Environmental DNA

WILEY

Environmental DNA
Open Access
Dedicated to the study and use of environmental DNA for basic and applied sciences

ORIGINAL ARTICLE OPEN ACCESS

Tag Jumping Produces Major Distortion on Metabarcoding-Based Reconstructions of Past and Present Environments

S. Rodriguez-Martinez[1] | M. A. Morlock[2] | D. Y.-T. Huang[3] | J. Klaminder[1]

[1]Department of Forest Ecology and Management, Swedish University of Agricultural Sciences, Umeå, Sweden | [2]Department of Ecology and Environmental Science, Umeå University, Umeå, Sweden | [3]School of Geography, Politics, and Sociology, Newcastle University, Newcastle upon Tyne, UK

Correspondence: S. Rodriguez-Martinez (saul.rodriguez.martinez@slu.se; saul.rguezm@gmail.com)

Received: 27 July 2023 | **Revised:** 30 May 2025 | **Accepted:** 9 June 2025

Funding: This work was supported by Svenska Forskningsrådet Formas (Grant Dnr 2017-04548) and Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung (Grant P2BEP2-188256).

Keywords: ancient DNA | DNA barcoding | ecological parameter monitoring | environmental DNA | high-throughput nucleotide sequencing

ABSTRACT

Analysis of environmental DNA (eDNA) through DNA metabarcoding has become an important technique for environmental science as it allows precise reconstructions of species communities in a fast, cheap and non-invasive way. In this study, we scrutinize how environmental reconstructions derived from metabarcoding data may be affected by a process in which sample specific labels (tags), added to sequences for identification of individual samples, are changed unintentionally during adapter ligation causing translocation of sequences between samples ('tag jumping'). We compare animal and plant communities reconstructed using sedimentary eDNA records processed according two different protocols: (i) a twin-tagging approach (control) where all amplicons received the same tag on both sides ($N = 102$); and (ii) a combinatorial tagging protocol (affected by tag jumps) where each amplicon received a unique combination, but where some tags on each side were reused to form new combinations ($N = 102$). We analyzed six different sediment matrices and observed higher average number of taxa in the combinatorial tagging dataset in comparison to our twin-tagged dataset serving as a reference for results unaffected by tag jumps. In the control dataset with twin tagged amplicons, reconstructed animal communities were statistically different in 14 out of 15 pairwise comparisons, while only 8 out of 15 of the comparisons were different when samples were analyzed using the combinatorial tagging protocol. All of the inferred plant communities were statistically different when analyzed with a twin-tagging approach, while 20% of these plant communities were not different in our combinatorial tagged dataset. Our results clearly show that tag jumps added species to samples where they were not originally present and affects interpretations of species diversity and time-trends for whole communities. We conclude that tag jumping, being rarely discussed in metabarcoding studies, constitutes a concern in parity with direct sample contamination.

1 | Introduction

The use of molecular markers as a proxy for communities of different organisms has become state-of-the-art when monitoring ecosystems (Fernández et al. 2018) or reconstructing past environments from analysis of sedimentary deposits (Haile et al. 2009; Kjær et al. 2022). One of the most prevalently used approaches when analyzing eDNA is metabarcoding (Taberlet et al. 2012). In this approach, small fragments of DNA (barcodes) of species are sequenced after being amplified using polymerase chain reaction (PCR). It is a well-established approach that has gained popularity as high-throughput sequencing technologies

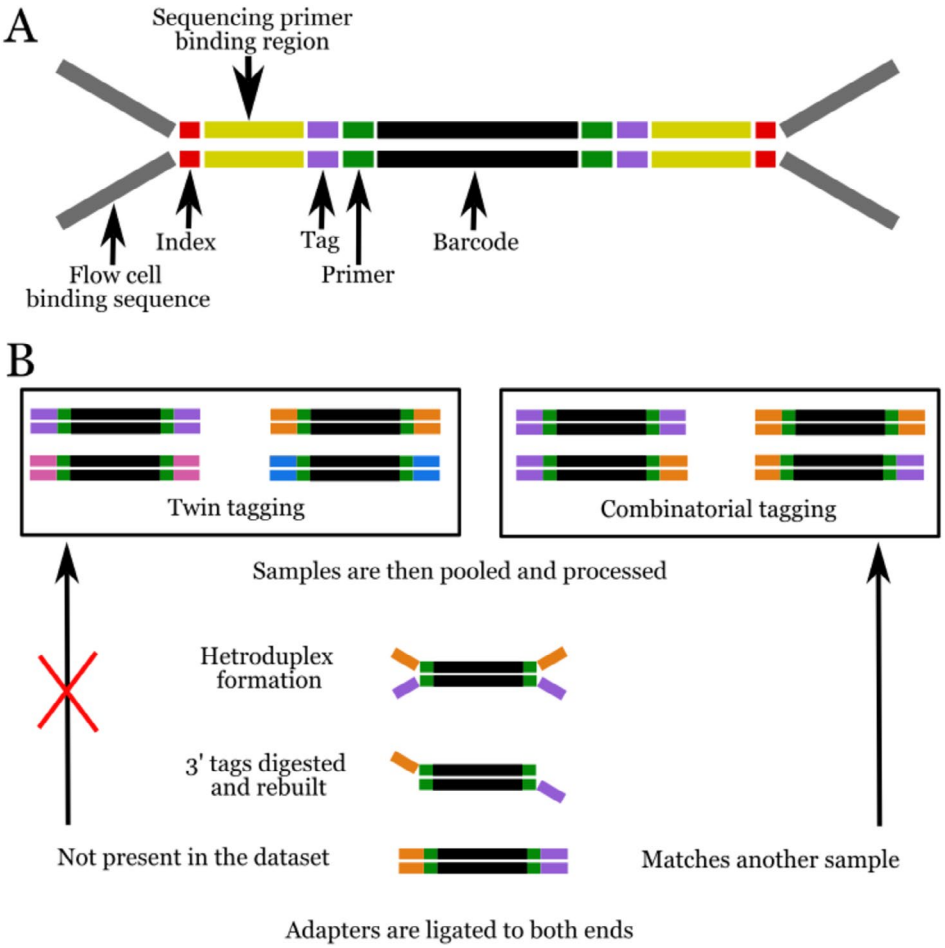

FIGURE 1 | Conceptual representation of the study. Panel A: Model of the DNA at the moment of sequencing with the naming as used in the text. Panel B: Conceptual design of the experiment with the end repair step that produces the tag jumps. During that process, single-stranded DNA from different amplicons interact, forming a heteroduplex. Then, T4 DNA polymerase effectuates the change. The resultant combination does not exist in the twin-tagged version of the dataset, but it matches a sample on the combinatorial-tagged one. After that, it cannot be separated from the actual sample.

evolved and became cheaper. Metabarcoding approaches have been used during the last decade to reconstruct past environments using sedimentary DNA preserved in deposits up to 45,000 years old (Zale et al. 2018). However, metabarcoding is still being perfected and purged of errors (Zaiko et al. 2021), and methodological uncertainties implicit in the technique have been highlighted and used as an argument for risks of flawed interpretations in eDNA-based paleoecological reconstructions (Birks and Birks 2016).

