully, the advent of the shotgun approach permits statistical data analyses to detect specific substitutions that are normally present at the ends of ancient DNA fragments, therefore confirming whether a sequence or set of sequences is relatively ancient and not modern contamination, as well as improving the specificity and sensitivity of taxonomic identification [81, 82]. In some cases, as for eukaryotes in *sed*aDNA, if the targeted DNA is rare compared to the total genomic DNA, producing large numbers of short sequencing reads [83] is required to recover sufficient genetic information and perform meaningful statistical analyses, particularly useful for *a*DNA analysis for its fragmentation and degradation [84]. Usefully, the ends of older sequences retrieved using a shotgun approach will show deamination damage, which can confirm whether a sequence or set of sequences is relatively ancient and not modern contamination. Although whole ancient genomes are becoming more readily accessible, mitochondrial [13, 85, 86] or chloroplast [12, 54, 78] genomes are an alternative choice in *a*DNA studies dealing with samples with high DNA degradation, and low DNA yields. Before sequencing, another alternative option applies the hybridization capture technique [78, 87]. The constraint of shotgun sequencing might be solved by using the hybridization capture approach before sequencing to enrich the DNA of the targeted species in the samples. To do this, small segments of DNA from the species and target sites of interest can be used as baits, with the matching sites of interest in ancient

DNA libraries being hybridized. This technique, originally developed for modern DNA, is commonly applied in ancient DNA studies, particularly for use on single specimens [88] and with a focus on mammals, mostly using mitochondrial DNA [89, 90], chloroplast and nuclear DNA [78, 91-93], cave sediments [19], permafrost samples [22].

4.3. Bioinformatics considerations

Now the shotgun approach provides an alternative approach to metabarcoding for determining for taxonomic and functional profiling of metagenome-assembled genomes. The amount of genetic data has risen exponentially and vast amounts of that are mostly uploaded to and stored on public archives, for example, European Bioinformatic Institute's (EBI) European Nucleotide Archive (ENA, https://www.ebi.ac.uk/ena/) or the US National Center for Biotechnology Information (NCBI)'s Sequence Read Archive (SRA, https://www.ncbi.nlm.nih.gov/sra). However, it brings huge challenges at the stage of bioinformatics for its analysis. A vast of bioinformatics tools, protocols and studies have been introduced to improve efficiency in analyzing ancient metagenomic data. Bioinformatics tools designed for *a*DNA metagenomics as mapDamage [94-96], PyDamage [97] or open-sourced/mapping guidelines pipeline [98, 99] for estimating DNA damage, SourceTracker [100] for identifying the proportions of endogenous and contaminant signals in each sample; resolving the sequencing errors [96, 101]; MEGAN [102, 103], PIA [104] for taxonomic identification; KEGG [105], EGGnog [106], SEED [107] protein databases for functional profiles can be analysed in MEGAN, reference-free alternative approaches based on k-mer counts [108] to annotate metagenomes. However, differences between metagenomic analysis pipelines produce systematic biases [25], which will require the development of more accurate analysis pipelines for ancient DNA.

Nevertheless, several issues currently limit the shotgun sequencing approach. Cytosine deamination patterns of *seda*DNA molecules impede *de novo* assembly of contigs [10, 109]. The limitation of sufficiently curated genome-scale reference data substantially reduces the potential for success of the bioinformatic analyses with metagenomic data, for example, plants [77, 110], and eukaryotic [111, 112]. The large fraction of taxa present in the environment, but not represented in databases is still problematic. In these cases, metagenomic data can vary in content across samples from the same or similar environments. In contrast, there are more than 130,000 genome or near-complete sequences available from different phyla that have been sequenced along with a variety of microorganisms, including archaea, fungi, and viruses [113-115]. Based on the annotated reference genomes or clade-specific [116] or universal markers [117], appropriate normalization by genome size [55], and taxon relative abundances can be estimated. This led to the development of the field of paleomicrobiology [1, 32], to the analysis of deposited microbial DNA to study microbial diversity, ecology, and evolution in environmental archives.

