Treated like dirt

Robust forensic and ecological inferences from soil eDNA after challenging sample storage

Frøslev, Tobias Guldberg; Ejrnæs, Rasmus; Hansen, Anders J.; Bruun, Hans Henrik; Nielsen, Ida Broman; Ekelund, Flemming; Vestergård, Mette; Kjøller, Rasmus

Published in:
Environmental DNA

DOI:
10.1002/edn3.367

Publication date:
2023

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC-ND

Citation for published version (APA):
Treated like dirt: Robust forensic and ecological inferences from soil eDNA after challenging sample storage

Tobias Guldberg Frøslev1 | Rasmus Ejrnæs2 | Anders J. Hansen1 | Hans Henrik Bruun3 | Ida Broman Nielsen1,4 | Flemming Ekelund3 | Mette Vestergård3,5 | Rasmus Kjøller3

Abstract

Biodiversity of soil is routinely assessed with environmental DNA—most often by massive parallel sequencing of marker genes (eDNA metabarcoding). Soil biodiversity may be investigated in relation to biodiversity research or as a tool in forensic investigations. After sampling, the taxonomic composition of soil biotic communities may change. In order to minimize community changes, it is desirable to reduce biological activity, e.g., by freezing immediately after sampling. However, this may be impossible due to remoteness of study sites or, in forensic cases, where soil has been attached to an item of interest for protracted periods of time. Here, we investigated the effect of storage duration and conditions on the assessment of the soil biota with eDNA metabarcoding. We extracted eDNA from freshly collected soil samples and again from the same samples after storage under contrasting temperature conditions and contrasting exposure (open/closed tubes). We used four different primer sets targeting bacteria, fungi, protists ( cercozoans), and general eukaryotes. We quantified differences in richness, evenness, and community composition. Subsequently, we tested whether we could correctly infer habitat type and original sample identity after storage using a large reference dataset. We found stronger community composition differences with extended storage time and with higher storage temperature, and differences between open and closed tubes. However, for samples stored <28 days at a maximum of 20°C, changes were generally insignificant. Classification models successfully assigned most samples to their exact location of origin and correct habitat type even after 480 days storage. Even samples showing larger changes generally retained the original sample as the best match. For most biodiversity and forensic applications, storage of samples for days and even several weeks may thus not be a problem, if storage temperature does not exceed 20°C.

KEYWORDS

community ecology, DNA metabarcoding, microbial diversity, sample matching, sample provenancing, sample storage
1 | INTRODUCTION

A teaspoon of soil may contain more than a billion bacterial cells, meters of fungal hyphae and profuse numbers of protists, nematodes, and small arthropods (Bardgett & van der Putten, 2014). Moreover, the phylogenetic diversity of soil is stunning not only at global scale, but also at local scales (Bang-Andreasen et al., 2017; Bender et al., 2016; Harder et al., 2016; Lekberg et al., 2014). Today, high-throughput sequencing—often with the approach called eDNA metabarcoding—is the standard tool for mapping this enormous soil biodiversity. DNA metabarcoding has shown that soil microbial biodiversity varies at scales from global to local, with a strong impact of habitat (Bahram et al., 2018; Tedersoo et al., 2014). The high soil biodiversity in combination with different habitat requirements for most soil-dwelling organisms, make the biological composition of any soil sample unique, and with a compositional signature that reflects the habitat and sampling location. The continuous introduction and extinction of species to any specific site further contributes to the uniqueness of any point or snapshot sample of the soil community. This ecological fingerprint may be used for making inferences about the wider community surrounding the sampling location, of potential use in ecological studies (e.g., Bahram et al., 2018; Ejrnæs et al., 2018; Frøslev et al., 2021), as well as in forensics (e.g., Fløjgaard et al., 2019; Grantham et al., 2020).

Almost a century ago, Edmond Locard stated that a perpetrator of a crime will bring something to the crime scene, and leave with something from it (Locard, 1930). Soil is ubiquitous and has thus been of forensic interest for a long time, as it has the potential to link persons or objects to a crime scene (Bull et al., 2006). Until the advent of DNA-based methods, biotic forensic soil analyses were restricted to a relatively small proportion of the actual biotic component, and generally dependent on the skills of a few highly trained experts (Ritz et al., 2008). One of the interesting aspects of DNA metabarcoding using high-throughput sequencing is that it extends the scope to all biotic components—especially the diverse microbial component—and provides "easy" reproducible data when methods are standardized (Allwood et al., 2020; Oliveira & Amorim, 2018; Young & Linacre, 2021).

Two basic types of forensic cases can be identified—matching and provenance prediction. In cases of matching (or discrimination), the likelihood that two soil samples share the same origin in space is assessed—e.g., soil from a suspect’s shoe sole and soil from a crime scene. Here, DNA metabarcoding has a huge potential (Damaso et al., 2018; Demanèche et al., 2017; Habtom et al., 2017; Jesmok et al., 2016; Young et al., 2015, 2017, 2019). Provenance prediction can be used when no potential crime scenes have been identified. Here, the likely origin(s) of the sample of interest is narrowed down in terms of a potential geographical area or habitat/location type. Provenance prediction using soil DNA metabarcoding has so far only been explored in a single study (Fløjgaard et al., 2019), but the same overall approach also has proven useful for dust samples (Grantham et al., 2015).