DNA contamination during sampling or laboratory handling is a commonly discussed source of errors when interpreting eDNA datasets (Birks et al. 2012; Gilbert et al. 2005) together with chimera formation (Fonseca et al. 2012). One additional source of errors is 'tag jumping' (Rodriguez-Martinez et al. 2022; Schnell et al. 2015). This process affects the 'tags', which are short oligonucleotides used as flags to match a sequence with its sample of origin (Figure 1A). The use of these tags allows for the pooling of tens to hundreds of samples in one single library, greatly increasing the throughput and maximizing sequencing cost efficiency. Tag jumps occur during library preparation where the multiplexed samples undergo the end-repair step (Schnell et al. 2015). During the end-repair step, single stranded sequences of DNA of two different amplicons can interact, forming a so called heteroduplex, in which each of the amplicons have their 3′ tags changed. If the new artefactual tag combination matches that of an actual sample, the sequence cannot be separated from authentic reads in downstream bioinformatic analysis.

The activity of the T4 DNA-polymerase used during the end-repair has been suggested to be the driving agent that changes the 3′ tags in heteroduplexes (Schnell et al. 2015), a theory that was recently supported by the topological patterns that tag jumps seems to generate within prepared libraries (Rodriguez-Martinez et al. 2022). These topological patterns do not only pinpoint toward T4 DNA polymerase activity being the cause behind tag jumping, but also exclude other cause of artefactual sequences, such as 'cluster bleeding', a process previously identified to distort metabarcoding datasets (Kircher et al. 2012). T4 DNA-polymerase driven tag jumps imply that there are likely methodological approaches that are not sensitive to tag jumps. For example, methods in which adapters are not ligated to tagged sequences (i.e., two PCR based methods) or methods in which the ligation is done on the adapter, making the tags protected from the polymerase activity (Rodriguez-Martinez et al. 2022). However, these protocols suffer from

other shortcomings including chimera amplification and reduced primer efficiency and specificity. The effect of tag jumping can also be circumvented by using each of the tags for only one reaction (twin-tagging) rather than reusing them for new combinations (combinatorial tagging) to reduce the number of tags needed for a batch of samples (Esling et al. 2015). With the former approach, all the artefactual tag-combinations that may form will not match any real samples and thus, these samples will be discarded during the bioinformatic analysis of the sequenced data.

The model of tag jumping outlined above predicts that sequences with the same tag combination may proceed from different samples and thus, that this process can potentially add DNA from species to samples where they were not originally present. It has also been suggested that tag jumping can make eDNA-inferred communities from two different samples appear more similar than they are as sequences may be shared between samples (Rodriguez-Martinez et al. 2022). This was proven by a cluster analysis in which the replicates clustered according to their tags rather than their sample of origin. However, the extent to which tag jumps affect metabarcoding data and potential interpretations remains unclear as the study by Rodriguez-Martinez et al. (2022) lacked a control counterpart to evaluate it. In this study, we produce a direct comparison between results with and without artifacts added by tag jumps and assess eventual differences in reconstructed communities when using twin-tagging protocols (circumventing tag jumping) and combinatorial tagging (tag jump sensitive) as a first direct measure for how tag jumping may affect between-sample differences and interpretation of eDNA data. We assess impacts of tag jumping on a sedimentary record from the Swiss Alps by comparing plant and mammal communities. A conceptual outline of the study design can be found in Figure 1B. For mammal DNA, we used the 16S mitochondrial rRNA gene (Taylor 1996), and for plant DNA, we used the trnL molecular markers (Taberlet et al. 2007). These two molecular markers have been frequently used in metabarcoding studies (Seersholm et al. 2020; van Vugt et al. 2022) making our comparison highly relevant for the published literature. We hypothesized that tag jumping: (i) adds species to samples where they were not originally present; and (ii) reduces between-sample variance for a randomly organized set of samples. We discuss our findings in a perspective of their implications on fundamental measures of eDNA based reconstructions.

2 | Material and Methods

Sediment cores were recovered in 2017 from the central basin of Lake Grosssee (47°04′43.8″ N, 9°14′47.1″ E, 1619 m.a.s.l.) situated in Flumserberg, Switzerland, with a percussion piston-coring system (Uwitec, Austria). Overlapping 3-m-long core sections were recovered with a horizontal offset of 1 m and a vertical offset of 0.5 m down to 8 m below the lake floor. Cores were split lengthwise and visually aligned to a composite of 7.73 m by using clearly identifiable flood event layers as markers (Glaus 2018). An age-depth model was established based on 17 radiocarbon dates from terrestrial plant macrofossils (Dwileski et al. 2025). The lake and its surroundings have experienced substantial vegetation change over the Holocene due to climate and human land use (Dwileski et al. 2025; Morlock et al. 2023). We selected six sediment samples from the core composite covering an age range from the early Holocene to sediment deposited during the last century to generate different analytical matrices typical for long-term palaeoecological studies. Sediments were subjected to the same subsampling and extraction protocols, but we applied two different tagging protocols: all samples were tagged using a twin-tagging protocol, and, in a separate analysis, all samples were processed using a combinatorial tagging protocol. With this approach, we could infer impacts of tag jumping on complex environmental matrices typical for the field of paleoecology by directly measuring eventual additional taxa (hypothesis 1) and altered between-sample differences (hypothesis 2) in the latter tag jump affected dataset. Subsampling, extraction, and tagging schemes are outlined in detail below.

Sub-sampling of the core was done in a clean room using standard ancient DNA precautions. DNA extractions were carried out in a dedicated ancient DNA lab at Umeå University that is isolated from other modern DNA labs and has a positive air pressure system accompanied by a HEPA air filter system. We followed protocols to avoid contamination during the extraction process (Paijmans et al. 2019), including extensive personal protective gear (facemask, gloves, and clean suit). Multiple extraction blanks and control amplifications were processed alongside all samples to screen for potential contamination. We extracted DNA from six depths along the sediment composite (13, 132, 232, 235, 352 and 529 cm composite depth). For each depth, three subsamples (extraction replicates, 0.5 g of material each) were extracted using the DNeasy Power Soil kit (QIAGEN, Germany), but incubating the samples overnight on a rotator at 60°C (lysis step). Five extraction blanks were extracted alongside the sediment samples. Both samples and blanks were handled in a similar way during all downstream processes including PCR amplification and sequencing. Together with the samples, positive controls were amplified and sequences (cow for animals and Picea sp. for plants). Quantifications of each extract can be found in Table S1.

In the libraries for mammals, we used primers (16Smam1 and 16Smam2) targeting mitochondrial 16S that amplify a ~150 bp fragment (Taylor 1996). The amplification was done on 25 μL solutions containing 1X PCR buffer (QIAGEN), 2 mM $MgCl_2$, 0.2 mM of dNTPs (QIAGEN), 0.4 μM of each primer, 0.625 U of HotStarTaq DNA polymerase (QIAGEN), 2 μg of bovine serum albumin (BSA) (Thermo Scientific) and 2 μL DNA template. In addition, we added 2 μM of a human specific blocker (Boessenkool et al. 2012) to the reaction to prevent the amplification of human contaminant DNA during PCR. PCR underwent the enzyme activation at 95°C for 15 min and then 45 cycles consisting of 95°C for 30 s, 60°C for 30 s, and 72°C for 30 s, followed by final elongation at 72°C for 10 min.

In the libraries for plants, the P6 loop region of the trnL (~140 bps) gene from the chloroplast was amplified with primers presented elsewhere (Taberlet et al. 2007). Amplification was carried out in a 20 μL reaction containing 1X PCR buffer (QIAGEN), 2 mM $MgCl_2$, 0.2 mM of dNTPs (QIAGEN), 0.4 μM of each primer, 1.25 U of HotStarTaq DNA polymerase

(QIAGEN), 4 μg of bovine serum albumin (BSA) (Thermo Scientific), and 1 μL DNA template. PCR underwent the enzyme activation at 95°C for 15 min and then 45 cycles consisting of 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s, followed by final elongation at 72°C for 10 min.