4.4. Applications of ancient environmental metagenomics

The shotgun of *seda*DNA in paleoecology from lake sediment cores combined a multi-proxy approach [14], and marine environments [37, 40], which has provided greater taxonomic resolution and extended the historical record of aquatic ecosystems to centennial or even millennial time scales. These *seda*DNA archives can be used to characterize biodiversity trends, illuminate past food web dynamics, and reconstruct long-term environmental changes in aquatic ecosystems. As ecology and paleoecology merge, both short-term and long-term trends as a consequence of human actions on aquatic ecosystems have been traced using paleogenomic research in freshwater ecosystems [118-120] and marine sediments [121, 122].

Paleogenomics is a branch of research concerned with reconstructing and analyzing genetic data from extinct organisms. Ancient genomes may be used to explore the evolution of present species in great detail by

sequencing ancient DNA preserved in subfossil remains [54, 123] or environmental archives [1, 12]. By analyzing large-scale environmental DNA metagenomic study of ancient plant and mammal communities, tracking the ancient population origins, movements and interrelationships, the evolutionary genomic changes at both macro- and micro-evolutionary temporal scales of the microbiome, vegetation, animals and *Homo species* [12, 13], as well as identification of phenotypic features over large temporal and geographical scales [89, 90, 124]. For example, a study on DNA retrieved from Arctic permafrost and lake sediment samples by Wang et al. [13] demonstrated that steppe–tundra flora dominated the Arctic during the Last Glacial Maximum, followed by the regional divergence of vegetation during the Holocene epoch. The extinction of several now-extinct megafauna species enabled the survival of some ancient plants and animals. Moreover, analysis of mammoth environmental DNA reveals a previously unsampled mitochondrial lineage. Additionally, the genetic material preserved in sedimentary archives offers a unique way to uncover the role of microorganisms in past ecosystems and their responses to environmental perturbations. Genomic reconstruction of historical and present microbial communities from ancient permafrost samples in Siberian broadened our understanding of biogeochemical changes [32]. Furthermore, this study provides insights into microorganisms' long-term survival strategies from the past paleoenvironment to present-day freezing-temperature conditions.

V. Summary

In conclusion, the fields of *a*DNA are increasingly turning to the environmental archives and provide great potential for entire paleoecosystems and paleoclimate reconstructions. As technology advances and procedures are optimized, metagenomic-based approaches, from metabarcoding (amplicon-based) to shotgun and true ancient metagenomics, are part of the next breakthrough in paleogenetic, offering the potential for better species identification and quantitative estimations of their abundances in large-scale biodiversity comparisons over both time and place. Importantly, further basic studies are needed to use a full understanding of its potential and limitations for applications of the use of metagenomics for ancient eDNA.

Acknowledgement

The research was financially supported by the Norwegian Financial Mechanism for 2014-2021, project no 2019/34/H/ST10/00682, full title: "Sedimentary ancient DNA - a new proxy to investigate the impact of environmental change on past and present biodiversity in Nordic Seas".

Reference

[1] Capo, E., et al., Environmental paleomicrobiology: using DNA preserved in aquatic sediments to its full potential. Environmental Microbiology, 2022.

[2] Pawlowski, J., et al., Environmental DNA for biomonitoring. Molecular Ecology, 2021. **30**(13): p. 2931-2936.

[3] Ruppert, K.M., R.J. Kline, and M.S. Rahman, Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Global Ecology and Conservation, 2019. **17**: p. e00547.

[4] Taberlet, P., et al., Environmental DNA: For Biodiversity Research and Monitoring. 2018.

[5] Fellows Yates, J.A., et al., Community-curated and standardised metadata of published ancient metagenomic samples with AncientMetagenomeDir. Scientific Data, 2021. **8**(1): p. 31.

[6] Orlando, L., et al., Ancient DNA analysis. Nature Reviews Methods Primers, 2021. **1**(1): p. 14.

[7] Pedersen, M.W., et al., Ancient and modern environmental DNA. Philos Trans R Soc Lond B Biol Sci, 2015. **370**(1660): p. 20130383.

[8] Armbrecht, L.H., et al., Ancient DNA from marine sediments: Precautions and considerations for seafloor coring, sample handling and data generation. Earth-Science Reviews, 2019. **196**: p. 102887.

[9] Armbrecht, L., et al., An optimized method for the extraction of ancient eukaryote DNA from marine sediments. Molecular Ecology Resources, 2020. **20**(4): p. 906-919.