For the biodiversity studies and forensic applications alike, it is important that the detected community of the sample reflects the biotic composition at the sampling site with a level of representativity adequate for the research question. For any soil sample, the final detected community will depend on its actual taxonomic composition, analytical bias and variance from the laboratory procedures, and finally the selected bioinformatic and statistical approaches. For eDNA metabarcoding, a number of sources of variance and errors relate to the last-mentioned points: e.g., DNA extraction method, PCR setup, sample tagging, library building approach, contamination, sequencing platform, and sequence processing/filtering, definition and delimitation of OTUs/ASVs (=molecular operational taxonomic units), and statistical approaches (Alberdi et al., 2018; Bálint et al., 2016; Zinger et al., 2019). These sources of variance/error mainly influence comparability between data from different studies, and to a great extent they can be controlled and standardized by the researcher. In contrast, what happens to a sample before it arrives in the lab may be less easy to control and standardize. Pre-analytical handling and storage are known to result in changes in the taxonomic composition, especially for heterotrophic microorganisms sensitive to the altered conditions (Badgley et al., 2018; Guerrieri et al., 2021; Pasternak et al., 2019).

To minimize biotic activity immediately after sampling, most sampling protocols prescribe to cool/freeze samples or add a buffer that inactivates biotic activity (Lindahl et al., 2013; Rissanen et al., 2010). In forensic applications, a soil sample recovered from an object or from a suspect has usually been removed from the crime scene for days, weeks or months and therefore has been subjected to desiccation or temperatures different from its original conditions. These “sample storage” conditions potentially change the biotic composition of the sample, which will ultimately affect the interpretation of laboratory results. Thus, it is important to establish a range of storage times and conditions that allow a valid interpretation of the different biotic components of soil samples. Pasternak et al. (2019) found that air-dried and freeze-dried soils resulted in significantly different communities after subsequent storage, whereas samples stored at −80°C remained unchanged. A study investigating the storage effect on three different soils using a small set of realistic storage scenarios for biodiversity studies concluded that the different approaches only marginally impaired the inferred richness measures and community patterns of bacteria, fungi and eukaryotes (Guerrieri et al., 2021). In their proposed guidelines, they advocated storage at 4°C for shorter periods (if possible), and otherwise desiccation of the sample with silica gel. Another study examined the forensic application of bacterial soil communities with a set of samples and locations, mimicking realistic evidence samples, and subjected samples to storage at 4 or 24°C and for different time periods (Badgley et al., 2018). They found consistent biological change with storage time and condition, but samples could still be assigned to the correct origin with supervised classification (random forest) among the studied sites. Lauber et al. (2010) also concluded that environmental factors are more important to microbial communities than differences in storage conditions of up to 14 days, and that many
samples collected and stored under field conditions can be used for community analyses. Rubin et al. (2013) found that soil bacterial community composition was significantly affected by storage times of up to 14 days and storage temperature, but that changes were nonsubstantial in a wider ecological context. It is, however, still unclear how longer storage of soils impacts basic biodiversity measures and compositional signatures of the original sampling site and habitat type in a broader ecological context, and if other soil organism groups are similarly affected as bacteria. Although soil microbial community composition changes with time (e.g., Martinová et al., 2016), season (e.g., Vorišková et al., 2014) and at small spatial scales (e.g., Schmidt et al., 2013), several studies indicate that the compositional differences at larger scales (i.e., between localities) are more systematic (e.g., Barberán et al., 2015; Damaso et al., 2018; Fournier et al., 2020; Martinović et al., 2021). Most forensic studies of soil biodiversity have used bacterial data and suggest that bacteria may be useful to identify the source of soil samples. However, bacterial communities vary at small spatial scales, and are strongly influenced by microhabitat conditions, and the inclusion of other groups of organisms as markers of geolocation may improve the accuracy (Allwood et al., 2020).

Here, we quantify the changes in taxonomic composition for a soil sample, which was divided into multiple subsamples and stored under a range of conditions, with focus on storage time, and temperature and exposure (in closed containers or in open containers allowing sample desiccation). We assessed soil biodiversity using eDNA metabarcoding targeting bacteria, but also fungi, protists (cercozoa), and general eukaryotes by use of taxon-specific primers (Figure 1). The soil sample was taken from a mature beech forest on relatively poor, not strongly leached till with a top soil, representing one of the more common natural habitats in north western Europe.

Our overall study aim was to assess the effect of sample storage on derived biodiversity metrics with a focus on biodiversity assessment and forensic applications. We approached this objective by investigating the following question:

How does sample storage affect basic biotic patterns such as richness, evenness, and taxonomic composition? And more specifically for the investigated habitat/soil type:

a. To what degree does sample storage change the signature of the sampling location and reduce the possibility of inferring the exact site of origin—i.e., sample matching?

b. To what degree does sample storage change the wider ecological signature and reduce the possibility of inferring the habitat type of the sampling site—i.e., sample provenance prediction?

We were also interested in indications of whether other markers than bacteria (fungi, protists, and eukaryotes) could be more stable location markers. Overall, we expected to see more change with longer storage time and higher storage temperature. To elucidate the effect of sample storage on basic biotic patterns, we looked for significant changes in basic biodiversity measures, such as sample richness and evenness, community dissimilarity, as well as change in taxonomic composition. To address points a and b, we applied supervised learning. We employed a reference dataset, which covers all major terrestrial habitat types in the study area (Denmark). We used k-nearest neighbors (KNNs) as a simple supervised classification approach applied directly on community dissimilarity measures, as we were interested in seeing the effect of storage on the full community, and did not aim for results directly dependent on presence/changes of particular taxa.