Two different tagging PCRs were carried out for each of the two primer sets (mammal and plant primers) using different tagging strategies, but the same extraction products. Note that this approach makes each primer set function as a separate experiment where the impact of the tagging is measured both for mammal and plant DNA. We designed 12-bp tags following Parameswaran et al. (2007), with any two tags differing in at least 5 nt. This design helps to mitigate misidentification of tags due to sequencing errors. In one of the tagging systems, each of the tags was used only once per library, matching forward and reverse tags (twin-tagging), a strategy that has proven to remove tag jumping errors from the results. In the other tagging PCR, we used combinations of tags, in which each forward tag is used multiple times with different reverse tags and vice versa (combinatorial tagging), which is sensitive to tag jumping. Three PCR replicates for each extract and extraction blanks were placed consecutively in 96 well plates (an example of the tag arrangement can be found in Figure S1), forming blocks with all replicates from a single extract using different tags for each. During this process, some (11) of the tag couples were skipped, and their respective wells contained neither primers nor DNA sample, which we call PCR blanks in the text. These blanks were placed in a way that there was at least one of them in each column and row in the tagging matrix (i.e., there is at least one blank for each forward and reverse tags). We used 5–10 blanks per plate in combination with 8–15 PCR negatives and 1–2 positives. For the first tagging, PCR was done using each of the tags only once with the protocol referred to as twin-tagging. The samples were assigned tags randomly, and PCR replicates were placed consecutively. In this protocol, there is no reason to include blanks as there are numerous unused tag combinations by the nature of the tagging arrangement, but we still had PCR negative and positive controls. To remove operator biases, all plates were prepared by the same person. Each of the libraries was run separately.

After PCR, 5 μL of all negatives and blanks were checked on a 2% agarose gel to confirm no visible contamination of the samples. Together, some of the samples and the positive control were checked as well to confirm positive amplification. Only if all negatives were clean and some samples confirmed the success of the PCR, the amplicons were further processed. After the gel visualization, amplicons were pooled equivolume for each specific region and purified. All pools were purified with QiaQuick PCR purification kit (QIAGEN, Germany) following the manufacturer's protocol.

Library preparation was conducted by Novogene (UK Company Limited) following the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB no. E7645) without the second PCR step. This was done following New England Biolabs instructions to use NEBNext Multiplex Oligos for Illumina with the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB no. E7645). The libraries were then sequenced on an Illumina NovaSeq 6000 PE150 platform. All our libraries were run in separate flow cells to avoid any chance of sample cross-talking between our own libraries during the sequencing.

2.1 | Bioinformatics Methods

Data quality was checked with fastQC (Andrews 2010) and all subsequent steps from raw data to assignment were carried out using OBITools3 (Boyer et al. 2016). In short, the following steps were carried out.

Reads merging: using the function *obi alignpairedend* (this function requires no parameters). Only aligned reads were kept, using the function *obi grep* with the parameter *-a mode: alignment*. Aligned reads were demultiplexed with the function *obi ngsfilter* allowing no mismatches on the tags with the parameter *–e 0*. The file was dereplicated by removing strictly identical sequences with *obi uniq*, keeping only the counts in each sample. Quality filtering of sequences shorter that 40 base pairs (bps) or with counts smaller than 10 in the whole library were removed using *obi grep*. Further removal of low repeated sequences was done manually after the assignment as explained below. Removal of PCR and sequencing errors was done with the function *obi clean*. Only head sequences (*-H* option) with no variants and a count greater than 5% of their own count (*–r 0.05*) were kept. Taxonomic assignment was done with the *obi ecotag* function to a 97% similarity threshold with the best match. The database was composed of the relevant genomic region for each primer pair and the whole kingdom, plants and animals respectively, to avoid misidentifications. These databases were downloaded from NCBI by searching the genome region and filtering the taxonomy to the adequate kingdom in each case. The 16S database for animals was downloaded on the 9th of December 2020 and the trn-L region database for plants on 14th on January 2021. All taxonomic classifications follow the NCBI Taxonomy (Schoch et al. 2020). Breakdown of reads after each step can be found in Table S1.

The results were exported to spreadsheets and further curated, starting by collapsing all assignments to different common taxonomic levels. Taxa that could not be resolved to the genus level were discarded. Then, taxa with low counts per sample were removed (counts lower than 10 in a sample) and finally, low repeated OTUs were removed as well (counts below 100 across the whole library). All denoising steps so far were done on all datasets. In addition, two taxa (*Picea* and *Alnus*) were eliminated from the twin tagged datasets as they were found in high counts in the negatives, while that was impossible in the tag-jumped datasets as that would mean eliminating virtually every taxon detected in the whole dataset.

2.2 | Statistical Analysis

All statistical analysis were conducted using R (R Core Team 2023) and the vegan package (Oksanen et al. 2022). To account for disparities in sequencing depth leading to biases in the number of reads in different libraries, we conducted a series of statistical tests that are very robust to said disparities. We also analyzed the libraries independently, which removes

dependency on reads per library of the analysis. We calculated a Bray–Curtis dissimilarity matrix (Bray and Curtis 1957) on the raw counts. These distances were corrected to produce a fully Euclidean representation without negative Eigenvalues in accordance to previously outlined methods (Legendre and Anderson 1999). The calculated distances were visualized on a biplot using a principal coordinates (PCoA) plot with centroids and one standard deviation bounds (Gower 2015). For testing statistical differences between samples we used a permutational multivariate analysis of variance (PERMANOVA) described in detail elsewhere (Anderson 2017). As multiple comparisons were carried out, which increase the risk of generating spurious positives, we applied the Holm correction when testing for significant effects (Holm 1979). The Eigenvalues of the PCoA were used to examine the variance. A permutation-based test of multivariate homogeneity of group dispersions was used to evaluate if the differences were attributed to centroid position (between sample differences) or to dispersion (within sample variance). These tests were carried out for each individual library and their respective results were then compared. In addition, multivariate distances between centroids and sum of eigenvalues in both twin-tagged and combinatorial-tagged datasets were compared to evaluate the differences in overall variance in the datasets.

3 | Results

In total, 18 mammal genera and 75 plant genera were found in the studied sediment record. In the twin-tagged datasets (our control), two taxa (*Picea* and *Alnus*) were found in the plant dataset PCR negative controls (up to 53,000 reads for *Picea* and 124,000 reads for *Alnus*) assumed to be due to reagent contamination, and these taxa were subsequently removed from the analysis. Besides *Picea* and *Alnus*, the blanks from the twin-tagged dataset were free from reads. In contrast, virtually every taxon detected in samples was also present in the blanks (range 55–36,664 reads in the mammal dataset and 0–22,356 reads in the plants dataset) and negatives of the dataset with combinatorial tagged samples (dataset sensitive to tag jumps). Dominant mammal taxa in both twin-tagged and combinatorial-tagged datasets were *Bos*, *Ovis*, and *Capra*, but others such as *Terricola* or *Lepus* were also found in some of the samples. In both plant datasets, the most abundant taxon was *Cirsium* after removing reagent contamination. A substantial number of other taxa, like *Poa* or *Gentiana*, were found in lower read numbers and less generally distributed. Breakdown of the taxa in each sample can be found in Table S2.