[10] Orlando, L., M.T.P. Gilbert, and E. Willerslev, Reconstructing ancient genomes and epigenomes. Nature Reviews Genetics, 2015. **16**(7): p. 395-408.

[11] Coolen, M.J.L., et al., Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake. Earth and Planetary Science Letters, 2004. **223**(1): p. 225-239.

[12] Lammers, Y., P.D. Heintzman, and I.G. Alsos, Environmental palaeogenomic reconstruction of an Ice Age algal population. Communications Biology, 2021. **4**(1): p. 220.

[13] Wang, Y., et al., Late Quaternary dynamics of Arctic biota from ancient environmental genomics. Nature, 2021. **600**(7887): p. 86-92.

[14] Pedersen, M.W., et al., Postglacial viability and colonization in North America's ice-free corridor. Nature, 2016. **537**(7618): p. 45-49.

[15] De Schepper, S., et al., The potential of sedimentary ancient DNA for reconstructing past sea ice evolution. ISME J, 2019. **13**(10): p. 2566-2577.

[16] Zimmermann, H.H., et al., Sedimentary Ancient DNA From the Subarctic North Pacific: How Sea Ice, Salinity, and Insolation Dynamics Have Shaped Diatom Composition and Richness Over the Past 20,000 Years. Paleoceanography and Paleoclimatology, 2021. **36**(4): p. e2020PA004091.

[17] Haile, J., et al., Ancient DNA Chronology within Sediment Deposits: Are Paleobiological Reconstructions Possible and Is DNA Leaching a Factor? Molecular Biology and Evolution, 2007. **24**(4): p. 982-989.

[18] Haouchar, D., et al., Thorough assessment of DNA preservation from fossil bone and sediments excavated from a late Pleistocene–Holocene cave deposit on Kangaroo Island, South Australia. Quaternary Science Reviews, 2014. **84**: p. 56-64.

[19] Slon, V., et al., Neandertal and Denisovan DNA from Pleistocene sediments. Science, 2017. **356**(6338): p. 605-608.

[20] Willerslev, E., et al., Ancient Biomolecules from Deep Ice Cores Reveal a Forested Southern Greenland. Science, 2007. **317**(5834): p. 111-114.

[21] Willerslev, E., et al., Diverse Plant and Animal Genetic Records from Holocene and Pleistocene Sediments. Science, 2003. **300**(5620): p. 791-795.

[22] Murchie, T.J., et al., Optimizing extraction and targeted capture of ancient environmental DNA for reconstructing past environments using the PalaeoChip Arctic-1.0 bait-set. Quaternary Research, 2020. **99**: p. 305-328.

[23] Willerslev, E., et al., Fifty thousand years of Arctic vegetation and megafaunal diet. Nature, 2014. **506**(7486): p. 47-51.

[24] Suyama, Y., U. Gunnarsson, and L. Parducci, Analysis of short DNA fragments from Holocene peatmoss samples. The Holocene, 2008. **18**(6): p. 1003-1006.

[25] Dommain, R., et al., The Challenges of Reconstructing Tropical Biodiversity With Sedimentary Ancient DNA: A 2200-Year-Long Metagenomic Record From Bwindi Impenetrable Forest, Uganda. Frontiers in Ecology and Evolution, 2020. **8**.

[26] Gould, B.A., et al., Evidence of a high-Andean, mid-Holocene plant community: An ancient DNA analysis of glacially preserved remains. American Journal of Botany, 2010. **97**(9): p. 1579-1584.

[27] Jørgensen, T., et al., A comparative study of ancient sedimentary DNA, pollen and macrofossils from permafrost sediments of northern Siberia reveals long-term vegetational stability. Molecular Ecology, 2012. **21**(8): p. 1989-2003.

[28] Pawlowska, J., et al., Ancient DNA sheds new light on the Svalbard foraminiferal fossil record of the last millennium. Geobiology, 2014. **12**(4): p. 277-88.

[29] Edwards, M.E., et al., Metabarcoding of modern soil DNA gives a highly local vegetation signal in Svalbard tundra. The Holocene, 2018. **28**(12): p. 2006-2016.

[30] Pedersen, M.W., et al., A comparative study of ancient environmental DNA to pollen and macrofossils from lake sediments reveals taxonomic overlap and additional plant taxa. Quaternary Science Reviews, 2013. **75**: p. 161-168.