We expect different groups of organisms to respond differently to storage. Essentially, when soil is detached from the natural plant-soil environment, mostly saprophytic organisms, which gains there, energy from dead organic matter, will be able to keep their metabolism running for some time. Also, higher trophic levels, e.g., protists and nematodes who feeds on bacteria and fungi may to some extent live on, as long as there are available resources bottom up. In contrast, organismal groups deriving their energy directly from living plants, e.g., symbiotic mycorrhizal fungi will stop their growth soon after detachment. We, therefore, predict that the diversity of the saprotrophic communities to be more resistant than symbiotic
communities while in contrast community similarity over time to be higher for symbiotic than for saprotrophic communities. DNA derived from larger organisms as plants and animals will also like the symbiotic organisms be “fixed” in time at sampling and then just slowly fade. In addition, some soil biota forms resistant structures that will allow them to stay for prolonged times also without energy input. Soil protists such as Cercozoa form resting cysts that are very resistant to unfavorable conditions, e.g., desiccation, low prey availability, etc. Usually, the major part of the protist population in a given soil will be in the encysted stage (Ekelund et al., 2002; Ekelund & Rann, 1994). Therefore, we expect that Cercozoa will be more persistent than other eukaryotes. Bacterial taxa also vary in terms of persistence. For instance, spore-forming bacteria (Firmicutes) can persist long periods of desiccation, high-temperatures and low-substrate availability. In contrast, typically r-strategic taxa such as Pseudomonads are less persistent to unfavorable conditions and their populations fluctuate with environmental changes (Zhao et al., 2021).

2 | MATERIALS AND METHODS

2.1 | Experimental setup

For this study we use soils from one location—a deciduous forest—representing one of the more common semi-natural habitat types in northwestern Europe. The soil was sampled in a mature beech (Fagus sylvatica) forest at the Strødam nature reserve in North Zealand (Nordsjælland), Denmark on August 31, 2017. The soil had a pH of 3.9 (H$_2$O), a water content of 25% and organic matter content of 10%. Loose and coarse litter was removed from the soil surface before soil sampling, and the upper 10 cm was sampled, which then included a thin, ≤ 1 cm organic layer O and the top of the A horizon. The soil sample was taken from a single pit, about 5 L in total in the middle of a permanently marked plot (SN081) established during the Biowide project (Brunbjerg et al., 2019).

Immediately after sampling, soil was stored cold (in an ice box) and transported to the laboratory (30min drive) and sieved (5-mm mesh). 50-ml centrifuge tubes acted as experimental units and 3.2 g fresh weight of the sieved soil was added to each centrifuge tube. Tubes were then stored in combinations of temperature and exposure. The experimental setup was completed within 2–3 h after field sampling (Figure 1).

Five sets of tubes were closed with a lid to avoid desiccation and stored at 0, 5, 10, 20, and 40°C, respectively. Furthermore, two sets of tubes were left open to allow desiccation and stored at 5 and 20°C, respectively. The open tubes and temperatures were chosen to reflect room temperature and colder ambient temperature to mimic intentional desiccation (20°C) and the conditions of most forensic relevant samples (5 and 20°C). Tubes were harvested (destructively sampled) after 0 days (1 h), 1 day, 7 days (1 week) and 28 days (4 weeks) and, further for tubes incubated at 20°C, after 60, 120, 240, and 480 days (~2, 4, 8, and 16 months). All 36 treatments (experimental combinations of storage time, temperature, and exposure) were in triplicate (i.e., n = 108). Prior to storage, an 8-mm hole had been drilled into the lids and fitted with a rubber plug to allow for subsequent gas measurements. We used CO$_2$ emission over time as a measure of total biological activity in our tubes. At each harvest event, CO$_2$ emission was measured for all tubes. After the gas measurement, the harvested tubes were placed at −80°C for later DNA analyses, as this storage method is sufficient to expect no further significant change in the microbial community (Pasternak et al., 2019), while all remaining tubes were placed back at their respective incubation temperatures.

2.2 | Measuring of CO$_2$

We sampled gas from the headspace air from each of the closed sample tubes with a gas-tight syringe inserted through the rubber plug. The 0.5 ml air sample was injected into a gas chromatograph equipped with a thermal conductivity detector (Mikrolabolaborietet, Århus) for the determination of CO$_2$ concentration. Gases were separated before detection on a 1.8-m Haysep Q column operated at 45°C. During each CO$_2$ measuring event, we measured the CO$_2$ concentration of atmospheric air and CO$_2$ standards as appropriate.

2.3 | Sequence data

2.3.1 | DNA extraction

The DNA was extracted from 107 soil samples (originally 108, but one tube—a replicate of “20°C open 480 days”—was empty upon harvest, reason unknown) in batches of up to 23 samples. Sample tubes were shaken by hand before subsampling for DNA extraction. From each sample, 0.25 g of soil was subjected to DNA extraction using DNeasy PowerSoil Kit (Qiagen), following the manufacturer’s protocol, except for the elution step, where 105 μl 1 x TET-buffer was used. For contamination control, an extraction blank was included per batch was included. Prior to extraction, the samples were homogenized using a TissueLyser II (Qiagen) at 30 Hz for 10 min. DNA concentrations were measured with Qubit dsDNA HS (High Sensitivity) Assay Kit (Invitrogen) and samples were normalized to a concentration of 1 ng/μl prior to PCR amplification. The samples were not randomized among batches, which pose a potential confounding factor, but there were no appreciable batch effects in the data.

2.3.2 | DNA amplification and sequencing

DNA was amplified using four different markers targeting bacteria, fungi, protists (Cercozoa), and general eukaryotes, respectively (see Table S1 for primer and PCR information). The reason to include both a general eukaryote marker and specifically address fungi and
Bioinformatic steps followed the general procedures of earlier studies (Frøslev et al., 2017, 2019) with minor modifications. Demultiplexing of samples was done with a custom script that keeps R1 and R2 separate for DADA2 processing, and is based on Cutadapt (Martin, 2011) searching for a sequence pattern matching the full length combined tag and primer allowing for errors, and removing possible remnants of the other primer at the 3' end. We used DADA2 (v 1.8) to identify amplicon sequence variants (ASVs) and to remove chimeras (bismeras). For highly length variable markers (ITS2 for fungi), the script included a sliding window truncation of sequences from the 5' end with Sickle (Joshi & Fass, 2011) (with options: pe -l 50 -q 28 -x -f -t sanger) to maximize output and quality of the ITS2 sequences that have length variation and therefore large differences in the onset of the quality drop towards the 3' end. For the other markers where amplicon length is homogeneous, we applied a fixed length cutoff of the 5' end that allowed for ample overlap between R1 and R2 reads. Sequences were filtered and matched between R1 and R2 reads with DADA2 (using fastqPairedFilter with options maxN = 0, maxEE = 2, truncQ = 2, matchIDs = TRUE).

