



HAL
open science

Effects of soil preservation for biodiversity monitoring using environmental DNA

Alessia Guerrieri, Aurélie Bonin, Tamara Münkemüller, Ludovic Gielly, Wilfried Thuiller, Gentile Francesco Ficetola

► **To cite this version:**

Alessia Guerrieri, Aurélie Bonin, Tamara Münkemüller, Ludovic Gielly, Wilfried Thuiller, et al.. Effects of soil preservation for biodiversity monitoring using environmental DNA. *Molecular Ecology*, 2021, 30, pp.3313-3325. 10.1111/mec.15674 . hal-03013928

HAL Id: hal-03013928

<https://hal.univ-grenoble-alpes.fr/hal-03013928>

Submitted on 19 Nov 2020

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

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

1 **Effects of soil preservation for biodiversity monitoring using environmental DNA**

2 Running title: Preserving soil for eDNA analyses

3

4 Alessia Guerrieri^{1, *}, Aurélie Bonin¹, Tamara Münkemüller², Ludovic Gielly², Wilfried Thuiller², Gentile

5 Francesco Ficetola^{1, 2}

6

7 Affiliations:

8 ¹Department of Environmental Science and Policy, Università degli Studi di Milano, Via Celoria 26, 20133

9 Milano, Italy

10 ²Univ. Grenoble Alpes, CNRS, Univ. Savoie Mont Blanc, LECA, Laboratoire d'Ecologie Alpine, F-38000,

11 Grenoble, France

12 *Correspondence author. Email: alessia.guerrieri@unimi.it

13 **ABSTRACT**

14 Environmental DNA metabarcoding is becoming a key tool for biodiversity monitoring over large
15 geographical or taxonomic scales and for elusive taxa like soil organisms. Increasing sample sizes and
16 interest in remote or extreme areas often require the preservation of soil samples and thus deviations from
17 optimal standardized protocols. However, we still ignore the impact of different methods of soil sample
18 preservation on the results of metabarcoding studies and there is no guideline for best practices so far.
19 Here, we assessed the impact of four methods of soil sample preservation that can be conveniently used
20 also in metabarcoding studies targeting remote or difficult to access areas. Tested methods include:
21 preservation at room temperature for 6h, preservation at 4°C for three days, desiccation immediately after
22 sampling and preservation for 21 days, and desiccation after 6h at room temperature and preservation for
23 21 days. For each preservation method, we benchmarked resulting estimates of taxon diversity and
24 community composition of three different taxonomic groups (bacteria, fungi and eukaryotes) in three
25 different habitats (forest, river bank and grassland) against results obtained under ideal conditions (i.e.
26 extraction of eDNA right after sampling). Overall, the different preservation methods only marginally
27 impaired results and only under certain conditions. When rare taxa were considered, we detected small but
28 significant changes in MOTU richness of bacteria, fungi and eukaryotes across treatments, but MOTUs
29 richness was similar across preservation methods if rare taxa were not considered. All the approaches were
30 able to identify differences in community structure among habitats, and the communities retrieved using
31 the different preservation conditions were extremely similar. We propose guidelines on the selection of the
32 optimal soil sample preservation conditions for metabarcoding studies, depending on the practical
33 constraints, costs and ultimate research goals.

34

35 **KEYWORDS:** eDNA metabarcoding, sample storage, MOTU richness, α and β diversity, microbial
36 communities, eukaryotes

37

38 INTRODUCTION

39 Environmental DNA (hereafter referred to as "eDNA") can be defined as the mixture of complex, sometimes
40 degraded, DNA that microorganisms (e.g. bacteria and fungi) or macroorganisms (e.g. animals, plants) leave
41 behind in their environment (i.e. soil, water, sediments, etc.). By studying short, taxonomically-informative
42 DNA fragments obtained from eDNA samples, it is possible to identify the associated taxa and therefore to
43 survey biodiversity. Coined as "eDNA metabarcoding", this approach has revolutionized several branches of
44 ecology and environmental sciences during the last decade, by providing relatively quick and non-invasive
45 assessments of present or past biodiversity of animals, plants and microorganisms (Taberlet, Bonin, Zinger,
46 & Coissac, 2018). Metabarcoding is particularly valuable for monitoring biodiversity over large geographical
47 or taxonomic scales (De Vargas et al., 2015; Delgado-Baquerizo et al., 2018; Zinger et al., 2019b).
48 Furthermore, it gives access to biodiversity components that are elusive to conventional survey methods.
49 For instance, it allows the rapid assessment of microbial soil biodiversity, which is extremely complex, time-
50 consuming and imperfect when using direct observations, culturing techniques or microscopy (Giovannoni,
51 Britschgi, Moyer, & Field, 1990; Ward, Weller, & Bateson, 1990).

52 Metabarcoding relies on a succession of several steps: 1) sampling; 2) preservation of the collected
53 material until lab processing; 3) DNA extraction; 4) PCR amplification of a selected genomic region; 5) high-
54 throughput sequencing of amplicons; and 6) analysis of sequences using bioinformatics and statistical tools
55 (Zinger, Bonin, et al., 2019). Each step is critical to obtain robust taxonomic inventories and diversity
56 estimates, and an increasing number of studies has assessed how methodological choices across the
57 different steps could influence the conclusions of a study (Calderón-Sanou, Münkemüller, Boyer, Zinger, &
58 Thuiller, 2020; Cantera et al., 2019; Chen & Ficetola, 2020; Nichols et al., 2018; Taberlet et al., 2018).

59 Despite this growing body of literature, so far little attention has been devoted to the effect of different
60 preservation conditions of the collected environmental material before lab processing (i.e. step 2). We thus
61 know little about the optimal storage conditions of the collected material, and how long samples can be
62 stored to limit biases in taxonomic inventories.

63 Some recent studies have analyzed the preservation of sampling material obtained from water (see
64 e.g. Kumar, Eble, & Gaither, 2020; Majaneva et al., 2018). Conversely, methodological analyses on the
65 effects of sample preservation are largely scarce for soil, perhaps because the majority of metabarcoding
66 studies have so far been performed in temperate areas where access to lab facilities is often easy
67 (Hoffmann, Schubert, & Calvignac-Spencer, 2016; Huerlimann et al., 2020). In such cases, sample
68 preservation is sometimes not necessary at all, or at least not over long periods of time. However, one
69 great promise of metabarcoding is its potential for providing biodiversity data for remote areas, where
70 biodiversity monitoring is essential but difficult. When sampling in remote or inaccessible areas (e.g.
71 tropical and arctic areas; mountain chains), samples are rarely collected nearby lab facilities and an
72 immediate *in situ* DNA extraction is generally not possible due to logistic constraints (but see Zinger,
73 Taberlet, et al., 2019b for a notable exception). More generally, with the ever-increasing number of
74 samples analyzed during a typical metabarcoding study, sample preservation is more and more
75 indispensable, and the time lag between sample collection and subsequent molecular processing makes it
76 particularly relevant to understand the impact of sample preservation, and to identify preservation
77 strategies that do not bias the conclusions of studies.