The average detection rate in the mammals datasets was 6.8 (range 0–12) taxa per sample in the twin tagged version, while the combinatorial tagged average detection was 11.8 (8–15) taxa per sample. Similarly, the plants twin tagged dataset had an average detection of 20.3 (14–30) taxa per sample, while the combinatorial tagged dataset had an average detection of 34.9 (13–52) taxa per sample. The major difference in taxa detection was viewed as a measure of incoming reads due to tag jumping in the latter tag jump sensitive dataset. For example, in the mammals data, taxa such as *Cervus*, *Sus*, or *Ovis* were restricted to only a few samples, while in the combinatorial tagged dataset, they were ubiquitous (Table S2).

The ordination on a reduced space (Figure 2A–D) suggested that our control dataset contained more dissimilar samples (Figure 2A) in comparison to its combinatorial tagged counterpart (Figure 2C). For example, mammal communities in samples 529-cm (Early Holocene) and 13-cm (modern sediments) were both clearly separated from each other and from those of other samples in the twin tagged dataset, but had overlap with one or several samples in the combinatorial tagged dataset. That both old and recently deposited sediment showed mammal communities that overlapped with that of sediments with intermediate age was not due to differences in dispersions around the mean value (Figure 2B,D), but rather driven by shorter inter-centroid distances in the combinatorial tagged dataset (Figure S2A,B, Table S1). We also observed a reduction in the variance for the tag jump sensitive combinatorial tagged dataset, the variance of the twin tagged dataset being 200% higher. This homogenization effect on the combinatorial tagged datasets generated a clustering of the variance along the main coordinate (Figure S3A,B). A complete overview of the Bray-Curtis dissimilarities can be found in Table S2.

Differences between the twin tagged dataset and the combinatorial dataset were also seen for reconstructed plant communities (Figure 3A–D). Here, plant communities in the 529-cm (Early Holocene) and 13-cm (modern sediment) were clearly separated from those of other samples in the twin tagged dataset. These differences became less clear in the combinatorial tagged dataset, mainly due to single replicates from the 13 cm sample overlapping with older samples (132-, 234- and 231-cm) and a single replicate of the 352-cm sample plotting within one standard deviation from the centroid of the 13 cm sample. Distances between sample centroids for the plant communities were larger in the twin tagged dataset in comparison to the combinatorial tagged dataset, with no apparent differences in dispersions around the mean value for the two datasets. However, the variance was 25% higher in the twin tagged dataset (Figure S2C,D). The homogenization effect seen in the combinatorial dataset generated a clustering of the variance along the main coordinate (Figure S3C,D), which produced a strong arch effect in the PCoA plots (Figures 2C and 3C).

In the twin-tagged data, our PERMANOVA confirmed significantly different mammal communities for 14 out of 15 couples of sediment samples ($p < 0.05$, d.f = 1). In contrast, only 8 out of 15 comparisons showed significant differences in the combinatorial tagged dataset. Like the mammal communities, plant communities in the twin-tagged dataset showed higher dissimilarities between samples than what was observed for the tag jumping sensitive data (Figure 3A–D). In the twin-tagged dataset, all samples were significantly different, while 12 out of 15 possible comparisons in the combinatorial tagged were significantly different from each other ($p < 0.05$, d.f = 1). Full breakdown of the results is shown in Table S2.

We plotted first principal coordinate (PCoA-1) scores in a chronological order (increasing sediment depth) to assess how temporal trends in the mammal and plant communities differed between the two tagging protocols (Figure 4A,B). In the twin-tagged dataset, it was apparent that the mammal and plant communities follow the same 'hump-shaped' temporal trend with major transitions in species composition in Early Holocene

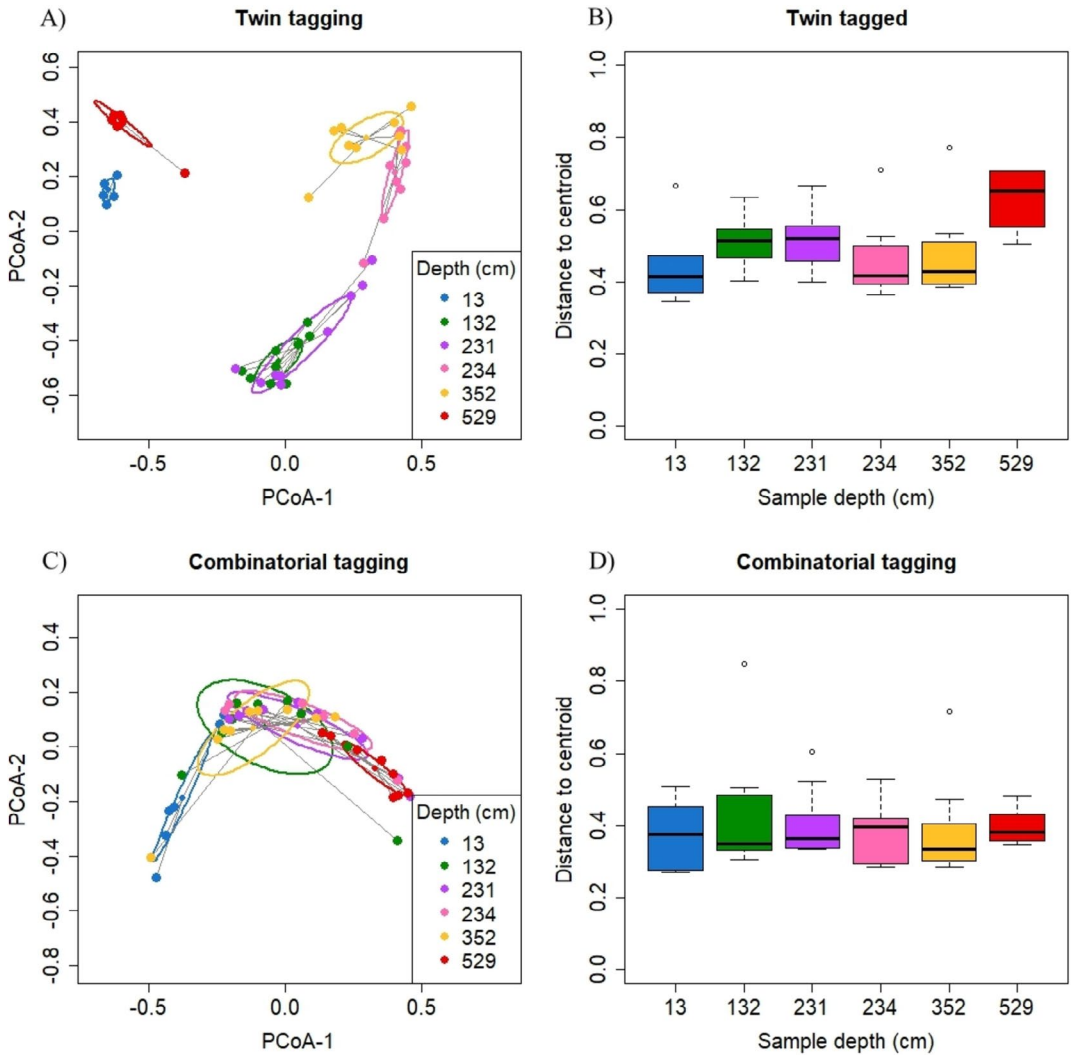

FIGURE 2 | Mammals datasets. Panels A (top left) and B (top right) correspond to the twin-tagged dataset principal coordinates analysis (PCoA) and dispersion, respectively. Panels C (bottom left) and D (bottom right) correspond to the tag-jumped dataset principal coordinates and dispersions, respectively. Panels A and C: Ellipses mark one standard deviation from the centroid. Panels B and D: Bar-median, box-interquartile range Q1–Q3, whiskers-all data points, circles-outliers. Numbers indicate sampled sediment core depth in cm.

sediment and in more recently deposited sediment (Figure 4A). In contrast, in the combinatorial tagged datasets (Figure 4B), the PCoA-1 score for mammal communities showed an increasing trend with increasing sediment depth. However, the PCoA-1 score for the plant communities was clearly separated from that describing the mammal communities, a result in clear contrast to the tight coupling between plant and mammal communities seen for the control dataset. To exemplify the origin of the differences we plotted the detection of 2 taxa, cow and deer, across the sediment column.