2.4.2 | Taxonomic assignment

For taxonomic assignment of ASVs, we used several different approaches. The bacterial dataset was assigned using the assignTaxonomy command in dada2 using the dada2-formatted version (DOI: 10.5281/zenodo.1172782) of the Silva Project’s version 132 database (Quast et al., 2013). The fungal, protist and general eukaryotes datasets were each matched against reference databases using vs-ear (Rognes et al., 2016) and a custom script that uses the top 10 matches to assign a majority rule taxonomy. Assignment of the fungal dataset was annotated by matching the ASVs against the UNITE database for fungi (v. 8.3) (Abarenkov et al., 2021) and annotation of the protist and eukaryote datasets was done by matching against the PR2 database (Guillou et al., 2013). For the analyses, where we combined the study datasets with the reference datasets, we only assigned taxonomy to the fungal data in order to exclude non-target sequences.

Some biodiversity studies work with focal taxonomic lineages and/or only with ASVs that can be reliably annotated at species level and are thus limited by the quality and completeness of reference databases for annotating taxonomic identity. Other approaches are based on, e.g., community dissimilarity metrics and do not as such need taxonomic annotation of the ASVs, and forensic application would ideally utilize all data produced by a primer set to maximize reproducibility. In this study, we only removed non-target sequences from the fungal dataset before downstream analyses, as these primers amplify a substantial amount of non-target (plant) sequences.

2.5 | Statistical analyses

For the all analyses relying on ASV tables, the relevant table was resampled to the 25th percentile to get even sequencing depth (but allowing a minor part of the samples to have lower read counts). To verify that sequencing depth was adequate to interpret ASV richness from the resampled data, rarefactions curves were plotted as well as plots of unmodified ASV richness against ASV richness of resampled data (Figure S7). Data was Hellinger transformed (square root on relative abundances) to downweight the impact of ASVs with high read numbers. As community dissimilarity measure, Bray-Curtis was used. Non-metric multidimensional scaling (NMDS) was done using the settings k = 2, try = 500, trymax = 4000 (using functions rarefy, decostand, vegdist and metaMDS from vegan package; Oksanen et al., 2007). We chose to show the ordination in two dimensions to be able to show all data in one graph, but only if stress
was sufficiently low (<0.2). All the statistical analyses were run in R version 4.0.3 (2020-10-10).

2.5.1 | Absolute sample change from time zero with storage

Data from time zero samples (n = 21, i.e., three from each of the seven combinations of temperature and exposure) were used for the time zero population (reference) when analyzing effects of storage with time. To address changes in richness and diversity with storage, ASV richness was used as a proxy for total taxonomic/species richness. Following the findings of Guerrieri et al. (2021), we also measured the change in richness of dominant ASVs (ASVs registered with ≥1% of the reads in each sample in the resampled tables). Pilou's evenness index was used as a measure of evenness/diversity. To address change in community composition, we calculated the Bray–Curtis dissimilarity between any stored sample and the centroid of all time zero samples. The centroid of time zero was calculated with the dist_to_centroid function (usedist package). For each particular treatment set (i.e., combination of temperature, exposure, storage time), we assessed significant changes in richness, evenness, and community composition compared with the time zero communities, using t-tests with Bonferroni correction for multiple tests (i.e., 29 tests, excluding time zero combinations). Significant differences in community composition (compared to time 0) were also assessed with pairwise PERMANOVA as implemented in the function pairwise.adonis (Arbizu, 2020) with the argument "reduce" to compare only against time 0. We considered p-values of <0.05 as significant. Community change was visualized with NMDS ordination. Furthermore, we evaluated the taxonomic composition for each treatment group, combining the ASVs of all the triplicates per treatment.

2.5.2 | Relative change—Habitat signature and forensic application

Despite of significant absolute changes in biodiversity metrics for stored samples, the change might still be insignificant for several applications, as the sample may have retained its signature—in terms of biological composition—of the exact sampling location or at least of the habitat type in a broader context. Therefore, to address the relative stability of the biotic signal of the stored samples, and thus, the forensic utility and robustness of biodiversity measures, the stored samples were analyzed together with a reference dataset. The reference dataset stems from and contains sequence data from 130 40m×40m plots across Denmark. The 130 plots represent major gradients of moisture, fertility, and succession, and thus include representatives of most natural to semi-natural terrestrial habitat types in Denmark, as well as some agricultural and silvicultural land-use types. Soil samples from the reference dataset were collected and processed like the samples in this study, except that each of the 130 samples were constructed from a bulk sample of 81 smaller samples, that the soil was thoroughly mechanically homogenized (potentially releasing more intracellular DNA), that 4g of soil was used for the DNA extraction, and that the soil was sampled 3 years earlier in 2014 (November–December). The bulk sample used for the storage samples in this study was taken in the middle of one (SN081) of the plots used for the reference dataset, and this plot was excluded from those analyses where it could bias the interpretation.