78 In an ideal metabarcoding study, communities recovered from preserved samples should match
79 those retrieved if samples had been processed immediately after sampling. However, inappropriate
80 preservation conditions can cause both DNA degradation and the proliferation of certain taxonomic groups
81 with respect to others, before DNA extraction (Cardona et al., 2012; Orchard, Standish, Nicol, Dickie, &
82 Ryan, 2017). This can in turn affect taxa detection and also the relative contributions of different taxonomic
83 groups to the estimated biodiversity. A recent review suggested that the majority of eDNA metabarcoding
84 studies does not provide accurate information about sample treatment before processing (Dickie et al.,
85 2018). Almost half of the studies do not report how samples were stored and conserved, and 30% of them
86 store samples at 0-4°C, and thus at a temperature where many bacteria and fungi continue to be active and
87 potentially affecting the whole sample. About 15% of the studies stored samples in a range of 5-35°C, which

88 can be considered a poor practice when no preservatives are added (Dickie et al., 2018), and only 10%
89 stored them below 0°C (Dickie et al., 2018).

90 So far, the consequences of preservation practices and the resulting deviations from immediate
91 processing and analyses have rarely been studied quantitatively. Some studies suggested that liquid-based
92 preservatives have limited usefulness for soil eDNA (Tatangelo et al., 2014). Yet, Lauber, Zhou, Gordon,
93 Knight, & Fierer (2010) tested the effect of storing samples from soil, human gut and skin at different
94 temperatures and did not detect any significant effect on bacterial communities, while Orchard et al.
95 (2017) found that storage time and temperature can affect colonization by arbuscular mycorrhizal fungi,
96 with subsequent impacts on the reconstruction of communities. Differences between these studies may be
97 due to their different protocols. However, they also focused on different taxonomic groups, which may
98 react differently to storage period and temperature. Consequently, these studies are difficult to compare,
99 highlighting the importance of formal assessments of preservation methods. Desiccation is a further
100 approach that can efficiently conserve high-quality DNA for genomic studies (e.g. Alsos et al., 2020; Chase &
101 Hills, 1991). Although not widely used for metabarcoding samples, desiccation through silica gel has good
102 potential for soil, as it allows removing >25% of its weight in water in a few hours (P. Taberlet, pers.
103 communication), is cost-effective, easy to transport, and is not an issue for aircraft transportation (no
104 flammable or dangerous preservatives). A clear understanding of the effect of different preservation
105 methods, especially across various groups of taxa, is thus pivotal for a robust application of eDNA
106 metabarcoding to biodiversity monitoring in general, and that of remote areas in particular.

107 Here, using eDNA metabarcoding of different taxonomic groups in soil systems, we tested: (i) how
108 preservation methods influence overall richness estimates and what the role of rarely observed taxa is; (ii)
109 how preservation methods influence identified community structure and its turnover between different
110 habitats; and (iii) what the best practices are under limited laboratory access. More specifically, we first
111 selected three soil preservation methods (room temperature, 4°C, desiccation in silica gel) because they are
112 commonly used in the literature (room temperature and 4°C) or because they are easy to implement in the
113 field (desiccation and room temperature). Then, we assessed the impact of these preservation methods

114 applied to different durations in order to mimic logistic constraints (see Supplementary Material, Appendix
115 A for details on experimental design), and compared the communities obtained with those observed in
116 ideal conditions, i.e. when eDNA is extracted immediately after sampling (within less than one hour). We
117 examined bacterial, fungal and eukaryotic communities to cover a broad taxonomic range, since different
118 taxa can be differentially affected by sample preservation conditions (Cardona et al., 2012; Orchard et al.,
119 2017).

120 MATERIALS AND METHODS

121 Soil preservation and experimental treatments

122 In April 2019, we collected soil samples from three habitats: a grassland (N 45.194° E 5.776°), a broadleaved
123 forest (N 45.196° E 5.774°), and a vegetated river bank (N 45.195° E 5.780°). The study design was
124 optimized to allow DNA extraction immediately after sampling, which hampered using distant study sites.
125 All sites were within 400 m from the Laboratoire d'Écologie Alpine (LECA) in Grenoble, France. We choose
126 habitats with some differentiation to allow different communities but not too extreme and relatively close
127 together so that we expect some overlap between communities. This mimics what is commonly done in the
128 field when gradients are sampled. We established five plots within each habitat; the minimum distance
129 between nearby plots was about 20 m. Within each plot, we collected four soil samples (with a minimum
130 distance of one meter) at a depth of 0–20 cm and then pooled them together, for a total of five pooled
131 samples per habitat (approx. 200 g each pooled sample). Soil litter was not included in the samples. Pooled
132 samples (15 in total) were homogenized; subsequently, from each of them we took five subsamples of 15 g
133 of soil (total: 75 subsamples; Fig. 1).

134 The five soil subsamples of each pooled sample were subjected to five different treatments: 1) eDNA was
135 extracted immediately after sampling (within 1 h; treatment hereafter referred to as “control”); 2) samples
136 were preserved at room temperature (21–23°C) and eDNA was extracted 6 h after sampling; 3) samples
137 were inserted in sterile 50-mL falcon tubes and preserved at 4°C. eDNA was extracted three days after
138 sampling; 4) samples were inserted in hermetic, sterile boxes with 20 g of silica gel immediately after
139 sampling, then stored at room temperature, and eDNA was extracted 21 days after sampling; 5) samples
140 were inserted in hermetic, sterile boxes with 20 g of silica gel 6h after sampling, then stored at room
141 temperature, and eDNA was extracted 21 days after sampling.

142 We did not test full-factorial combinations of different preservation periods and conditions, which
143 was not feasible in terms of time and costs. We did not consider freezing or storage in liquid nitrogen,
144 which is unrealistic when dealing with large numbers and / or volumes of samples, as is the case for large-
145 scale metabarcoding studies. Furthermore, freezing is generally impossible in remote areas (Dickie et al.,

146 2018), where maintaining a cold chain cannot be ensured given the logistical challenges and this is often
147 replaced by preservation in a cool box (4°C). Previous studies showed that preservative solutions have
148 limited usefulness for soil or sediment samples ((Rissanen et al., 2010; Tatangelo et al., 2014).

149

150 **Molecular analyses**

151 For all sample treatments, eDNA extraction was performed in a dedicated room using the NucleoSpin® Soil
152 Mini Kit (Macherey-Nagel, Germany), after a preliminary step where 15 g of soil were mixed with 20 ml of
153 phosphate buffer for 15 minutes as described in (Taberlet, Coissac, Pompanon, Brochmann, & Willerslev,
154 2012); and with a final elution in 150 µl. We also included one extraction negative control per treatment.

155 Environmental DNA of bacteria, fungi and eukaryotes was amplified using primers designed for
156 markers Bact02 (Taberlet et al., 2018), Fung02 (Epp et al., 2012; Taberlet et al., 2018) and Euka02
157 (Guardiola et al., 2015), respectively. Bact02 and Fung02 amplify fragments of about 220-250 bp, while
158 Euka02 generally amplifies fragments <150 bp. The three markers are well suited for metabarcoding
159 analyses, as all have a very low number of mismatches in the priming region across target organisms, and
160 the relatively short length of amplified fragments allows their use with potentially degraded DNA (Taberlet
161 et al., 2018). To allow bioinformatic discrimination of PCR replicates after sequencing, eight-nucleotide long
162 tags were added on the 5' end of both forward and reverse primers, so that each PCR replicate was
163 represented by a unique combination of forward and reverse tags. Tags had at least five nucleotide
164 differences among them (Coissac, 2012). Samples were randomized on a 96-well plate, along with the five
165 extraction controls, eight bioinformatic blanks, six PCR negative controls and two PCR positive controls. PCR
166 positive controls were included to check for potential cross-contaminations and to monitor amplification
167 and sequencing performances. The positive control was a 1:10 dilution of the ZymoBIOMICS™ Microbial
168 Community DNA Standard II (Zymo Research, USA) constituted of genomic DNA of eight bacterial and two
169 fungal strains (i.e., *Pseudomonas aeruginosa*, *Escherichia coli*, *Salmonella enterica*, *Lactobacillus*
170 *fermentum*, *Enterococcus faecalis*, *Staphylococcus aureus*, *Listeria monocytogenes*, *Bacillus subtilis*,
171 *Saccharomyces cerevisiae*, *Cryptococcus neoformans*) at known concentrations.