4 | Discussion

Published literature has not yet fully recognized the problem with tag jumping and often lacks clear information to evaluate whether the process may have affected the data or not. While studies have previously identified tag jumping as a process (Esling et al. 2015; Rodriguez-Martinez et al. 2022; Schnell et al. 2015), the scarcity of studies evaluating the strength of distortion on metabarcoding data may explain the limited discussion about tag jumps in the literature. In our study, we provide direct measures illustrating how tag jumps can affect metabarcoding data, by comparing twin-tagged and combinatorial-tagged eDNA datasets. In line with our first hypothesis, tag jumps caused substantial alteration of the mammal and plant communities by introducing species in several samples where they were not originally present. Note that the higher average number of species in the combinatorial dataset was not driven by the addition of rare species in single samples, an effect that can be caused by small differences in PCR amplifications or sequencing performance, but rather driven by an increase in species numbers for almost all samples. This strong effect is evident from the > 70% increase in mammal and plant taxa per sample in the whole dataset. In other words, the number of taxa per sample was increased without increasing the total number of taxa in the library. In line with our second hypothesis, tag jumping reduced between-sample variation. The observed reduction in between-sample variation in the combinatorial tagged eDNA dataset in comparison

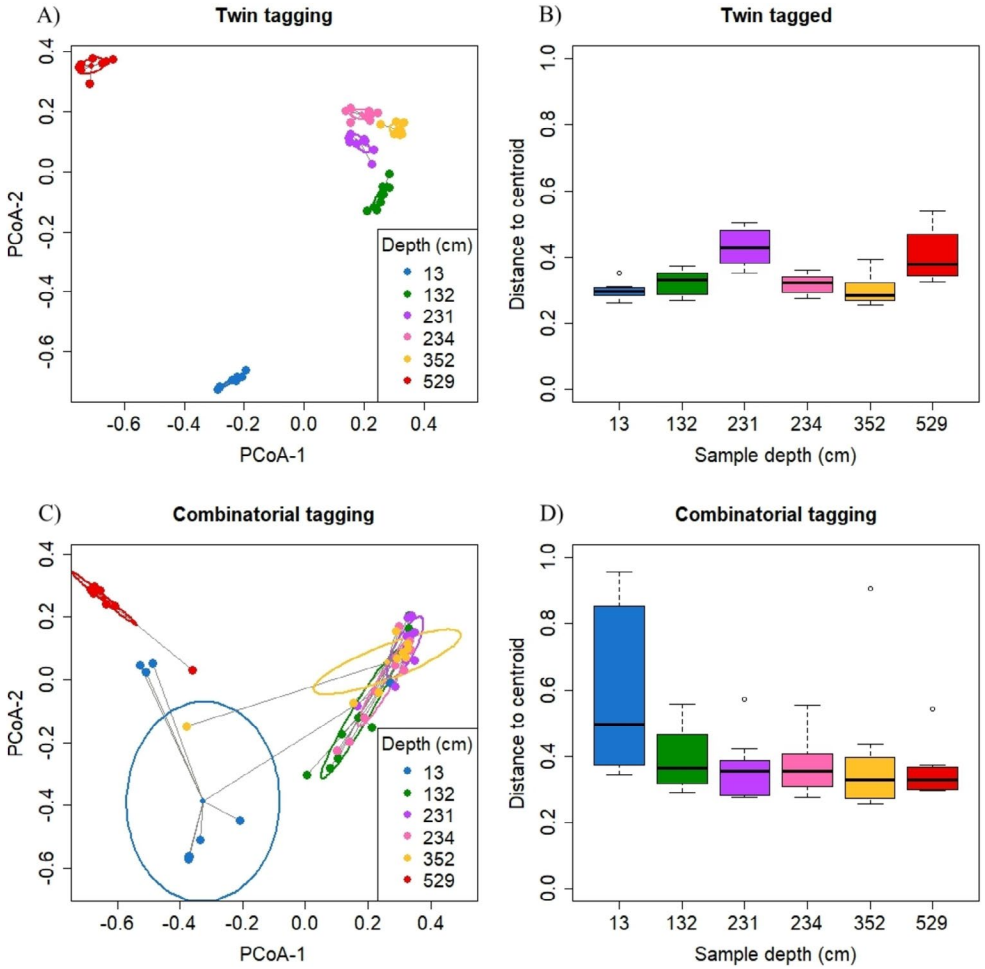

FIGURE 3 | Plants datasets. Panels A (top left) and B (top right) correspond to the twin-tagged dataset principal coordinates analysis (PCoA) and dispersion, respectively. Panels C (bottom left) and D (bottom right) correspond to the tag-jumped dataset principal coordinates and dispersions, respectively. Ellipses mark one standard deviation from the centroid. Panels B and D: Bar-median, box-interquartile range Q1–Q3, whiskers-all data points, circles-outliers. Numbers indicate sample depth in cm.

to the twin tagged data, supports an explanatory model where tag jumps produce cross-contamination between different amplicons reducing the dissimilarity between samples as they share barcodes from a common DNA pool. As a result, many of the samples in our combinatorial tagged eDNA dataset showed insignificant differences, even though the twin tagged data clearly revealed that the sediment they originated from contained very different eDNA communities (Figures 2 and 3). However, insignificant differences between samples from the combinatorial tagged eDNA dataset were not only caused by shared DNA, but also by the increased within replicate variation due to the non-homogenous distribution of the tag jumping patterns. Increased within-sample variance also reduces the likelihood of finding statistically significant differences between sediment layers. Importantly, tag jumps depend on the sample's position in the tagging matrix (Rodriguez-Martinez et al. 2022) and we are aware that there is a theoretical chance that tag jumps could reduce differences between replicates (reduce within-sample dispersion) if they all share one of the tags. During conditions where replicates experience reduced within-sample variability, it seems likely that tag jumping can also, in contrast to what we observed in our results, create samples that appear statistically different from other samples even though they are not.

In our experiment, high levels of translocated sequences were found in the tag jumping affected datasets (both for mammals and plants). However, the twin tagged datasets were free of artefactual sequences in the blanks. Given that the main difference between the two datasets is the introduction of T4 polymerase driven tag jumps in the results of the combinatorial tagged eDNA, it seems evident that the artefactual sequences in the blanks are a product of tag jumping. The magnitude of the contamination is difficult to evaluate in read numbers, but the fact that more taxa are found on average in the samples of the combinatorial datasets and their higher similarity constitute very strong evidence of the potential repercussions of tag jumping in datasets. It is important to mention here that the finding of more taxa in samples of the combinatorial dataset is not paired with higher diversity of the dataset. In fact, the combinatorial datasets present a slightly lower overall diversity than the twin-tagged ones-possibly due to small differences in sequencing depth-while the average number of taxa per sample and the range of taxa per sample are both higher. Effects of tag jumping were more severe in the mammal dataset in comparison to the plant dataset. That is, eigenvalue-inferred variance differed by around 200% in the mammals datasets (Figure S2C,D), while this difference was around 25% in the plants dataset (Figure S2A,B). While