Sequence data (ASV tables) from the present study and the reference dataset were combined for each of the four organism groups. Taxonomy was only assigned for the fungal data, to allow for exclusion of non-target sequences. For these combined analyses, we discarded ASVs with <10 reads in the reference dataset, and thereby excluded ASVs unique to the stored samples, that could otherwise make these samples more similar due to unique ASVs in that dataset. For supervised classification, we used k-nearest neighbor analysis (KNN) on community dissimilarity measures (the Bray–Curtis dissimilarity of Hellinger transformed ASV tables), by evaluating the classes (habitat type or location) of the KNNs (the k samples in the training dataset with the lowest Bray–Curtis dissimilarity measure to the investigated stored sample). This was performed with a custom r-script.

**Signature of exact location**

Using KNN, we investigated to which degree the stored samples retained characteristics of the exact location where they were collected, in the context of our reference dataset of terrestrial habitats in Denmark. Data from the reference dataset acted as outgroup. To avoid inflating classification success, we used only nine of the 21 time zero samples (triplicates of 0°C closed and 5°C closed and open) as ingroup. The soil used for the storage samples was sampled in the middle of one of the plots (SN081) from the reference dataset, so this sample could reasonably have been coded as ingroup. We chose, however, to exclude it from the models to not impose any biases.

We calculated the proportion of ingroup and outgroup samples among the seven nearest neighbors of the 129 reference plots and nine time zero samples) as the classification probability. As a direct visualization of the relative dissimilarities underlying this approach, we calculated and plotted a dissimilarity ratio for each stored sample in the form of the Bray–Curtis dissimilarities between the stored sample and the time zero centroid, compared with the dissimilarity between stored sample and each of the 129 reference plots.

**Signature of habitat type**

Using a similar approach as above, we investigated to which degree the stored samples retained characteristics of the broader habitat type, to which the un-stored original sample was assigned. We used the survey dataset of 36,323 observations of 5464 species (of vascular plants, bryophytes, macrofungi, lichens, and insects) recorded across the 130 reference sites (Brunbjerg et al., 2019) to define nine strata (from hereon: habitat types), eight natural types and one agricultural. These habitat types were defined by supervised
classification (see Data S1) and encompassed the following: Mor forest (acidic forest), Mull forest (alkaline forest), Bog forest, Swamp forest, Heathland, Grassland, Moor (acidic wetland), Fen (alkaline wetland), and Agriculture.

We then calculated the proportion of different natural strata among the nine nearest neighbors (of the 129 reference plots, excluding the reference sample from the sampling site of the stored samples, as well as all stored samples) as the classification probability. For comparison, we also calculated the classification probability for the original (SN081) reference sample from the sampling site. We established the variance of the ingroup stratum by calculating the Bray–Curtis dissimilarity of the cluster members to the habitat cluster centroid. Subsequently, the dissimilarity of stored samples to habitat centroids was related to the said variance in order to assess probability of correct habitat type assignment.

3 | RESULTS

3.1 Absolute sample change from time zero with storage

Overall, measures of richness, evenness and community composition were relatively stable for all organism groups and systems ≤20°C for up to 28 days (Figure 2a–d, Table S2), whereas measures diverged gradually for most systems stored at 40°C or stored at 20°C for 28 days or more. Generally, richness started to decrease after 1 to 28 days while evenness was more stable except for a few treatments. Community compositional dissimilarity to time zero started to increase at days 1 to 28. The concentration of DNA extracted was decreasing for 40°C samples from day 1, and for 20°C samples stored for more than 28 days (Figure 2e, Table S4). CO₂ emission per hour increased with storage temperature and decreased gradually with time (Figure 2f, Table S4).

Richness (Figure 2): Generally, we found relatively large variation in richness estimates, and thus relatively few changes with time were significant though the trends were common for most taxa/treatment comparisons. The pattern for total richness and richness of dominant ASVs, were similar. Bacteria had the highest number of significant differences (20°C open and closed at 60 days or more, all p < 0.022, and 40°C at 7 days or more, all p < 0.0005), whereas eukaryotes was the group with fewest significant differences (40°C 28 days, p = 0.009, and 20°C (open) at 240 days or more, p < 0.0005, and 5°C (open) at 7 days, p = 0.002). Despite the lack of significance, the downward trend was evident for all 20°C samples stored for a long time, seemingly with a difference between open and closed tubes for fungi, protists, and eukaryotes, where the open tubes showed a faster and more pronounced decrease in richness.

Evenness (Figure 2): The evenness of bacterial communities did not change significantly with time, although the figure shows a clear declining trend for 40°C (and partly 20°C closed) samples. For the other groups there were some significant differences, but generally, evenness was relatively stable with time. The protist data showed a marked difference for 20°C samples, where only the closed systems saw a mainly significant drop in evenness from day 28 (p = 0.396 to p < 0.0005).

Divergence from time 0 (Figure 2): All treatments gradually showed increased the Bray–Curtis dissimilarity to time 0 community composition (the calculated centroid), with 40°C (and partly 20°C) samples increasing faster and more. All 40°C samples showed a clear and significant trend, being significantly more dissimilar from time zero already after 1 day (all p < 0.001). For most 20°C samples, divergence from time zero was apparent to significant from day 28 (all p > 0.151 to p < 2e⁻¹⁶). For bacteria and fungi, the closed 20°C tubes changed faster and more than the corresponding open tubes, whereas protists showed the opposite pattern. For protists and bacteria, the long term stored 20°C samples changed as much or more than the 40°C (28 day) samples.

Community change (Figure 3): In the NMDs ordinations (all with stress <0.12) of the communities, the samples stored at 0, 5, and 10°C for up to 28 days, displayed no systematic change, reflecting the low level of change observed in the other metrics. However, for the samples stored at 20°C we observed a systematic change from day 28 and onwards, with open and closed tubes clearly showing different trajectories (least evident for fungi). The 40°C samples showed a clearly changed position already after 1 day of storage, and a different trajectory compared to the 20°C samples. For samples exhibiting evident change (20°C for 28 or more days and 40°C), the change was of a 31214308-manner (although the protists displayed some variation in the 20°C open samples at day 240 and 480). The pairwise PERMANOVA (Table S3) had low (<0.1) to significant (<0.05) p-values for most samples stored at 20°C for 28 days or more (but also for 1 day/open), and for 40°C after 1 day, corresponding well with the NMDS and divergence from time 0.