172 In order to avoid over-amplification of template DNA and to limit chimera formation, we determined the
173 optimal number of amplification cycles and DNA extract dilution using qPCR. The qPCR assay was
174 conducted on 48 randomly selected samples, using 1 μ l of 1:1000 diluted SYBR[®] Green I nucleic acid gel
175 stain (Invitrogen[™], USA), with a real-time PCR thermal cycler set to standard mode. qPCR was performed
176 for both 1:10 diluted and undiluted template eDNA.

177 For Bact02 and Fung02, PCR reactions were performed on 1:10 diluted template DNA, using 32 and 44
178 cycles respectively. For Euka02, we performed 34 cycles on undiluted DNA. All PCR reactions consisted of
179 10 μ l of AmpliTaq Gold 360 Master Mix 2X (Applied Biosystems[™], Foster City, CA, USA), 2 μ l of primers mix
180 at initial concentration of 5 μ M of each primer, 0.16 μ l of Bovine Serum Albumin (corresponding to 3.2 μ g;
181 Roche Diagnostic, Basel, Switzerland) and 2 μ l of DNA extract, for a final volume of 20 μ l. The PCR profiles
182 had an initial step of 10 min at 95°C, followed by several cycles of a 30 s denaturation at 94°C, a 30 s
183 annealing at 53°C (Bact02), 56°C (Fung02) or 45°C (Euka02), and a 90 s elongation for Bact02 and Fung02,
184 or a 60 s elongation for Euka02 at 72°C, followed by a final elongation at 72°C for 7 minutes. The
185 amplification was performed in 384-well plates, with four replicates for each sample. After amplification,
186 PCR products of the same marker were pooled together in equal volumes and a 5- μ l aliquot of the pooled
187 amplicons was visualized by high-resolution capillary electrophoresis (QIAxcel Advanced System, QIAGEN,
188 GERMANY) to verify the expected fragments length and to monitor primer dimers. Pooled amplicons were
189 purified using the MinElute PCR Purification Kit (QIAGEN, GERMANY) following the manufacturer's
190 protocol. Six subsamples of the pool of amplicons were purified separately for each marker, and then
191 combined again before being sent for library preparation and sequencing to Fasteris (SA, Geneva,
192 Switzerland). One library was prepared per marker using the MetaFast protocol (Taberlet et al., 2018) and
193 then sequenced using the MiSeq (Fung02 and Bact02) or HiSeq 2500 (Euka02) platforms (Illumina, San
194 Diego, CA, USA) with a paired-end approach (2 \times 250 bp for Fung02 and Bact02, and 2 \times 150 bp for Euka02).

195

196 **Bioinformatic treatment**

197 The bioinformatic treatment of sequence data was performed using the OBITools software suite (Boyer et
198 al., 2016). First, forward and reverse reads were assembled using the *illumina-paired-end* program, keeping
199 only sequences with an alignment score higher than 40. Aligned sequences were assigned to the
200 corresponding PCR replicate using the program *ngsfilter*, by allowing two and zero mismatches on primers
201 and tags, respectively. After sequence dereplication using *obiuniq*, bad-quality sequences (i.e. containing
202 “N”), sequences whose length fell outside the expected size interval (below 45 bp for Bact02, below 68 bp
203 Fung02 and below 36 bp for Euka02) and singletons were filtered out. The *obiclean* program was run to
204 detect potential PCR or sequencing errors with the *-r* option set at 0.5: in a PCR reaction, sequences are
205 tagged as “heads” when they are at least twice as abundant as other related sequences differing by one
206 base. Only the sequences tagged as “heads” in at least one PCR were kept.

207 Taxonomic assignment was conducted using the *ecotag* program based on a reference database
208 constructed from EMBL (version 136) by running the *ecoPCR* program (Ficetola et al., 2010). More
209 specifically, *ecoPCR* carried out an *in silico* PCR with the primer pair used for the experiment and allowing
210 three mismatches per primer. The obtained reference databases were further curated by keeping only the
211 sequences assigned at least at the family level.

212 Further data filtering was performed in R version 3.6.1 (R Core Team, 2018) to remove spurious
213 sequences that can bias ecological conclusions drawn from DNA metabarcoding data (Calderón-Sanou et
214 al., 2020). More specifically, we discarded all MOTUs with best identity <85% (Fung02, Bact02) or <80%
215 (Euka02). These MOTUs were indeed rare (31%, 1.7% and 7.3% of reads for Fung02, Bact02 and Euka02,
216 respectively) and their effect on the NMDS was marginal, as observed in other studies (e.g. Botnen et al.,
217 2018). Furthermore, we removed MOTUS with less than five occurrences in the overall dataset, detected in
218 more than one extraction or PCR negative control (Zinger, Bonin, et al., 2019a), or that were detected in
219 less than two PCR replicates of the same sample, as they often represent false positives (Ficetola et al.,
220 2015).

221

222 **Statistical analyses**

223 For all taxonomic groups, we used Generalized Linear Mixed Models (GLMMs) to test if the different
224 treatments lead to differences in the observed MOTU richness. In GLMMs, the number of MOTUs per
225 sample was calculated and used as a dependent effect, the five treatments were used as predictors, and
226 sample identity was used as a random factor. The model was performed with the generalized poisson
227 distribution error using the R package *glmmTMB* (Brooks et al., 2017), in order to consider overdispersion
228 (Consul & Famoye, 1992). If GLMM detected significant differences among treatments, we used treatment
229 contrasts to test if each treatment led to communities significantly different from those unraveled by the
230 “control” condition. Treatment contrasts are standard non-orthogonal contrasts, in which each category
231 (treatment) is compared to a user-defined reference category, and are appropriate to compare multiple
232 treatments against one single control category (in this case, immediate extraction; (Field, Miles, & Field,
233 2015). The uncorrected number of MOTUs tends to overestimate the actual taxonomic richness (Calderón-
234 Sanou et al., 2020). Therefore, we repeated this analysis twice: considering all the observed MOTUs, and
235 considering only MOTUs with frequency $\geq 1\%$ in each sample (hereafter referred to as "common MOTUs").

236 Subsequently, we used multivariate analyses to assess the variation of bacteria, fungi and
237 eukaryotic communities across habitats and treatments. Before running multivariate analyses, we
238 calculated the proportion of reads of each MOTU in each sample. Relative abundance values were then
239 transformed using the Box-Cox transformation, which simultaneously solves the double-zero problem and
240 improves the multivariate normality of data (Legendre & Borcard, 2018).

241 First, we used Nonmetric MultiDimensional Scaling (NMDS) to describe differences in communities
242 among the three habitats, and check whether different treatments yield different interpretations of
243 ecological relationships among samples. NMDS uses an optimization process to find a configuration of
244 points (samples) in a space with a small number of dimensions, and is suitable for metabarcoding analyses
245 that aim to reconstruct variation in community composition as well as possible, without preserving any
246 particular distance measure among objects (Borcard, Gillet, & Legendre, 2011; Chen & Ficetola, 2020; Paliy
247 & Shankar, 2016). Given its robustness and flexibility, NMDS is often used as the first step to characterize
248 the similarity of communities in metabarcoding studies (Chen & Ficetola, 2020; Paliy & Shankar, 2016).