[31] Warinner, C., et al., A Robust Framework for Microbial Archaeology. Annual Review of Genomics and Human Genetics, 2017. **18**(1): p. 321-356.

[32] Liang, R., et al., Genomic reconstruction of fossil and living microorganisms in ancient Siberian permafrost. Microbiome, 2021. **9**(1): p. 110.

[33] Warinner, C., et al., Ancient human microbiomes. Journal of Human Evolution, 2015. **79**: p. 125-136.

[34] Wibowo, M.C., et al., Reconstruction of ancient microbial genomes from the human gut. Nature, 2021. **594**(7862): p. 234-239.

[35] Capo, E., et al., Lake Sedimentary DNA Research on Past Terrestrial and Aquatic Biodiversity: Overview and Recommendations. Quaternary, 2021. **4**(1): p. 6.

[36] Taberlet, P., et al., Towards next-generation biodiversity assessment using DNA metabarcoding. Mol Ecol, 2012. **21**(8): p. 2045-50.

[37] Armbrecht, L., et al., Paleo-diatom composition from Santa Barbara Basin deep-sea sediments: a comparison of 18S-V9 and diat-rbcL metabarcoding vs shotgun metagenomics. ISME Communications, 2021. **1**(1): p. 66.

[38] Hagan, R.W., et al., Comparison of extraction methods for recovering ancient microbial DNA from paleofeces. American Journal of Physical Anthropology, 2020. **171**(2): p. 275-284.

[39] Pearman, J.K., et al., Comparing sediment DNA extraction methods for assessing organic enrichment associated with marine aquaculture. PeerJ, 2020. **8**: p. e10231.

[40] Orsi, W.D., et al., Climate oscillations reflected within the microbiome of Arabian Sea sediments. Sci Rep, 2017. **7**(1): p. 6040.

[41] Coolen, M.J., et al., Evolution of the plankton paleome in the Black Sea from the Deglacial to Anthropocene. Proc Natl Acad Sci U S A, 2013. **110**(21): p. 8609-14.

[42] Hofreiter, M., et al., DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Research, 2001. **29**(23): p. 4793-4799.

[43] Pääbo, S., et al., Genetic analyses from ancient DNA. Annu. Rev. Genet., 2004. **38**: p. 645-679.

[44] Torti, A., M.A. Lever, and B.B. Jørgensen, Origin, dynamics, and implications of extracellular DNA pools in marine sediments. Marine Genomics, 2015. **24**: p. 185-196.

[45] Briggs, A.W., et al., Patterns of damage in genomic DNA sequences from a Neandertal. Proceedings of the National Academy of Sciences, 2007. **104**(37): p. 14616-14621.

[46] Dabney, J., M. Meyer, and S. Pääbo, Ancient DNA Damage. Cold Spring Harb Perspect Biol 5: a012567. 2013.

[47] Gansauge, M.-T. and M. Meyer, Selective enrichment of damaged DNA molecules for ancient genome sequencing. Genome research, 2014. **24**(9): p. 1543-1549.

[48] Shendure, J. and H. Ji, Next-generation DNA sequencing. Nature Biotechnology, 2008. **26**(10): p. 1135-1145.

[49] Quince, C., et al., Shotgun metagenomics, from sampling to analysis. Nature Biotechnology, 2017. **35**(9): p. 833-844.

[50] Dormontt, E.E., et al., Advancing DNA Barcoding and Metabarcoding Applications for Plants Requires Systematic Analysis of Herbarium Collections—An Australian Perspective. Frontiers in Ecology and Evolution, 2018. **6**.

[51] Lejzerowicz, F., et al., Eukaryotic Biodiversity and Spatial Patterns in the Clarion-Clipperton Zone and Other Abyssal Regions: Insights From Sediment DNA and RNA Metabarcoding. Frontiers in Marine Science, 2021. **8**(536).

[52] Hirai, J., et al., DNA/RNA metabarcoding and morphological analysis of epipelagic copepod communities in the Izu Ridge off the southern coast of Japan. ICES Journal of Marine Science, 2021. **78**(9): p. 3444-3456.

[53] Laroche, O., et al., A cross-taxa study using environmental DNA/RNA metabarcoding to measure biological impacts of offshore oil and gas drilling and production operations. Mar Pollut Bull, 2018. **127**: p. 97-107.