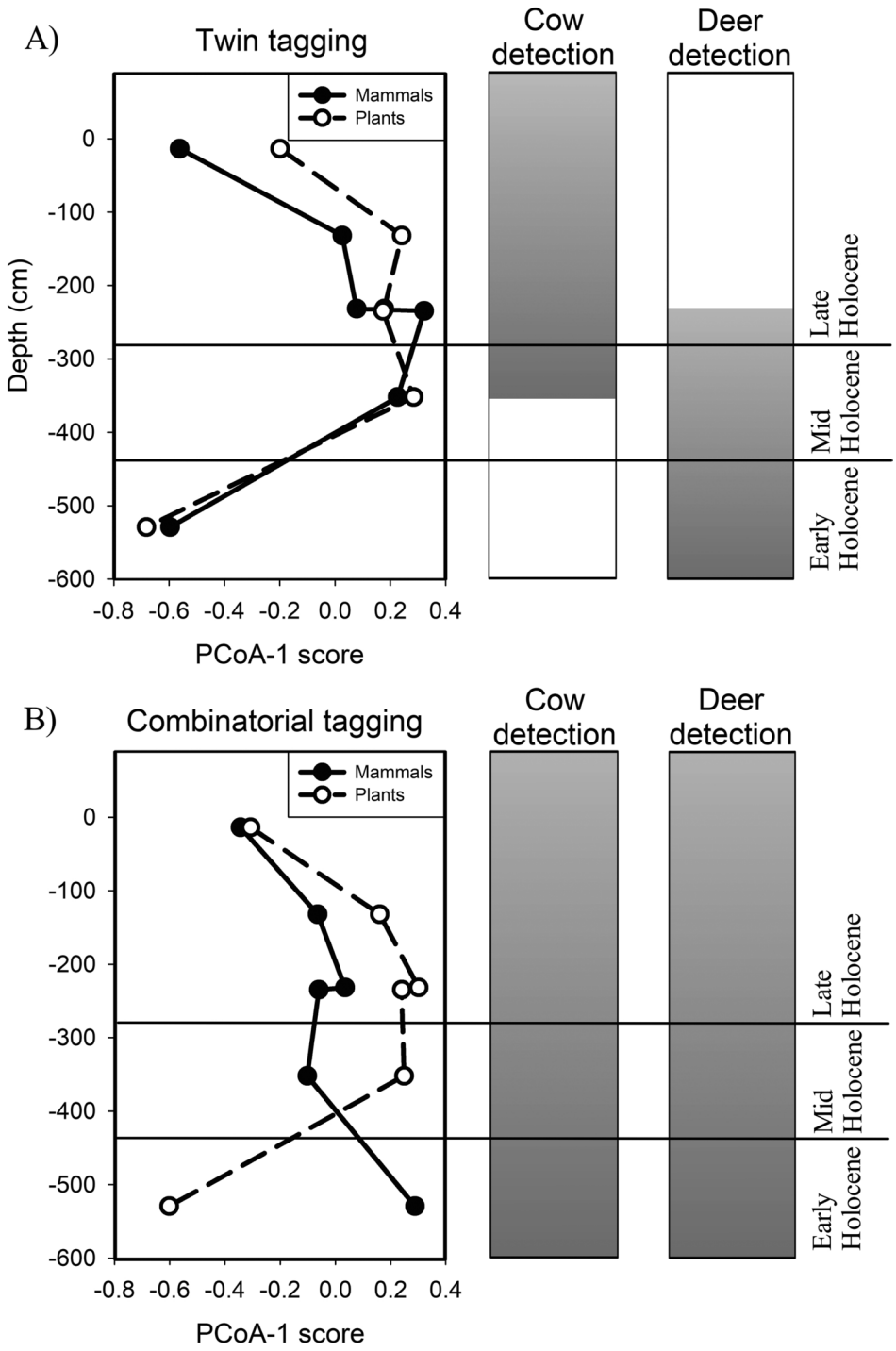

FIGURE 4 | First principal coordinate (x-axis) plotted against sediment depth and detections of two of the mammal species; the shadowed part indicates positive detection of the species. Data for control twin-tagged datasets (top) and affected combinatorial-tagged datasets (bottom).

we hesitate to generalize our finding to be universal for all metabarcoding studies, we highlight that the result may suggest that different genetic markers exhibit different levels of tag jumping. That is, some genetic regions may be more prone to form heteroduplexes as it has been noticed in other heteroduplex driven processes (Shin et al. 2014; Wang and Wang 1997). Differences in tag jumping impacts in our datasets cannot be attributed to tag properties because these properties unlikely explain the heteroduplex formation driving the jumps. It should be noted, however, that the arrangement of samples within the tagging matrix and hence, slight differences in tag arrangements may explain some observed differences in effects caused by tag jumping between the mammal and plant datasets. Moreover, impacts from tag jumps will be more severe for low diversity datasets, i.e., as the introduction of a few species to a low diversity community will have a larger relative impact. Therefore, we suggest that the effects of tag jumping should preferably be evaluated on a case-by-case basis. Within paleoecological reconstructions, the first or last appearance of species-specific DNA within a record is often

interpreted in a context of ecological change. Here, the oldest DNA is often used to detect the arrival of migrating species (Nota et al. 2022), extinction (Graham et al. 2016), first use of agricultural crops (Smith et al. 2015), introduction of domestic mammals (Giguet-Covex et al. 2014) or the onset of human activities (Pansu et al. 2015). In similar, the youngest DNA of extinct species can be used to temporally constrain historic extinction events (Allentoft et al. 2010; Haile et al. 2009). As illustrated by our results, tag jumping introduces positive detections of species in samples where they were not originally present. By adding species both in old and recently deposited sediments, tag jumping can affect ecological important measures such as the 'earliest' and 'youngest' detections of species. For example, if we were to interpret the mammal data without being aware of the tag jumping in the combinatorial-tagged dataset, we would have concluded that domestic animal DNA, such as cow, was present in sediments deposited shortly after deglaciation. As expected, the control dataset shows a much later arrival, and thus, this former interpretation is clearly flawed.

One of the strengths of eDNA-based palaeoecological studies is their ability to detect long-term environmental change by interpreting changes in whole communities rather than interpreting the presence of single indicator species. Our results showed that tag jumping is a process potent enough to reduce between-sample variance at a whole community scale and fundamentally alter reconstructed time trends for mammal and plant communities. First, tag jumping made samples from sediment deposited during the Holocene more similar; hence, this artificial homogenization of samples blurred apparent changes in the mammal and plant communities. Second, tag jumps increased within-sample variation, an effect that made it more difficult to detect significant community changes for sections of the sediment record. The latter effect, in combination with the artefactual homogenization of the communities that tag jumps generated, caused an asymmetrical impact with the strongest effects in the older, less concentrated, and less diverse samples in our dataset. Temporal and spatial changes in plant and mammal communities are the central theme of eDNA research focusing on natural history (Pansu et al. 2015) or ecosystem changes caused by external drivers such as human land use (Smith et al. 2015) and climate change (Rosa et al. 2022). Tag jumping may not only mask substantial environmental changes in metabarcoding studies, but also decouple existing common trends for mammal and plant eDNA communities, as seen by the contrasting temporal trends reconstructed in the combinatorial tagged dataset.