249 NMDS was run on the Euclidean distance computed on Box–Cox-chord-transformed data (Legendre &
250 Borcard, 2018), by building 1,000 ordinations.

251 Second, we used *ProcMod*, a Procrustes-based analysis (Coissac & Gonindard-Melodelima, 2019),
252 to measure the multivariate correlations between the communities obtained using the different
253 treatments. *ProcMod* can be used to measure the shared variation between matrices, and is particularly
254 appropriate to test relationships between datasets obtained through DNA metabarcoding and
255 metagenomics (Coissac & Gonindard-Melodelima, 2019). Procrustes analyses tend to overfit the data,
256 therefore we used a modified version of Procrustes correlation that is robust to highly-dimensional data
257 and allows a correct estimation of the shared variation between data sets (Coissac & Gonindard-
258 Melodelima, 2019). The Procrustes-based correlation tests were performed using the *corls* function in the R
259 package *ProcMod*, using 1,000 randomizations to test the mean covariance between random matrices
260 (Coissac & Gonindard-Melodelima, 2019).

261 Third, we used redundancy analysis (RDA) to measure the amount of variation among communities
262 that is explained by differences in habitat and treatments (Legendre & Legendre, 2012; Ter Braak, 1986).
263 With habitat typology and treatment as constraining matrices, we used treatment contrasts to test if each
264 treatment led to communities significantly different from those unraveled by the control treatment. Thus,
265 significant treatment contrasts indicate that results between control and experimental treatments differ in
266 an important way, while non-significant results mean that deviation from ideal conditions is not specifically
267 pronounced. Significance of RDA and treatment contrasts was tested through 10,000 permutations using
268 the *vegan* package in R (Borcard et al., 2011; Oksanen et al., 2019).

269 For bacteria only, RDA detected significant differences between the control and some of the
270 treatments. We thus ran a similarity percentage analysis with the *simper* R function (Clarke, 1993) from
271 *vegan* to identify the taxa contributing to the overall pairwise treatment difference (Geyer et al., 2014).
272 Significance was tested using 50,000 permutations. Given the large number of tests performed, the
273 significance of tests was corrected using the False Discovery Rate (FDR) method with the *fdrtool* package
274 (Strimmer, 2008). FDR has greater power than traditional approaches (e.g. Bonferroni correction) when

275 performing multiple comparisons (Benjamini & Hochberg, 1995). All statistical analyses were performed in
276 the R environment.
277

278 RESULTS

279 A total of 6.3, 7.9 and 25.7 million reads were obtained from the Bact02, Fung02 and Euka02 libraries,
280 respectively. After read assembly, quality filtering, spurious sequence and contaminant removal, 481,411;
281 2,511,721 and 13,232,441 good-quality sequences remained, consisted of 660 (Bact02), 1,075 (Fung02) and
282 3,611 (Euka02) unique sequences (i.e. MOTUs).

283

284 Differences in MOTU richness among treatments

285 Generalized Linear Mixed Models allowed identifying shifts in the richness of observed MOTUs. When we
286 considered all the detected MOTUs, GLMM detected significant differences in MOTUs richness among
287 treatments for all the markers considered (Bact02: $\chi_4 = 38.9$, $P < 0.001$; Fung02: $\chi_4 = 18.2$, $P = 0.001$;
288 Euka02: $\chi_4 = 21.7$, $P < 0.001$; Fig. 2). Compared to the control, contrasts showed small but significant
289 changes in MOTUs richness under the 4°C treatment (Bact02: $z = 2.54$, $P = 0.010$; Fung02: $z = -2.17$, $P =$
290 0.029 ; Euka02: $z = 2.65$, $P = 0.008$), the silicagel treatment (Bact02: $z = -2.93$, $P = 0.003$; Fung02: $z = -3.99$, P
291 < 0.001 ; Euka02: $z = 3.92$, $P < 0.001$), and the silicagel+6h treatment (Bact02: $z = -3.74$; Fung02: $z = -4.02$;
292 Euka02: $z = 4.18$; all $P < 0.001$). The 6h treatment caused a small but significant decrease in MOTUs richness
293 for fungi ($z = -2.42$; $P = 0.015$), but not for bacteria and eukaryotes ($P = 0.456$, $P = 0.283$, respectively; for all
294 contrasts: Table S1).

295 Nevertheless, when we repeated analyses by excluding MOTUs with a frequency $< 1\%$, differences
296 in richness were much smaller, and were only significant for bacteria and fungi (Bact02: $\chi_4 = 9.69$, $P = 0.045$;
297 Fung02: $\chi_4 = 14.1$, $P = 0.006$; Euka02: $\chi_4 = 2.22$, $P = 0.693$; Fig. 2). Compared to the control, MOTUs richness
298 decreases for Bact02 under the 4°C treatment ($z = -2.91$; $P = 0.003$) and increases for Fung02 under the two
299 silicagel treatments ($z = 2.77$; $P = 0.005$; $z = 1.75$; $P = 0.080$; respectively), while no significant effect was
300 detected for Euka02 under any of the treatments (all $P > 0.170$; for all contrasts: Table S1).

301 Habitat caused a significant effect in MOTUs richness only for Fung02 both before and after
302 removing rare MOTUs (before: $\chi_1 = 11.8$, $P < 0.001$; after: $\chi_1 = 20.5$, $P < 0.001$).

303

304 **Ecological similarity of communities among treatments**

305 Nonmetric MultiDimensional Scaling showed a stress value of 0.13 for Bact02, 0.14 for Fung02 and 0.12 for
306 Euka02. For each of the three markers, the NMDS plots obtained for the five sample preservation
307 treatments were extremely similar, and the ecological differences among the three habitats were clearly
308 identified by all the preservation treatments (Fig. 3).

309 The multivariate correlation between the communities obtained with the five treatments was
310 always very strong (Procrustes-modified correlation: for all comparisons between “control” and treatments
311 $r \geq 0.84$, $P < 0.0001$; Fig. 4) indicating, for all markers, that most of the variation of retrieved communities
312 was shared across all the treatments. Procrustes correlations were particularly high for Fung02 and Euka02
313 (all $r \geq 0.9$), and between the control and the treatments 6h and 4°C (all $r \geq 0.93$; Fig. 4).

314

315 **Differences between the obtained communities**

316 Redundancy analysis allowed us to measure the amount of variation explained by differences among
317 habitats and by treatments. Overall, 33%, 24%, and 33% of variability was explained by differences in
318 habitat for bacteria, fungi, and eukaryotes, respectively. The community differences among habitats were
319 strongly significant for the three taxonomic groups (permutation test: all $P \leq 0.001$). Differences among
320 treatments were much weaker, and explained 9%, 2% and 2% of variation only for bacteria, fungi and
321 eukaryotes, respectively. Differences between treatments were significant for bacteria (permutation test: P
322 < 0.0001), but not for fungi and eukaryotes (both $P = 1$).