[54] Pont, C., et al., Paleogenomics: reconstruction of plant evolutionary trajectories from modern and ancient DNA. Genome Biology, 2019. **20**(1): p. 29.

[55] Schubert, M., et al., Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. Nature Protocols, 2014. **9**(5): p. 1056-1082.

[56] Leray, M., et al., A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Front Zool, 2013. **10**: p. 34.

[57] Lecroq, B., et al., Ultra-deep sequencing of foraminiferal microbarcodes unveils hidden richness of early monothalamous lineages in deep-sea sediments. Proc Natl Acad Sci U S A, 2011. **108**(32): p. 13177-82.

[58] Pawlowski, J. and B. Lecroq, Short rDNA barcodes for species identification in foraminifera. J Eukaryot Microbiol, 2010. **57**(2): p. 197-205.

[59] Pawłowska, J., et al., Ancient DNA sheds new light on the Svalbard foraminiferal fossil record of the last millennium. Geobiology, 2014. **12**(4): p. 277-288.

[60] Talas, L., et al., Sedimentary Ancient DNA (sedaDNA) Reveals Fungal Diversity and Environmental Drivers of Community Changes throughout the Holocene in the Present Boreal Lake Lielais Svētiņu (Eastern Latvia). Microorganisms, 2021. **9**(4): p. 719.

[61] Ratnasingham, S. and P.D.N. Hebert, The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes, 2007. **7**(3): p. 355-364.

[62] Taberlet, P., et al., Power and limitations of the chloroplast trn L (UAA) intron for plant DNA barcoding. Nucleic Acids Research, 2006. **35**(3): p. e14-e14.

[63] Deiner, K., et al., Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Mol Ecol, 2017. **26**(21): p. 5872-5895.

[64] Voldstad, L.H., et al., A complete Holocene lake sediment ancient DNA record reveals long-standing high Arctic plant diversity hotspot in northern Svalbard. Quaternary Science Reviews, 2020. **234**: p. 106207.

[65] Huang, S., et al., Plant Sedimentary Ancient DNA From Far East Russia Covering the Last 28,000 Years Reveals Different Assembly Rules in Cold and Warm Climates. Frontiers in Ecology and Evolution, 2021. **9**.

[66] Di Capua, I., et al., Metazoan diversity and seasonality through eDNA metabarcoding at a Mediterranean long-term ecological research site. ICES Journal of Marine Science, 2021. **78**(9): p. 3303-3316.

[67] Leduc, N., et al., Comparing eDNA metabarcoding and species collection for documenting Arctic metazoan biodiversity. Environmental DNA, 2019. **1**(4): p. 342-358.

[68] Cao, X., et al., Sedimentary ancient DNA metabarcoding delineates the contrastingly temporal change of lake cyanobacterial communities. Water Research, 2020. **183**: p. 116077.

[69] Epp, L.S., et al., New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. Mol Ecol, 2012. **21**(8): p. 1821-33.

[70] Webster, G., et al., Assessment of bacterial community structure in the deep sub-seafloor biosphere by 16S rDNA-based techniques: a cautionary tale. Journal of Microbiological Methods, 2003. **55**(1): p. 155-164.

[71] Ficetola, G.F., et al., Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. Molecular Ecology Resources, 2015. **15**(3): p. 543-556.

[72] Ziesemer, K.A., et al., Intrinsic challenges in ancient microbiome reconstruction using 16S rRNA gene amplification. Scientific Reports, 2015. **5**(1): p. 16498.

[73] Truong, D.T., et al., Microbial strain-level population structure and genetic diversity from metagenomes. Genome Research, 2017. **27**(4): p. 626-638.

[74] Albanese, D. and C. Donati, Strain profiling and epidemiology of bacterial species from metagenomic sequencing. Nature Communications, 2017. **8**(1): p. 2260.

[75] Quince, C., et al., DESMAN: a new tool for de novo extraction of strains from metagenomes. Genome Biology, 2017. **18**(1): p. 181.

[76] Nurk, S., et al., metaSPAdes: a new versatile metagenomic assembler. Genome research, 2017. **27**(5): p. 824-834.

[77] Estrada, O., et al., Ancient plant DNA in the genomic era. Nature Plants, 2018. **4**(7): p. 394-396.