Our observation of strong effects of tag jumping on taxa composition, sample heterogeneity, time trends, and the decoupling between plant and mammal diversity –all important ecological measures –highlights that tag jumping is a process of great concern for eDNA-based studies. While our study emphasizes tag jumping as a problem for metabarcoding studies, it seems plausible that other techniques using combinatorial tags ligated in a library (e.g., metagenomic studies) may as well be affected. Moreover, all efforts to remove the noise induced by tag jumping have proven futile simply because this process can produce a high number –spanning into the tens of thousands– of translocated sequences that account for high proportions in each sample (Rodriguez-Martinez et al. 2022). Removing such high numbers would effectively remove most of the dataset. This means that the original information can no longer be inferred from the affected datasets. With this realization in mind, an intuitive question becomes: are there published studies that have made interpretations based on datasets affected by tag jumping? Tag jumping is a heteroduplex-driven process, intimately tied to early pooling in methodological pipelines in which adapters are ligated rather than added in a second PCR. These ligation-based pipelines have been used in roughly ten thousand papers over the last decade according to the Illumina databases (https://emea.illumina.com/techniques/sequencing/ngs-library-prep/ligation.html). In other words, if only a fraction of these studies followed tag jumping-sensitive protocols, there is a high risk that a substantial number of scientific papers may include datasets affected by tag jumps. Unfortunately, method descriptions in metabarcoding studies often fail to clearly specify the details of library construction and preparation needed to review the likelihood of tag jumping by the use of extensive sets of blanks embedded in the tagging matrix. Hence, for most published studies, it is not possible to assess to what extent their conclusions might have been affected by tag jumping. Importantly, our study demonstrates impacts of tag jumping on some ligation-based libraries while also revealing that an alternative ligation-based strategy can circumvent the issue. We recognize the ligation methodology has advantages, like being less prone to chimera amplification, but considering the strong effects that tag jumping can have on eDNA datasets, better methodological descriptions of the pipelines and acknowledgment of the potential errors that tag jumping can cause seem of critical importance. Data distortion of the magnitude that we observed would affect interpretations independent of the nature of the metabarcoding study. We underline the need for recognizing tag jumps as a process of concern for metabarcoding reconstructions, similar to the more established concern for sample contamination during lab handling. Substantial efforts are currently devoted toward avoiding laboratory and field contamination when studying eDNA (Fulton 2012; Giguet-Covex et al. 2014). In perspective of our results, tag jumping appears to be at least an equally important source of errors in metabarcoding studies as direct sample contamination.

5 | Conclusion

Our study clearly shows that twin-tagging effectively removes all noise caused by tag jumping and that tag jumping is potent enough to fundamentally change whole plant and mammal communities inferred from eDNA. Tag jumping introduced false positive detections of different taxa and made plant and mammal communities in our samples more similar. Moreover, the tag jumping process influenced replicates differently, causing increased within-sample variance that made detection of between-sample differences in species composition more difficult. Our study highlights the need to recognize tag jumping not only in palaeoecological context but in all metabarcoding studies where species and the biodiversity of communities are of interest. With our result in mind, it seems plausible that tag jumping may have masked substantial environmental changes in studies that have used combinatorial tagging instead of twin-tagging approaches. Therefore, we urge the research community to recognize tag jumping as a serious problem and highlight

the need to improve methodological descriptions related to tagging approaches.

Author Contributions

S.R.-M., M.A.M. and J.K. designed the study. S.R.-M., M.A.M. and D.Y.-T.H. did the data acquisition and analysis. S.R.-M., M.A.M. and J.K. wrote the manuscript.

Acknowledgments

We thank the Swedish research council for founding J.K. (Dnr 2017-04548) and the Swiss National Science Foundation for founding M.A.M. (grant no. P2BEP2-188256).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study is openly available in Dryad at DOI: 10.5061/dryad.41ns1rnqq.

References

Allentoft, M. E., M. Bunce, R. P. Scofield, M. L. Hale, and R. N. Holdaway. 2010. "Highly Skewed Sex Ratios and Biased Fossil Deposition of Moa: Ancient DNA Provides New Insight on New Zealand's Extinct Megafauna." *Quaternary Science Reviews* 29: 753–762. https://doi.org/10.1016/j.quascirev.2009.11.022.

Anderson, M. J. 2017. "Permutational Multivariate Analysis of Variance (PERMANOVA)." In *Wiley StatsRef: Statistics Reference Online*, edited by N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri, and J. L. Teugels, 1–15. Wiley. https://doi.org/10.1002/9781118445112.stat07841.

Andrews, S. 2010. *FastQC: A Quality Control Tool for High Throughput Sequence Data*.

Birks, H. H., T. Giesecke, G. M. Hewitt, P. C. Tzedakis, J. Bakke, and H. J. B. Birks. 2012. "Comment on 'Glacial Survival of Boreal Trees in Northern Scandinavia'." *Science* 338: 742. https://doi.org/10.1126/science.1225345.

Birks, H. J. B., and H. H. Birks. 2016. "How Have Studies of Ancient DNA From Sediments Contributed to the Reconstruction of Quaternary Floras?" *New Phytologist* 209: 499–506. https://doi.org/10.1111/nph.13657.

Boessenkool, S., L. S. Epp, J. Haile, et al. 2012. "Blocking Human Contaminant DNA During PCR Allows Amplification of Rare Mammal Species From Sedimentary Ancient DNA: Blocking Primers for Environmental Ancient DNA." *Molecular Ecology* 21: 1806–1815. https://doi.org/10.1111/j.1365-294X.2011.05306.x.

Boyer, F., C. Mercier, A. Bonin, Y. Le Bras, P. Taberlet, and E. Coissac. 2016. "Obitools: A Unix -Inspired Software Package for DNA Metabarcoding." *Molecular Ecology Resources* 16: 176–182. https://doi.org/10.1111/1755-0998.12428.

Bray, J. R., and J. T. Curtis. 1957. "An Ordination of the Upland Forest Communities of Southern Wisconsin." *Ecological Monographs* 27: 325–349. https://doi.org/10.2307/1942268.

Dwileski, A. R., F. Rey, M. A. Morlock, et al. 2025. "Holocene Vegetation Change at Grosssee, Eastern Swiss Alps: Effects of Climate and Human Impact." *Vegetation History and Archaeobotany* 34: 331–348. https://doi.org/10.1007/s00334-024-01014-7.

Esling, P., F. Lejzerowicz, and J. Pawlowski. 2015. "Accurate Multiplexing and Filtering for High-Throughput Amplicon-Sequencing." *Nucleic Acids Research* 43: 2513–2524. https://doi.org/10.1093/nar/gkv107.

Fernández, S., S. Rodríguez, J. L. Martínez, Y. J. Borrell, A. Ardura, and E. García-Vázquez. 2018. "Evaluating Freshwater Macroinvertebrates From eDNA Metabarcoding: A River Nalón Case Study." *PLoS One* 13: e0201741. https://doi.org/10.1371/journal.pone.0201741.

Fonseca, V. G., B. Nichols, D. Lallias, et al. 2012. "Sample Richness and Genetic Diversity as Drivers of Chimera Formation in nSSU Metagenetic Analyses." *Nucleic Acids Research* 40: e66. https://doi.org/10.1093/nar/gks002.

Fulton, T. L. 2012. "Setting up an Ancient DNA Laboratory." In *Ancient DNA, Methods in Molecular Biology*, edited by B. Shapiro and M. Hofreiter, 1–11. Humana Press. https://doi.org/10.1007/978-1-61779-516-9_1.

Giguet-Covex, C., J. Pansu, F. Arnaud, et al. 2014. "Long Livestock Farming History and Human Landscape Shaping Revealed by Lake Sediment DNA." *Nature Communications* 5: 3211. https://doi.org/10.1038/ncomms4211.