323 For bacteria, contrasts did not detect significant differences between control and the 6h or 4°C
324 treatments. Differences between control and the two silicagel treatments were significant but explained a
325 limited amount of variation (for both treatments, $\approx 3\%$ of variation explained; $P < 0.0001$; Table 1). We thus
326 used similarity percentage analysis to identify the MOTUs significantly contributing to these differences.
327 Only one single MOTU showed a significant contribution ($P = 0.03$ after FDR correction) to the differences
328 between control and silicagel treatment; this MOTU (belonging to the Bacteroidetes phylum) showed a

329 very limited frequency under the silicagel treatment (Fig. S1). After FDR correction, no MOTU showed a
330 significant contribution to the differences between control and the silicagel+6h treatment.
331

332 **DISCUSSION**

333 Monitoring soil biodiversity with eDNA metabarcoding over large geographical and taxonomic scales and
334 sometimes in remote places **is increasingly** important in ecological research. Understanding how
335 preservation conditions affect estimates of taxonomic richness and community composition is essential to
336 ensure sound conclusions. Our study shows that soil metabarcoding results are surprisingly robust to
337 preservation conditions, as we observed limited differences in community structure and diversity estimates
338 when samples were preserved using different strategies. However, some taxonomic groups and diversity
339 components are more sensitive than others to certain preservation conditions. This allowed **us to develop**
340 **guidelines** for preservation depending on the aims of monitoring programs and on focal taxa.

341 The aim of this study was comparing realistic approaches to soil preservation against an ideal
342 situation. Immediate extraction was our reference approach, as it avoids both DNA degradation (i.e.
343 potential under-representation of certain taxa) and continued growth of certain taxonomic groups (i.e.
344 potential over-representation of other taxa). Unfortunately, immediate extraction is only possible if
345 sampling occurs nearby facilities, or when a mobile eDNA laboratory is available (e.g. Zinger, Taberlet, et al.
346 2019b), and logistical constraints often hamper its application in remote areas. We selected preservation
347 conditions among the most achievable, cost-effective and frequent practices to sampling soil for eDNA
348 studies (Dickie et al., 2018), **although additional storage methods (e.g. liquid nitrogen, dry ice, RNA later)**
349 **are available.**

350

351 **Influence of preservation methods on richness estimates**

352 Preservation methods generated some small but significant differences in MOTUs richness compared to
353 what is observed in the “control”, with some contrasting effects across taxa. When considering all the
354 MOTUs, none of the preservation conditions yielded estimates of alpha-diversity identical to the “control”.
355 For instance, just six hours at room temperature caused a significant decrease of MOTUs richness in fungi.
356 It has been shown that estimates of alpha-diversity using metabarcoding are extremely sensitive to
357 methodological choices (Calderón-Sanou et al., 2020). Our study underlines that even preservation for a

358 very short time can affect the detection of rare MOTUs and highlights the sensitivity of fungi to
359 preservation at room temperature (Delavaux, Bever, Karppinen, & Bainard, 2020). MOTUs richness of all
360 the taxa was also affected by preservation at 4°C, which caused a slight increase of MOTUs richness for
361 bacteria and eukaryotes, and a slight decrease for fungi. The effect of temperature and time storage in
362 fungal and bacterial growth **has already been proven** (see e.g. Orchard et al., 2017; Pettersson & Bååth,
363 2003). Despite this, in addition to temperature, we can expect that other parameters such as initial soil
364 moisture and pH influence bacterial growth (Bååth & Arnebrant, 1994; Drenovsky, Vo, Graham, & Scow,
365 2004; Fernández-Calviño & Bååth, 2010; Kaiser et al., 2016) with a combined effect. Finally, drought affects
366 the richness of microbial communities in soil ecosystems with differential effects across taxa depending on
367 their ecology (Evans, Wallenstein, & Burke, 2014; Meisner, Jacquiod, Snoek, Ten Hooven, & van der Putten,
368 2018; Ochoa-Hueso et al., 2018), and three weeks of preservation with silica gel generally reduced the
369 observed MOTUs richness in bacteria and fungi, while it increased the richness of eukaryotes.

370 However, our study also shows that specific caution is mostly necessary when rare MOTUs are of
371 interest. The exclusion of rare **and uncommon** MOTUs strongly reduced differences between optimal
372 conditions and different preservation. The **remaining effects were much weaker for bacteria and fungi,**
373 **while disappeared for eukaryotes (Fig. 2), and can be due to the heterogeneous growth across taxonomic**
374 **groups, or to differential DNA degradation under different preservation conditions.** This suggests that the
375 effect of preservation approach on taxonomic richness mostly occurs on rare species, as already suggested
376 for microbial communities (Meisner et al., 2018). Several authors have shown that eDNA metabarcoding
377 does not represent the best tool for the detection of rare MOTUs, as some rare MOTUs remain undetected,
378 while many sequences detected at rare frequency are **artifacts** (Brown et al., 2015). Estimates of α -diversity
379 should therefore **always be taken** with caution, and indices that underweight rare MOTUs (e.g. Shannon or
380 Simpson diversity) can provide more robust estimates (Brown et al., 2015; Calderón-Sanou et al., 2020;
381 Bálint et al., 2016).

382

383 **Differences in community structure**

384 If the study interest is in community structure and not in richness estimates, then preservation choices
385 become even less important. In fact, the similarity of communities obtained through the different
386 preservation conditions is surprisingly high (see Procrustes correlation coefficients; Fig. 4); the amount of
387 variation explained by preservation conditions was much lower than the observed differences among
388 habitats (see redundancy analysis), and multivariate analyses consistently allowed to detect community
389 differences among habitats (Fig. 3). In other words, metabarcoding is able to identify the ecological
390 differences among sites, independently of the preservation approach. Even though metabarcoding analyses
391 are sensitive to methodological choices, estimates of relationships between diversity and the environment
392 are often robust (Calderón-Sanou et al., 2020; Ji et al., 2013), and this is a very good news if we want to
393 apply these approaches to broad-scale monitoring programs, aiming at assessing the effects of
394 environmental changes.

395 Bacteria were the only taxon for which we detected significant differences between the “control”
396 and the preservation conditions, with ≈3% of variability explained by differences between the “control” and
397 the desiccation treatments. The observed differences most likely refer to some taxa that are affected by
398 the dry conditions and could lead to an overrepresentation of some taxa that are more resistant under
399 these conditions. We expect some taxa to better survive and grow in dry treatments with respect to others,
400 and this would make their DNA more available in the samples. Differences between the desiccation
401 treatments and the control were small, and only one out of 660 MOTUs showed a significant variation in
402 abundance with the control. This MOTU (belonging to the Bacteroidetes phylum of bacteria, see Supporting
403 Information) was generally abundant in the control and preservation conditions 2 and 3 (average frequency
404 of reads around 10%) while it drastically decreased under preservation conditions 4 and 5 (Fig. S1). This
405 agrees with studies showing that different genera belonging to this phylum respond differently to drought
406 (Meisner et al. 2018). In fact, the Bacteroidetes *Flavobacterium* shows an increase in abundance over dried-
407 treatments, even though differences after three weeks were not significant (see Fig. S2).

408 The significant differences observed for some taxa and preservation conditions stress the
409 importance of selecting the preservation method before starting a monitoring program and using it

410 consistently through the whole monitoring, to avoid confusion between the effects of methods and of
411 environmental changes.