[78] Schulte, L., et al., Hybridization capture of larch (Larix Mill.) chloroplast genomes from sedimentary ancient DNA reveals past changes of Siberian forest. Molecular Ecology Resources, 2021. **21**(3): p. 801-815.

[79] Diroma, M.A., et al., New Insights Into Mitochondrial DNA Reconstruction and Variant Detection in Ancient Samples. Frontiers in Genetics, 2021. **12**.

[80] Vernot, B., et al., Unearthing Neanderthal population history using nuclear and mitochondrial DNA from cave sediments. Science, 2021. **372**(6542): p. eabf1667.

[81] Overballe-Petersen, S., L. Orlando, and E. Willerslev, Next-generation sequencing offers new insights into DNA degradation. Trends in Biotechnology, 2012. **30**(7): p. 364-368.

[82] Binladen, J., et al., Assessing the Fidelity of Ancient DNA Sequences Amplified From Nuclear Genes. Genetics, 2006. **172**(2): p. 733-741.

[83] Yin, Z., et al., Computing Platforms for Big Biological Data Analytics: Perspectives and Challenges. Computational and Structural Biotechnology Journal, 2017. **15**: p. 403-411.

[84] Gutaker, R.M. and H.A. Burbano, Reinforcing plant evolutionary genomics using ancient DNA. Current Opinion in Plant Biology, 2017. **36**: p. 38-45.

[85] Almathen, F., et al., Ancient and modern DNA reveal dynamics of domestication and cross-continental dispersal of the dromedary. Proceedings of the National Academy of Sciences, 2016. **113**(24): p. 6707-6712.

[86] Meyer, M., et al., A mitochondrial genome sequence of a hominin from Sima de los Huesos. Nature, 2014. **505**(7483): p. 403-406.

[87] Armbrecht, L., et al., Hybridisation capture allows DNA damage analysis of ancient marine eukaryotes. Scientific Reports, 2021. **11**(1): p. 3220.

[88] Ávila-Arcos, M.C., et al., Application and comparison of large-scale solution-based DNA capture-enrichment methods on ancient DNA. Scientific Reports, 2011. **1**(1): p. 74.

[89] Dabney, J., et al., Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proceedings of the National Academy of Sciences, 2013. **110**(39): p. 15758-15763.

[90] Enk, J., et al., Mammuthus Population Dynamics in Late Pleistocene North America: Divergence, Phylogeography, and Introgression. Frontiers in Ecology and Evolution, 2016. **4**.

[91] Parducci, L., et al., Ancient plant DNA in lake sediments. New Phytologist, 2017. **214**(3): p. 924-942.

[92] Kistler, L., et al., Transoceanic drift and the domestication of African bottle gourds in the Americas. Proceedings of the National Academy of Sciences, 2014. **111**(8): p. 2937-2941.

[93] Schmid, S., et al., HyRAD-X, a versatile method combining exome capture and RAD sequencing to extract genomic information from ancient DNA. Methods in Ecology and Evolution, 2017. **8**(10): p. 1374-1388.

[94] Ginolhac, A., et al., mapDamage: testing for damage patterns in ancient DNA sequences. Bioinformatics, 2011. **27**(15): p. 2153-2155.

[95] Jónsson, H., et al., mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics, 2013. **29**(13): p. 1682-1684.

[96] Kistler, L., et al., A new model for ancient DNA decay based on paleogenomic meta-analysis. Nucleic Acids Research, 2017. **45**(11): p. 6310-6320.

[97] Borry, M., et al., PyDamage: automated ancient damage identification and estimation for contigs in ancient DNA de novo assembly. PeerJ, 2021. **9**: p. e11845.

[98] Collin, T.C., et al., An open-sourced bioinformatic pipeline for the processing of Next-Generation Sequencing derived nucleotide reads: Identification and authentication of ancient metagenomic DNA. bioRxiv, 2020.

[99] Xu, W., et al., An efficient pipeline for ancient DNA mapping and recovery of endogenous ancient DNA from whole-genome sequencing data. Ecology and Evolution, 2021. **11**(1): p. 390-401.

[100] Knights, D., et al., Bayesian community-wide culture-independent microbial source tracking. Nature Methods, 2011. **8**(9): p. 761-763.