Gilbert, M. T. P., H.-J. Bandelt, M. Hofreiter, and I. Barnes. 2005. "Assessing Ancient DNA Studies." *Trends in Ecology & Evolution* 20: 541–544. https://doi.org/10.1016/j.tree.2005.07.005.

Glaus, N. 2018. *Holocene Flood Events and Sediment Flux in Alpine Lake Grosssee*. University of Bern. Master Thesis.

Gower, J. C. 2015. "Principal Coordinates Analysis." In *Wiley StatsRef: Statistics Reference Online*, edited by N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri, and J. L. Teugels, 1–7. Wiley. https://doi.org/10.1002/9781118445112.stat05670.pub2.

Graham, R. W., S. Belmecheri, K. Choy, et al. 2016. "Timing and Causes of Mid-Holocene Mammoth Extinction on St. Paul Island, Alaska." *Proceedings of the National Academy of Sciences of the United States of America* 113: 9310–9314. https://doi.org/10.1073/pnas.1604903113.

Haile, J., D. G. Froese, R. D. E. MacPhee, et al. 2009. "Ancient DNA Reveals Late Survival of Mammoth and Horse in Interior Alaska." *Proceedings of the National Academy of Sciences* 106: 22352–22357. https://doi.org/10.1073/pnas.0912510106.

Holm, S. 1979. "A Simple Sequentially Rejective Multiple Test Procedure." *Scandinavian Journal of Statistics* 6: 65–70.

Kircher, M., S. Sawyer, and M. Meyer. 2012. "Double Indexing Overcomes Inaccuracies in Multiplex Sequencing on the Illumina Platform." *Nucleic Acids Research* 40, no. 1: e3. https://doi.org/10.1093/nar/gkr771.

Kjær, K. H., M. Winther Pedersen, B. De Sanctis, et al. 2022. "A 2-Million-Year-Old Ecosystem in Greenland Uncovered by Environmental DNA." *Nature* 612: 283–291. https://doi.org/10.1038/s41586-022-05453-y.

Legendre, P., and M. J. Anderson. 1999. "Distance-Based Redundancy Analysis: Testing Multispecies Responses in Multifactorial Ecological Experiments." *Ecological Monographs* 69: 1–24. https://doi.org/10.1890/0012-9615(1999)069[0001:DBRATM]2.0.CO;2.

Morlock, M. A., S. Rodriguez-Martinez, D. Y. Huang, and J. Klaminder. 2023. "Erosion Regime Controls Sediment Environmental DNA -Based Community Reconstruction." *Environmental DNA* 5: 1393–1404. https://doi.org/10.1002/edn3.458.

Nota, K., J. Klaminder, P. Milesi, et al. 2022. "Norway Spruce Postglacial Recolonization of Fennoscandia." *Nature Communications* 13: 1333. https://doi.org/10.1038/s41467-022-28976-4.

Oksanen, J., G. L. Simpson, F. G. Blanchet, et al. 2022. *vegan: Community Ecology Package*.

Paijmans, J. L., G. González Fortes, and D. W. Förster. 2019. "Application of Solid-State Capture for the Retrieval of Small-to-Medium Sized Target Loci from Ancient DNA." *Methods in Molecular Biology* 1963: 129–139.

Pansu, J., C. Giguet-Covex, G. F. Ficetola, et al. 2015. "Reconstructing Long-Term Human Impacts on Plant Communities: An Ecological

Approach Based on Lake Sediment DNA." *Molecular Ecology* 24: 1485–1498. https://doi.org/10.1111/mec.13136.

Parameswaran, P., R. Jalili, L. Tao, et al. 2007. "A Pyrosequencing-Tailored Nucleotide Barcode Design Unveils Opportunities for Large-Scale Sample Multiplexing." *Nucleic Acids Research* 35: e130. https://doi.org/10.1093/nar/gkm760.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Rodriguez-Martinez, S., J. Klaminder, M. A. Morlock, L. Dalén, and D.Y.-T. Huang. 2022. "The Topological Nature of Tag Jumping in Environmental DNA Metabarcoding Studies." *Molecular Ecology Resources* 23: 13745. https://doi.org/10.1111/1755-0998.13745.

Rosa, L. H., M. B. Ogaki, J. M. Lirio, et al. 2022. "Fungal Diversity in a Sediment Core From Climate Change Impacted Boeckella Lake, Hope Bay, North-Eastern Antarctic Peninsula Assessed Using Metabarcoding." *Extremophiles* 26: 16. https://doi.org/10.1007/s00792-022-01264-1.

Schnell, I. B., K. Bohmann, and M. T. P. Gilbert. 2015. "Tag Jumps Illuminated – Reducing Sequence-To-Sample Misidentifications in Metabarcoding Studies." *Molecular Ecology Resources* 15: 1289–1303. https://doi.org/10.1111/1755-0998.12402.

Schoch, C. L., S. Ciufo, M. Domrachev, et al. 2020. "NCBI Taxonomy: A Comprehensive Update on Curation, Resources and Tools." *Database* 2020: baaa062. https://doi.org/10.1093/database/baaa062.

Seersholm, F. V., D. J. Werndly, A. Grealy, et al. 2020. "Rapid Range Shifts and Megafaunal Extinctions Associated With Late Pleistocene Climate Change." *Nature Communications* 11: 2770. https://doi.org/10.1038/s41467-020-16502-3.

Shin, S., T. K. Lee, J. M. Han, and J. Park. 2014. "Regional Effects on Chimera Formation in 454 Pyrosequenced Amplicons From a Mock Community." *Journal of Microbiology* 52: 566–573. https://doi.org/10.1007/s12275-014-3485-6.

Smith, O., G. Momber, R. Bates, et al. 2015. "Sedimentary DNA From a Submerged Site Reveals Wheat in the British Isles 8000 Years Ago." *Science* 347: 998–1001. https://doi.org/10.1126/science.1261278.

Taberlet, P., E. Coissac, F. Pompanon, C. Brochmann, and E. Willerslev. 2012. "Towards Next-Generation Biodiversity Assessment Using DNA Metabarcoding." *Molecular Ecology* 21: 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x.

Taberlet, P., E. Coissac, F. Pompanon, et al. 2007. "Power and Limitations of the Chloroplast trnL (UAA) Intron for Plant DNA Barcoding." *Nucleic Acids Research* 35: e14. https://doi.org/10.1093/nar/gkl938.

Taylor, P. G. 1996. "Reproducibility of Ancient DNA Sequences From Extinct Pleistocene Fauna." *Molecular Biology and Evolution* 13: 283–285. https://doi.org/10.1093/oxfordjournals.molbev.a025566.

van Vugt, L., S. Garcés-Pastor, E. Gobet, et al. 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.

Wang, G. C., and Y. Wang. 1997. "Frequency of Formation of Chimeric Molecules as a Consequence of PCR Coamplification of 16S rRNA Genes From Mixed Bacterial Genomes." *Applied and Environmental Microbiology* 63: 4645–4650. https://doi.org/10.1128/aem.63.12.4645-4650.1997.

Zaiko, A., P. Greenfield, C. Abbott, et al. 2021. "Towards Reproducible Metabarcoding Data: Lessons From an International Cross-Laboratory Experiment." *Molecular Ecology Resources* 22: 13485. https://doi.org/10.1111/1755-0998.13485.

Zale, R., Y.-T. Huang, C. Bigler, et al. 2018. "Growth of Plants on the Late Weichselian Ice-Sheet During Greenland Interstadial-1?" *Quaternary Science Reviews* 185: 222–229. https://doi.org/10.1016/j.quascirev.2018.02.005.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.