412

413 **Conclusions: guidelines for optimizing preservation conditions**

414 Standardized protocols are essential for repeatable and reliable biodiversity monitoring, and our results
415 allow to propose guidelines to improve and standardize the preservation of soil samples for eDNA
416 metabarcoding analyses (Fig. 5):

- 417 1) If sampling occurs close to lab facilities, or a mobile lab is available, extracting DNA as soon as possible
418 is the best approach. Storing samples a few hours at room temperature does not have major impact
419 on the outcome of analyses, especially if the focus is not on rare MOTUs;
- 420 2) If lab facilities are available after a short-time transportation, storing samples in the fridge (0-4°C) for a
421 few days is a safe approach as it does not have a significant impact on community composition, and
422 only moderately affects MOTUs richness. However, this approach can be problematic if the aim is to
423 estimate MOTUs richness, and particularly the occurrence of rare MOTUs. The feasibility of this
424 strategy also depends on the number and volume of samples, and to the possibility of maintaining the
425 cold chain;
- 426 3) If monitoring in remote areas, sample desiccation (e.g. using silica gel) and long-term preservation at
427 room temperature is a reasonable approach, and it is particularly convenient when working with a
428 large volume of samples. This approach preserves ecological signal, but can affect the detection of
429 some taxa, particularly among the rarest ones. Therefore, this approach is suboptimal for monitoring
430 programs aiming at detecting rare MOTUs.

431 An effective application of eDNA metabarcoding to biodiversity monitoring is complex, and protocols of
432 sample preservation are key methodological choices that have to be taken into account when designing a
433 metabarcoding-based monitoring. When working in difficult and remote environments researchers are
434 faced with the trade-offs between a faithful representation of biodiversity, and multiple logistic constraints
435 in the field. Accurate a-priori planning is often the basis of successful monitoring programs and our

436 guidelines can help researchers and practitioners to identify the best approach to sample preservation,
437 depending on the studied taxa and research goals.
438

439 **ACKNOWLEDGMENTS**

440 We thank P. Taberlet for constructive discussions. This study is funded by the European Research Council
441 under the European Community's Horizon 2020 Programme, Grant Agreement no. 772284
442 (IceCommunities). WT, TM and LG also thank the rom the French Agence Nationale de la Recherche (ANR)
443 through the GlobNets (ANR-16-CE02-0009) project and from 'Investissement d'Avenir' grant managed by
444 the ANR (Montane: OSUG@2020: ANR-10-LAB-56).
445

446 **REFERENCES**

447

- 448 Alsos, I. G., Lavergne, S., Merkel, M. K. F., Boleda, M., Lammers, Y., Alberti, A., Pouchon, C., Deneud, F.,
449 Pitelkova, I., Puşcaş, M., Roquet, C., Hurdu, B. I., Thuiller, W., Zimmermann, N. E., Hollingsworth, P.
450 M., & Coissac, E. (2020). The treasure vault can be opened: Large-scale genome skimming works well
451 using herbarium and silica gel dried material. *Plants*, *9*(4), 432.
452 <https://doi.org/10.3390/plants9040432>
- 453 Botnen, S. S., Davey, M. L., Halvorsen, R., & Kausrud, H. (2018). Sequence clustering threshold has little
454 effect on the recovery of microbial community structure. *Molecular Ecology Resources*, *18*(5), 1064–
455 1076. <https://doi.org/10.1111/1755-0998.12894>
- 456 Chase, M. W., & Hills, H. H. (1991). Silica gel: An ideal material for field preservation of leaf samples for DNA
457 studies. *TAXON*, *40*(2), 215–220. <https://doi.org/10.2307/1222975>
- 458 Rissanen, A. J., Kurhela, E., Aho, T., Oittinen, T., & Tirola, M. (2010). Storage of environmental samples for
459 guaranteeing nucleic acid yields for molecular microbiological studies. *Applied Microbiology and*
460 *Biotechnology*, *88*(4), 977–984. <https://doi.org/10.1007/s00253-010-2838-2>
- 461 Tatangelo, V., Franzetti, A., Gandolfi, I., Bestetti, G., & Ambrosini, R. (2014). Effect of preservation method
462 on the assessment of bacterial community structure in soil and water samples. *FEMS Microbiology*
463 *Letters*, *356*(1), 32–38. <https://doi.org/10.1111/1574-6968.12475>
- 464 Bååth, E., & Arnebrant, K. (1994). Growth rate and response of bacterial communities to pH in limed and
465 ash treated forest soils. *Soil Biology and Biochemistry*, *26*(8), 995–1001. doi: 10.1016/0038-
466 0717(94)90114-7
- 467 Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful
468 Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1),
469 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
- 470 Borcard, D., Gillet, F., & Legendre, P. (2011). Numerical Ecology with R. In *Numerical Ecology with R*. doi:
471 10.1007/978-1-4419-7976-6
- 472 Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: A unix-inspired
473 software package for DNA metabarcoding. *Molecular Ecology Resources*, *16*(1), 176–182. doi:
474 10.1111/1755-0998.12428
- 475 Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., ... Bolker, B. M.
476 (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear
477 mixed modeling. *R Journal*, *9*(2), 378–400. doi: 10.32614/rj-2017-066
- 478 Brown, S. P., Veach, A. M., Rigdon-Huss, A. R., Grond, K., Lickteig, S. K., Lothamer, K., ... Jumpponen, A.
479 (2015). Scraping the bottom of the barrel: Are rare high throughput sequences artifacts? *Fungal*
480 *Ecology*, *13*, 221–225. doi: 10.1016/j.funeco.2014.08.006
- 481 Calderón-Sanou, I., Münkemüller, T., Boyer, F., Zinger, L., & Thuiller, W. (2020). From environmental DNA
482 sequences to ecological conclusions: How strong is the influence of methodological choices? *Journal*
483 *of Biogeography*, *47*(1), 193–206. doi: 10.1111/jbi.13681
- 484 Cantera, I., Cilleros, K., Valentini, A., Cerdan, A., Dejean, T., Iribar, A., ... Brosse, S. (2019). Optimizing
485 environmental DNA sampling effort for fish inventories in tropical streams and rivers. *Scientific*
486 *Reports*, *9*(1), 1–11. doi: 10.1038/s41598-019-39399-5
- 487 Cardona, S., Eck, A., Cassellas, M., Gallart, M., Alastrue, C., Dore, J., ... Manichanh, C. (2012). Storage
488 conditions of intestinal microbiota matter in metagenomic analysis. *BMC Microbiology*, *12*. doi:
489 10.1186/1471-2180-12-158
- 490 Chase, M. W., & Hills, H. H. (1991). Silica gel: An ideal material for field preservation of leaf samples for DNA
491 studies. *TAXON*, *40*(2), 215–220. doi: 10.2307/1222975
- 492 Chen, W., & Ficetola, G. F. (2020). Numerical methods for sedimentary-ancient-DNA-based study on past
493 biodiversity and ecosystem functioning. *Environmental DNA*, edn3.79. doi: 10.1002/edn3.79
- 494 CLARKE, K. R. (1993). Non-parametric multivariate analyses of changes in community structure. *Austral*
495 *Ecology*, *18*(1), 117–143. doi: 10.1111/j.1442-9993.1993.tb00438.x
- 496 Coissac, E. (2012). OligoTag: A program for designing sets of tags for next-generation sequencing of
497 multiplexed samples. *Methods in Molecular Biology*, *888*, 13–31. doi: 10.1007/978-1-61779-870-2_2