[101] Schubert, M., et al., Improving ancient DNA read mapping against modern reference genomes. BMC Genomics, 2012. **13**(1): p. 178.

[102] Huson, D.H., et al., MEGAN analysis of metagenomic data. Genome research, 2007. **17**(3): p. 377-386.

[103] Huson, D.H. and N. Weber, Microbial community analysis using MEGAN, in Methods in enzymology. 2013, Elsevier. p. 465-485.

[104] Cribdon, B., et al., PIA: More Accurate Taxonomic Assignment of Metagenomic Data Demonstrated on sedaDNA From the North Sea. Frontiers in Ecology and Evolution, 2020. **8**(84).

[105] Kanehisa, M., et al., KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Research, 2011. **40**(D1): p. D109-D114.

[106] Powell, S., et al., eggNOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. Nucleic Acids Research, 2011. **40**(D1): p. D284-D289.

[107] Overbeek, R., et al., The Subsystems Approach to Genome Annotation and its Use in the Project to Annotate 1000 Genomes. Nucleic Acids Research, 2005. **33**(17): p. 5691-5702.

[108] Edwards, R.A., et al., Real Time Metagenomics: Using k-mers to annotate metagenomes. Bioinformatics, 2012. **28**(24): p. 3316-3317.

[109] Prüfer, K., et al., Computational challenges in the analysis of ancient DNA. Genome Biology, 2010. **11**(5): p. R47.

[110] Tiley, G.P., C. Ané, and J.G. Burleigh, Evaluating and Characterizing Ancient Whole-Genome Duplications in Plants with Gene Count Data. Genome Biology and Evolution, 2016. **8**(4): p. 1023-1037.

[111] Dawson, S.C. and L.K. Fritz-Laylin, Sequencing free-living protists: the case for metagenomics. Environmental Microbiology, 2009. **11**(7): p. 1627-1631.

[112] Glöckner, G., et al., The Genome of the Foraminiferan Reticulomyxa filosa. Current Biology, 2014. **24**(1): p. 11-18.

[113] Mukherjee, S., et al., Genomes OnLine database (GOLD) v.7: updates and new features. Nucleic Acids Research, 2018. **47**(D1): p. D649-D659.

[114] Royo-Llonch, M., et al., Compendium of 530 metagenome-assembled bacterial and archaeal genomes from the polar Arctic Ocean. Nature Microbiology, 2021. **6**(12): p. 1561-1574.

[115] Arriola, L.A., A. Cooper, and L.S. Weyrich, Palaeomicrobiology: Application of Ancient DNA Sequencing to Better Understand Bacterial Genome Evolution and Adaptation. Frontiers in Ecology and Evolution, 2020. **8**.

[116] Segata, N., et al., Metagenomic microbial community profiling using unique clade-specific marker genes. Nature Methods, 2012. **9**(8): p. 811-814.

[117] Sunagawa, S., et al., Metagenomic species profiling using universal phylogenetic marker genes. Nature Methods, 2013. **10**(12): p. 1196-1199.

[118] Domaizon, I., et al., DNA-based methods in paleolimnology: new opportunities for investigating long-term dynamics of lacustrine biodiversity. Journal of Paleolimnology, 2017. **58**(1): p. 1-21.

[119] Giguet-Covex, C., et al., New insights on lake sediment DNA from the catchment: importance of taphonomic and analytical issues on the record quality. Scientific Reports, 2019. **9**(1): p. 14676.

[120] Keck, F., et al., Assessing the response of micro-eukaryotic diversity to the Great Acceleration using lake sedimentary DNA. Nature Communications, 2020. **11**(1): p. 3831.

[121] Siano, R., et al., Sediment archives reveal irreversible shifts in plankton communities after World War II and agricultural pollution. Current Biology, 2021. **31**(12): p. 2682-2689.e7.

[122] Shaw, J.L.A., et al., Retrospective eDNA assessment of potentially harmful algae in historical ship ballast tank and marine port sediments. Molecular Ecology, 2019. **28**(10): p. 2476-2485.

[123] Meyer, M., et al., Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. Nature, 2016. **531**(7595): p. 504-507.

[124] Wang, C.-C., et al., Ancient human genome-wide data from a 3000-year interval in the Caucasus corresponds with eco-geographic regions. Nature Communications, 2019. **10**(1): p. 590.