Taxonomic changes (Figure 4, all the taxonomic levels can be seen in Figures S1–S4). Bacteria: For most treatments of 20°C or lower, few major taxonomic changes occurred up to day 28. However, pronounced taxonomic changes took place in the 40°C samples where the Firmicute genus Alicyclobacillus increased to finally dominate the samples after 28 days. In the 20°C samples, gradual change in the proportions of several taxa was observed from day 60, and there was a clear difference between the open and closed tubes. The Firmicute genus Bacillus increased markedly after 120 days in the open tubes, whereas the closed tubes saw a corresponding increase of the Acidobacteria Acidipila. Fungi: For all treatments of 20°C or lower, Mortierellomycetes (Mortierella) systematically increased, whereas Agaricomycetes (Inocybe, Cortinarius, etc.) concomitantly decreased already after 7 days. In the 40°C samples, Aspergillus dominated already after 7 days. Protists: The taxonomic change of protists was less pronounced, but with a few systematic changes. The 20°C (open and closed tubes) displayed an increase of Allaspidae after 120 days, whereas, in the closed tubes, only Cryomonadida (Rhogostoma lineage) increased from day 7, and decreased again at day 240. Whereas the other organism groups displayed a drastically different taxonomic
FIGURE 2 Absolute change with sample storage. Rows show (from top to bottom) (a) Bacteria, (b) Fungi, (c) Protists and (d) Eukaryotes, and bottom row shows, (e) Changes in measured DNA concentrations and (f) Measured CO$_2$ emissions. Plots (a–d) show, column 1: Change in ASV richness; column 2: Community evenness as Pilou's evenness index; column 3: Community change as Bray–Curtis dissimilarity from the centroid of the time zero communities. Plots show mean value ± SEM for triplicates per treatment, with storage time on the x-axis, colors indicate storage temperature, and shape indicate exposure. Corresponding p-values for significant differences can be seen in Table S2.
composition in 40°C samples, the taxonomic composition in the 40°C samples of the protists was comparable to that of samples at lower temperatures. Eukaryotes: For most treatments of 20°C or lower at 28 days or less, few systematic taxonomic changes occurred. For 20°C closed there was a decline of metazoan (mainly in the form of Enoplea nematodes) and an increase of fungi. This was also the case (but less linear) in the 20°C open tubes. This was also seen very clearly for the 40°C, where the Metazoa ASVs disappeared at day 7, and where an increase of Apicomplexa was also seen.

### 3.2 | Signature of exact location

We used supervised KNN classification to test if the stored samples could be reclassified to the correct location (sampling site) as
represented by nine unstored (time zero) samples using a 129 sample reference dataset as outgroup. Using a criterion of 0.5 mean probability, the approach classified all stored samples correctly (Figure 5a–d). The dissimilarity ratio—defined as the Bray–Curtis dissimilarity of a stored sample to any of the 129 reference plots divided by the Bray–Curtis dissimilarity to time zero centroid of the stored samples—became smaller with storage and temperature (Figure 6), but the ratio never dropped below one. Thus, no sample changed to become more similar to other localities than to the origin. Figure S5 shows how the absolute dissimilarity of stored samples to any of the 129 samples from the reference data show a steady increase for 40°C and long term (from day 28) storage 20°C samples.

3.3 | Signature of habitat type

Using a supervised classification of the set of 130 reference sites into nine broadly circumscribed habitat types (see Data S1), we used the KNN classification to examine to which habitat type the stored samples were assigned. For all the datasets, the dominant habitat
type for unstored samples was Mor forest, followed by Mull forest. This assignment fitted well with the ecological properties of the focal soil sampling site (SN081), which is mature beech forest on relatively poor, but not strongly leached, till with a top-soil pH of 3.9. This slightly ambiguous classification as acidic Mor forest border-line to alkaline Mull forest was seen even at time zero (Figure 7), and also for the original reference sample from the focal site (SN081 indicated as “Ref” in Figure 7). Using a criterion of 0.5 mean probability, Figure 7a–d shows that this approach classified all stored samples correctly, except for bacteria at 40°C after 7 and 28 days, and at 20°C open after 240 and 480 days. The probability of correct assignment was generally constant with storage time and with temperature. Apparent decline in assignment success was only seen for the for 40°C samples and for the 20°C bacteria samples. Although the KNN approach was overall successful, Figure 7e–h shows that several stored samples exceeded the dissimilarity to the habitat type centroid by more than two SDs (for the reference data members of the habitat type). This was most evident for the protists although this group showed the highest and most stable mean classification success. Figure S6 shows the classification probabilities towards the second most probable habitat type (Mull forest).

4 | DISCUSSION

It is paramount in eDNA metabarcoding studies, that a sample adequately represents the community from which it was drawn. Ideally, it should be comparable to an immediately processed sample, and only show deviations corresponding to what one would expect from the chosen analytical workflow. However, immediate processing is often not possible, and taxonomic compositional changes of the sample may occur. This is particularly the case in forensic sampling where storage conditions of soil traces are beyond the control of the analyst, and for which storage under suboptimal conditions and for an unknown period of time is the norm. Depending on the objective of the study, some degree of community change during sample storage may be acceptable, but the uncertainties related to storage conditions are crucial to understand.