- 498 Coissac, E., & Gonindard-Melodelima, C. (2019). Assessing the shared variation among high-dimensional
 499 data matrices: a modified version of the Procrustean correlation coefficient. *BioRxiv Bioinformatics*, 1–
 500 16. doi: 10.1101/842070
- 501 Consul, P. C., & Famoye, F. (1992). Generalized poisson regression model. *Communications in Statistics -
 502 Theory and Methods*, 21(1), 89–109. doi: 10.1080/03610929208830766
- 503 De Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., ... Velayoudon, D. (2015). Eukaryotic
 504 plankton diversity in the sunlit ocean. *Science*, 348(6237). doi: 10.1126/science.1261605
- 505 Delavaux, C. S., Bever, J. D., Karppinen, E. M., & Bainard, L. D. (2020). Keeping it cool: Soil sample cold pack
 506 storage and DNA shipment up to 1 month does not impact metabarcoding results. *Ecology and
 507 Evolution*, (November 2019), 1–13. doi: 10.1002/ece3.6219
- 508 Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-González, A., Eldridge, D. J., Bardgett, R. D.,
 509 ... Fierer, N. (2018). A global atlas of the dominant bacteria found in soil. *Science*, 359(6373), 320–325.
 510 doi: 10.1126/science.aap9516
- 511 Dickie, I. A., Boyer, S., Buckley, H. L., Duncan, R. P., Gardner, P. P., Hogg, I. D., ... Weaver, L. (2018). Towards
 512 robust and repeatable sampling methods in eDNA-based studies. *Molecular Ecology Resources*, 18(5),
 513 940–952. doi: 10.1111/1755-0998.12907
- 514 Drenovsky, R. E., Vo, D., Graham, K. J., & Scow, K. M. (2004). Soil water content and organic carbon
 515 availability are major determinants of soil microbial community composition. *Microbial Ecology*, 48(3),
 516 424–430. doi: 10.1007/s00248-003-1063-2
- 517 Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., ... Brochmann, C. (2012). New
 518 environmental metabarcodes for analysing soil DNA: Potential for studying past and present
 519 ecosystems. *Molecular Ecology*, 21(8), 1821–1833. doi: 10.1111/j.1365-294X.2012.05537.x
- 520 Evans, S. E., Wallenstein, M. D., & Burke, I. C. (2014). Is bacterial moisture niche a good predictor of shifts in
 521 community composition under long-term drought. *Ecology*, 95(1), 110–122. doi: 10.1890/13-0500.1
- 522 Fernández-Calviño, D., & Bååth, E. (2010). Growth response of the bacterial community to pH in soils
 523 differing in pH. *FEMS Microbiology Ecology*, 73(1), 149–156. doi: 10.1111/j.1574-6941.2010.00873.x
- 524 Ficetola, G. F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., ... Pompanon, F. (2010). An In silico
 525 approach for the evaluation of DNA barcodes. *BMC Genomics*, 11(1), 434. doi: 10.1186/1471-2164-11-
 526 434
- 527 Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguët-Covex, C., De Barba, M., ... Taberlet, P. (2015).
 528 Replication levels, false presences and the estimation of the presence/absence from eDNA
 529 metabarcoding data. *Molecular Ecology Resources*. doi: 10.1111/1755-0998.12338
- 530 Field, A. P., Miles, J. 1968-, & Field, Z. (2012). *Discovering statistics Using r*. London ; Thousand Oaks, Calif. :
 531 Sage,.
- 532 Geyer, K. M., Altrichter, A. E., Takacs-Vesbach, C. D., Van Horn, D. J., Gooseff, M. N., & Barrett, J. E. (2014).
 533 Bacterial community composition of divergent soil habitats in a polar desert. *FEMS Microbiology
 534 Ecology*, 89(2), 490–494. doi: 10.1111/1574-6941.12306
- 535 Giovannoni, S. J., Britschgi, T. B., Moyer, C. L., & Field, K. G. (1990). Genetic diversity in Sargasso Sea
 536 bacterioplankton. *Nature*, 345(6270), 60–63. doi: 10.1038/345060a0
- 537 Guardiola, M., Uriz, M. J., Taberlet, P., Coissac, E., Wangenstein, O. S., & Turon, X. (2015). Deep-Sea, Deep-
 538 Sequencing: Metabarcoding Extracellular DNA from Sediments of Marine Canyons. *PLOS ONE*, 10(10),
 539 e0139633. doi: 10.1371/journal.pone.0139633
- 540 Guerrieri, Alessia et al. (2020), Effects of soil preservation for biodiversity monitoring using environmental
 541 DNA, Dryad, Dataset, <https://doi.org/10.5061/dryad.zkh189382>
- 542 Guo, Y., Yang, G. Q., Chen, Y., Li, D., & Guo, Z. (2018). A comparison of different methods for preserving
 543 plant molecular materials and the effect of degraded DNA on ddRAD sequencing. *Plant Diversity*,
 544 40(3), 106–116. doi: 10.1016/j.pld.2018.04.001
- 545 Hoffmann, C., Schubert, G., & Calvignac-Spencer, S. (2016, February 1). Aquatic biodiversity assessment for
 546 the lazy. *Molecular Ecology*, Vol. 25, pp. 846–848. doi: 10.1111/mec.13535
- 547 Huerlimann, R., Cooper, M. K., Edmunds, R. C., Villacorta-Rath, C., Le Port, A., Robson, H. L. A., ... Jerry, D. R.
 548 (2020). Enhancing tropical conservation and ecology research with aquatic environmental DNA
 549 methods: an introduction for non-environmental DNA specialists. *Animal Conservation*. doi:

550 10.1111/acv.12583

551 Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., ... Yu, D. W. (2013). Reliable,
552 verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*(10), 1245–
553 1257. doi: 10.1111/ele.12162

554 Kaiser, K., Wemheuer, B., Korolkow, V., Wemheuer, F., Nacke, H., Schöning, I., ... Daniel, R. (2016). Driving
555 forces of soil bacterial community structure, diversity, and function in temperate grasslands and
556 forests. *Scientific Reports*, *6*(August), 1–12. doi: 10.1038/srep33696

557 Kumar, G., Eble, J. E., & Gaither, M. R. (2020). A practical guide to sample preservation and pre-PCR
558 processing of aquatic environmental DNA. *Molecular Ecology Resources*, *20*(1), 29–39. doi:
559 10.1111/1755-0998.13107

560 Lauber, C. L., Zhou, N., Gordon, J. I., Knight, R., & Fierer, N. (2010). Effect of storage conditions on the
561 assessment of bacterial community structure in soil and human-associated samples. *FEMS
562 Microbiology Letters*, *307*(1), 80–86. doi: 10.1111/j.1574-6968.2010.01965.x

563 Legendre, P., & Borcard, D. (2018). Box-Cox-chord transformations for community composition data prior
564 to beta diversity analysis. *Ecography*, *41*(11), 1820–1824. doi: 10.1111/ecog.03498

565 Legendre, P., & Legendre, L. (2012). Numerical Ecology Ch 6 - Multidimensional qualitative data.
566 *Developments in Environmental Modelling*, *24*, 337–424. doi: 10.1016/B978-0-444-53868-0.50008-3

567 Majaneva, M., Diserud, O. H., Eagle, S. H. C., Boström, E., Hajibabaei, M., & Ekrem, T. (2018). Environmental
568 DNA filtration techniques affect recovered biodiversity. *Scientific Reports*, *8*(1), 4682. doi:
569 10.1038/s41598-018-23052-8

570 Meisner, A., Jacquiod, S., Snoek, B. L., Ten Hooven, F. C., & van der Putten, W. H. (2018). Drought legacy
571 effects on the composition of soil fungal and prokaryote communities. *Frontiers in Microbiology*,
572 *9*(MAR), 294. doi: 10.3389/fmicb.2018.00294

573 Nichols, R. V., Vollmers, C., Newsom, L. A., Wang, Y., Heintzman, P. D., Leighton, M., ... Shapiro, B. (2018).
574 Minimizing polymerase biases in metabarcoding. *Molecular Ecology Resources*, *18*(5), 927–939. doi:
575 10.1111/1755-0998.12895

576 Ochoa-Hueso, R., Collins, S. L., Delgado-Baquerizo, M., Hamonts, K., Pockman, W. T., Sinsabaugh, R. L., ...
577 Power, S. A. (2018). Drought consistently alters the composition of soil fungal and bacterial
578 communities in grasslands from two continents. *Global Change Biology*, *24*(7), 2818–2827. doi:
579 10.1111/gcb.14113

580 Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., ... Maintainer, H. W. (2019).
581 *Package “vegan” Title Community Ecology Package Version 2.5-6.*

582 Orchard, S., Standish, R. J., Nicol, D., Dickie, I. A., & Ryan, M. H. (2017). Sample storage conditions alter
583 colonisation structures of arbuscular mycorrhizal fungi and, particularly, fine root endophyte. *Plant
584 and Soil*, *412*(1–2), 35–42. doi: 10.1007/s11104-016-2867-4

585 Os B ´ Alint, M., Bahram, M., Murat Eren, A., Faust, K., Fuhrman, J. A., Bj¨, B., ... Tedersoo, L. (2016). Millions
586 of reads, thousands of taxa: microbial community structure and associations analyzed via marker
587 genes. *FEMS Microbiology Reviews*, *017*, 686–700. doi: 10.1093/femsre/fuw017

588 Paliy, O., & Shankar, V. (2016, March 1). Application of multivariate statistical techniques in microbial
589 ecology. *Molecular Ecology*, Vol. 25, pp. 1032–1057. doi: 10.1111/mec.13536

590 Pettersson, M., & Bååth, E. (2003). Temperature-dependent changes in the soil bacterial community in
591 limed and unlimed soil. *FEMS Microbiology Ecology*, *45*(1), 13–21. doi: 10.1016/S0168-
592 6496(03)00106-5

593 Spens, J., Evans, A. R., Halfmaerten, D., Knudsen, S. W., Sengupta, M. E., Mak, S. S. T., ... Hellström, M.
594 (2017). Comparison of capture and storage methods for aqueous microbial eDNA using an optimized
595 extraction protocol: advantage of enclosed filter. *Methods in Ecology and Evolution*, *8*(5), 635–645.
596 doi: 10.1111/2041-210X.12683

597 Strimmer, K. (2008). fdrtool: a versatile R package for estimating local and tail area-based false discovery
598 rates. *Bioinformatics (Oxford, England)*, *24*(12), 1461–1462. doi: 10.1093/bioinformatics/btn209

599 Taberlet, P., Bonin, A., Zinger, L., & Coissac, É. (2018). Environmental DNA: For Biodiversity Research and
600 Monitoring. In *Environmental DNA: For Biodiversity Research and Monitoring*. doi:
601 10.1093/oso/9780198767220.001.0001

602 TABERLET, P., COISSAC, E., POMPANON, F., BROCHMANN, C., & WILLERSLEV, E. (2012). Towards next-
603 generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, 21(8), 2045–2050.
604 doi: 10.1111/j.1365-294X.2012.05470.x

605 Ter Braak, C. J. F. (1986). Canonical correspondence analysis: a new eigenvector technique for multivariate
606 direct gradient analysis. *Ecology*, 67(5), 1167–1179. doi: 10.2307/1938672

607 Ward, D. M., Weller, R., & Bateson, M. M. (1990). 16S rRNA sequences reveal numerous uncultured
608 microorganisms in a natural community. *Nature*, 345(6270), 63–65. doi: 10.1038/345063a0

609 Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., ... Taberlet, P. (2019, April 1). DNA
610 metabarcoding—Need for robust experimental designs to draw sound ecological conclusions.
611 *Molecular Ecology*, Vol. 28, pp. 1857–1862. doi: 10.1111/mec.15060

612 Zinger, L., Taberlet, P., Schimann, H., Bonin, A., Boyer, F., De Barba, M., ... Chave, J. (2019). Body size
613 determines soil community assembly in a tropical forest. *Molecular Ecology*, 28(3), 528–543. doi:
614 10.1111/mec.14919
615

616 **DATA ACCESSIBILITY**

617 Raw sequences as well as filtered data are available in the DRYAD Digital Repository
618 (<https://doi.org/10.5061/dryad.zkh189382>)

619 .

620 **AUTHOR CONTRIBUTIONS**

621 TM, GFF, LG, AB, WT and AG designed the experiment. LG, TM and AG conducted the field work. AG
622 conducted all molecular analyses and performed the bioinformatic treatment of sequences with the help of
623 AB. AG, AB and GFF ran statistical analyses. AG, GFF and AB drafted the manuscript. All the authors
624 contributed substantially to the revision process, and accepted the final version.

625 **ORCID**

626 Alessia Guerrieri: <https://orcid.org/0000-0002-1519-3517>
627 Aurélie Bonin: <https://orcid.org/0000-0001-7800-8609>
628 Tamara Münkemüller: <https://orcid.org/0000-0001-9743-1322>
629 Ludovic Gielly: <https://orcid.org/0000-0001-5164-6512>
630 Wilfried Thuiller: <https://orcid.org/0000-0002-5388-5274>
631 Gentile Francesco Ficetola: <https://orcid.org/0000-0003-3414-5155>

632 **TABLES**

633 Table 1. Treatment contrasts assessing differences between the control (immediate extraction) and four
 634 approaches to soil conservation before eDNA extraction. Each conservation treatment was compared
 635 against the control in order to determine the percentage of explained variability.

	Bact02		Fung02		Euka02	
	Explained variability (%)	<i>P</i>	Explained variability (%)	<i>P</i>	Explained variability (%)	<i>P</i>
Treatment 2: room temperature, extraction after 6h;	0.58	0.956	0.53	1	0.44	0.993
Treatment 3: 4°C, extraction after 3 days;	0.81	0.563	0.58	0.976	0.54	0.949
Treatment 4: silica gel immediately inserted, extraction after 21 days;	3.14	<0.001	1.03	0.518	0.63	0.876
Treatment 5: silica gel inserted after 6h, extraction after 21 days;	3.16	<0.001	0.73	0.880	0.75	0.725

636

637 **FIGURES CAPTIONS**

638 Fig. 1. Experimental sampling design.

639 Fig. 2. MOTUs richness across the different treatments (control: immediate extraction; T2: extraction after
640 6h at room temperature; T3: extraction after three days at 4°C; T4: immediate preservation in tubes with
641 silica gel, extraction after 21 days; T5: preservation in tubes with silica gel after 6h at room temperature,
642 extraction after 21 days) before (left) and after (right) removing MOTUs with frequency < 1% in each
643 sample.

644 Fig. 3. Plots of non-metric dimensional scaling showing dissimilarities of communities among the three
645 habitats: broadleaved forest (black); grassland (green); vegetated riverbank (blue). Each plot shows the
646 results of metabarcoding analysis based on soil samples subjected to five different treatments.

647 Fig. 4. Procrustes correlation between communities obtained from metabarcoding analyses based on soil
648 samples across environmental conditions subjected to five sample treatments (control: immediate
649 extraction; RT+6h: extraction after 6h at room temperature; 4°C: extraction after three days at 4°C;
650 silicagel: immediate preservation in tubes with silica gel, extraction after 21 days; silicagel+6h: preservation
651 in tubes with silica gel after 6h at room temperature, extraction after 21 days. All correlation coefficients
652 are highly significant (all $P < 0.0001$).

653 Fig. 5. Guidelines for improving monitoring strategies with eDNA from soil.