In this study, we addressed community change across a combination of storage conditions and periods. As expected, we found temperature-dependent community changes during storage time. However, we also observed that changes in measures of richness and evenness, and changes in community structure and taxonomic composition, were small for storage temperatures of 10°C (to 20°C) or lower and storage times of 28 days or less. However, biodiversity measures and community patterns diverged gradually for treatments at 40°C already after 1 day and for samples stored at 20°C for 28 days or more. Despite significant taxonomic compositional changes, we could still refer most samples to original habitat type and exact location with supervised classification models.

It is important to note that our results only relate to storage of soil from one location and habitat type, i.e., a specific temperate beech forest. Although this soil type is likely to behave similar to soils from a wide set of more common habitats with comparable soil characteristics, e.g., city parks, coniferous forests and grasslands, we can only cautiously extrapolate to general rules and recommendations. This is supported by other studies of soil bacterial communities from other localities and habitats showing insignificant changes with shorter storage times at or below room temperature for and lower for soils from other localities and habitats (Guerrieri et al., 2021; Lauber et al., 2010; Rubin et al., 2013). Thus, it seems safe to conclude that for studies of major biodiversity patterns, soil samples can be collected and stored for shorter time periods (days) without the need of immediate freezing/cooling, as long as it is possible to store samples at 20°C or lower. This will often be possible in temperate regions, and is practical if no lab facilities are nearby and/or if working with bulk samples that must be transported for further processing, etc.

**FIGURE 5** Supervised classification of exact location using k-nearest neighbor analysis (KNN). Probability of stored samples being classified as belonging to the exact sampling site, using nine time-zero stored samples as ingroup and 129 samples representing a wide selection of terrestrial habitats in Denmark as outgroup. Classification probability was calculated as the proportion of ingroup samples among the seven closest neighbors—defined as those samples with the smallest Bray–Curtis dissimilarity to the examined stored sample. Cells show the mean value of the triplicate per treatment. Panels show (from left to right): (a) Bacteria, (b) Fungi, (c) Protists and (d) Eukaryotes.
FIGURE 6 Dissimilarity ratio of stored samples to reference data compared to time zero. Each point shows the ratio between the Bray–Curtis dissimilarity of a stored sample to one of the 129 reference plots divided by the Bray–Curtis dissimilarity to time zero centroid of the stored samples. Thus, for each set of replicates for each treatment, the plot has 129 points. x-axis and color indicate storage time, symbol indicates exposure (open vs. closed tubes), and faceting corresponds to storage temperature and exposure. If a point is below the dotted line, it means that the stored sample is more similar to a reference plot than to the time zero centroid. This is not the case for any comparisons, meaning that the stored samples all retain highest similarity (lowest dissimilarity) to the time zero centroid compared to all the reference plots.
The taxonomic changes over time and with increasing temperature mainly confirmed our expectations. Spore-forming Firmicute genera as *Bacillus* and *Alicyclobacillus* increased with time and for fungi saprotrophic Mortierellomycetes increased, whereas ectomycorrhizal genera such as *Inocybe* and *Cortinarius* decreased. Still, targeting of certain fast-growing taxa, e.g., molds such as *Mortierella* requires special consideration. On the other hand, our results show that higher temperatures (40°C) induce relatively early changes in taxonomic composition, as well as significant changes in other biodiversity measures already after 1 day, and this effect is very likely to also be relevant for most other soil types. Hence, work in the tropics needs special attention when there is no access to cooling. Desiccation is a good approach to conserve DNA (Alsos et al., 2020; Guerrieri et al., 2021) and has also been used in practice for soil DNA studies with a global scope (Tedersoo et al., 2014). In this study, we only investigated passive desiccation in the form of open 20°C (and 5°C) tubes, which clearly differed from their closed 20°C counterparts, and the divergence between closed and open treatments continued until the last sampling time. Whether this continuous change was due to differential growth of species present from the start, or partially from influx of new species to the open tubes is not clear. We expect that active desiccation with, e.g., silica gel followed by storage in closed containers may be the best approach, when cooling is not possible, as suggested by another study (Guerrieri et al., 2021). However, it is important to consider which further analyses samples are to be subjected to as other soil properties may need other...
storage regimes; e.g., Rhymes et al. (2021) investigated the effect of storage time and temperature on biochemical properties of soil and concluded that refrigeration is better than freezing.

Previous studies of biotic changes with sample storage have focused on bacteria. Here, we extended the focus to fungi, protists, and eukaryotes in general. The general effects on the four different organism groups were very similar, but there were some marked differences. As the only group, protists showed a lower response across all measures for 40°C storage than for long term storage at 20°C. A characteristic of a good marker for sample origin is low stochasticity related to sampling. Our time zero samples can be perceived as sampling replicates, and the dissimilarity of these unstored samples to the time zero centroid as a measure of stochasticity. This measure was lowest (around 0.25) for bacteria and fungi, whereas protists and eukaryotes showed values around 0.35. Another characteristic of a good location marker is the ratio between the dissimilarity to other sites and the origin. Figure 6 shows that this ratio generally was well above 1, except for some of the bacterial samples stored at 20°C for 240 days or more or at 40°C for 28 days. Among the other three groups, fungi showed the highest overall ratio, persistently higher than 1, except for 40°C, where protists had a higher ratio. This also translates into data from all four groups resulting in almost perfect reclassification success to the original location (Figure 5) and no decrease in habitat classification success with storage (Figure 7), except for bacteria. All in all, this indicates that fungi may be the best marker of location and habitat of the four organism groups examined here, supplemented by protists for samples stored at 40°C. Together with the results of Young et al. (2014) who found that fungi were better than plants, bacteria and eukaryotes in separating soils from two localities, this supports the proposition by Allwood et al. (2020), that other markers than bacteria may be more productive for geolocation.

In this study, we combined the stored samples with a reference dataset of 130 samples representing most major terrestrial habitat types in Denmark, including a sample from the same study site as the stored samples. Despite several differences in the sampling strategies of the two datasets, we could classify the stored sample to the correct habitat type using supervised classification in the form of simple KNN models.

The most important forensic lessons from this study, is that no stored sample gained higher similarity to any other sample after storage. Thus, all the samples retained higher similarity to the original unstored sample than to any sample from the reference dataset. Hence, the KNN models depending on compositional similarity could correctly match all stored samples to the correct exact location.

In forensic matching of samples—i.e., comparison of a trace sample to a crime scene—it is not permissible to get false matches, which may potentially lead to conviction of innocent persons. Thus, when employing community compositional approaches like this study, it is important to consider the strengths and weaknesses of the analytical approaches. The KNN approach uses dissimilarities to known observations, so if case evidence samples are merely investigated in the context of few other observations—or with observations from entirely different habitat types—false positives are likely. We suggest that real life forensic cases should not exclusively rely on approaches like KNN based on the closest match, but also consider whether the observed dissimilarities lie within or close to the variation seen for replicated samples from the same location, and ideally be combined with a score-based likelihood ratio-like measure. The matching approach applied here depends on a representative sampling with several replicates of the reference site. In the case of matching two trace samples, or one trace sample to single references from different localities, other approaches than KNN are needed.

Contrary to models for matching of forensic samples, models for provenance prediction will most likely only be used as an investigative tool in forensic cases—e.g., to narrow down areas of interest—and thus, some flexibility of models may be allowed as avoidance of false positive predictions is less critical. Here, we tested whether the stored soil samples could correctly be classified to a wider habitat type of the location where they were collected. The KNN models for all organism groups were very successful and only failed for 40°C after 7 and 28 days, and 20°C open after 240 and 480 days for bacteria. The success of such approaches for stored samples is likely to be influenced by the starting material, and extrapolations from the present study should be done with caution, as we examined only one type of soil.

For the real-life forensic applications, we recommend prioritizing large representative ecological reference databases—i.e., sequence data from soils of a wide selection of habitat types—to reduce uncertainty in ecological inferences and site matching. Further studies are needed to test if such ecological reference database should be based on single bulk samples constructed for maximal representation of larger localities as used in this study, or one based on several replicates of smaller soil samples representing smaller or larger localities. Along with this, other sources of variation, e.g., season, need to be addressed in future studies. It is also important to note that forensic investigations are most often a multi-disciplinary approach, where evidence from many different investigations feed into a unified conclusion. The eDNA based methods are still new, and need to be supported by other, more well-established methods for soil discrimination and/or other evidence in the specific case. We find it likely that non-DNA based, physicochemical analyses may be less impacted by storage, whereas eDNA may show resolution at other spatial scales.

The changes we see in the stored samples are systematic—i.e., the replicates change in the same direction, as also detected in another study (Foran & Badgley, 2020). Furthermore, we see that the direction of community change depends on temperature and exposure (open/closed tube). It may thus be possible to predict storage condition and time for a sample of interest. This could be a valuable approach for forensic samples, where time since removal from the original site may be of interest, parallel to the estimation of postmortem interval. We also see clear differences in the taxonomic composition related to temperature and exposure, and it may be possible and interesting to identify indicator taxa for storage conditions. On the other side of
this coin, it may also be possible to identify and extract those taxa that are least sensitive to storage and use these to build provenance and matching models that less impacted by sample storage.

This study shows that soil samples retain a large proportion of the original taxonomic compositional signature during relatively extended storage, and that the observed deviation—although deterministic—does not exceed the variance between replicated un-stored samples, if they are not stored warm or for a very long time. Still, this source of variation in biodiversity patterns from soil eDNA metabarcoding needs to be compared to other sources like seasonality, samples size, etc., to inform sampling strategies for biodiversity studies as well as making a solid foundation for interpretation of forensic analyses. In conclusion, this is important for both the forensic applications and biodiversity studies as it indicates that for most applications, a relaxed approach to sample storage during sampling can be exercised.

**AUTHOR CONTRIBUTIONS**

RK, TGF, FE, MV, AJH, and RE conceived the overall study. RK carried out the sampling and designed the storage setup. IBN did the lab work. TGF, RK, and RE decided on the analytical approach. TGF did the bioinformatics, statistical analyses, and plots. TGF and RK wrote the first draft. All the authors contributed to the finalizing of the manuscript.

**ACKNOWLEDGMENTS**

This is a contribution to a national research project to develop soil forensic methods supported by Innovation Fund Denmark Grand Solutions (grant no. 6151-000028, https://innovationsfonden.dk/), and a contribution to a project on optimizing strategies for eDNA based biodiversity surveys, DNAmark, supported by Aage V. Jensen Foundation. The funders did not participate in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Jens H. Petersen is thanked for designing three of the organism group icons.

**CONFLICT OF INTEREST**

The authors declare no conflict of interest.

**DATA AVAILABILITY STATEMENT**

Raw non-demultiplexed sequence data is available from Dryad (https://doi.org/10.5061/dryad.k0p2ngfbs) (Frøslev et al., 2022). Sample-wise demultiplexed paired-end reads are available in the ENA project PRJEB56039 (https://www.ebi.ac.uk/ena/browser/view/PRJEB56039) (also see Data S1 for specific information on reads in ENA). All further data and scripts necessary to reproduce this study is available in the form of an r-project on GitHub (https://github.com/tobiasgf/sample_storage).

**ORCID**

Tobias Guldberg Frøslev https://orcid.org/0000-0002-3530-013X

**REFERENCES**


Arbizu, P. M. (2020). pairwiseAdonis: Pairwise multilevel comparison using Adonis (version 0.4) [R package]. https://github.com/pmartinezarbizu/pairwiseAdonis


**SUPPORTING INFORMATION**

Additional supporting information can be found online in the Supporting Information section at the end of this article.

---