Received: 30 November 2021 | Revised: 4 April 2022 | Accepted: 12 April 2022

DOI: 10.1002/edn3.315

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

WILEY

ORIGINAL ARTICLE

Evaluation of lake sedimentary ancient DNA metabarcoding to assess fungal biodiversity in Arctic paleoecosystems

Peter Andreas Seeber[1] | Barbara von Hippel[2] | Hårvard Kauserud[3] | Ulrike Löber[4,5] | Kathleen Rosmarie Stoof-Leichsenring[2] | Ulrike Herzschuh[2,6,7] | Laura S. Epp[1]

[1]Department of Biology, Limnological Institute, University of Konstanz, Konstanz, Germany

[2]Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Polar Terrestrial Environmental Systems, Potsdam, Germany

[3]Section for Genetics and Evolutionary Biology, Department of Biosciences, University of Oslo, Oslo, Norway

[4]Experimental and Clinical Research Center, Charité Universitätsmedizin Berlin and Max Delbrück Center for Molecular Medicine, Berlin, Germany

[5]Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Berlin, Germany

[6]Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany

[7]Institute of Biochemistry and Biology, University of Potsdam, Potsdam, Germany

Correspondence
Laura S. Epp, Limnological Institute, Mainaustr. 252 78464 Konstanz, Germany.
Email: laura.epp@uni-konstanz.de

Funding information
European Research Council (ERC) Consolidator Grant 'Glacial Legacy', Grant/Award Number: 772852; BiodivScen ERA-Net COFUND; Natural Sciences and Engineering Research Council of Canada; National Science Foundation; Academy of Finland; Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning, Formas; Research Council of Norway; Agence Nationale de la Recherche; Deutsche Forschungsgemeinschaft, Grant/Award Number: EP-98/3-1

Abstract

Fungi are crucial organisms in most ecosystems as they exert ecological key functions and are closely associated with land plants. Fungal community changes may, therefore, help reveal biodiversity changes in past ecosystems. Lake sediments contain the DNA of organisms in the catchment area, which allows reconstructing past biodiversity by using metabarcoding of ancient sedimentary DNA. We re-evaluated various commonly used metabarcoding primers, and we developed a novel PCR primer combination for fungal metabarcoding to produce a short amplicon, thus accounting for amplification bias due to the degradation of ancient DNA. *In silico* PCRs showed higher diversity using this new primer combination, compared with previously established fungal metabarcoding primers. We analyzed data from sediment cores from four artic and one boreal lake in Siberia. These cores had been stored for 2–22 years after coring; we, therefore, examined the degradation effects of ancient DNA and storage time-related bias affecting fungal communities. Amplicon lengths showed considerable variation within and between the major divisions of fungi, for example, amplicons of Basidiomycota were significantly longer than those of Mucoromycota; however, we observed no significant effect of sample age on amplicon length and GC content, suggesting the robustness of our results. We also found no indication of post-coring fungal growth during storage regarding the proportions of common mold taxa, which would otherwise distort conclusions on past fungal communities. Terrestrial soil fungi, including mycorrhizal fungi and saprotrophs, were predominant in all lakes, whereas typical aquatic taxa were only represented to a negligible extent, which supports the use of lake sedimentary ancient DNA for reconstructing terrestrial communities.

KEYWORDS
fungi, *in silico* PCR, ITS, lake sediment, metabarcoding, sedimentary ancient DNA

1 | INTRODUCTION

Fungi constitute the third-largest kingdom on Earth in terms of biomass, after plants, and bacteria (Bar-On et al., 2018), and they have considerable effects on the structure and functioning of most ecosystems. Fungi are essential to the survival, growth, and fitness of many organisms with which they form associations, including enumerable plant species, in almost all ecosystems (Brundrett, 2004; Finlay, 2008). The biodiversity of mycorrhizal fungi is known to influence plant community structures, thereby affecting entire ecosystems (Clemmensen et al., 2013; Powell & Rillig, 2018; van der Heijden et al., 1998). Particularly in environments which are notoriously nutrient poor, such as the Arctic, plants are typically highly dependent on symbioses with mycorrhizal and endophytic fungi (Smith & Read, 2008). Apart from mycorrhizal symbioses, fungi exert various other key ecological functions in terrestrial and aquatic habitats, including decomposition of components of complex substrates such as cellulose and lignin, subsequent recycling of nutrients, and pathogenic effects on countless taxa of eukaryotes (Grigoriev, 2013). However, compared to other kingdoms such as Animalia and Viridiplantae, knowledge on fungal diversity and distribution of fungal taxa and functional groups is relatively limited (Baldrian et al., 2021). The comparable lag in studies on fungi, particularly within their natural habitats, may be attributed, in part, to the microscopic dimensions of their vegetative bodies with hyphae of only a few microns in diameter, but also to the extreme taxonomic diversity within the fungal kingdom with a currently estimated 1–4 million species (Baldrian et al., 2021; Blackwell, 2011; Hawksworth & Lücking, 2017), compared to approximately 400,000 species of vascular plants. However, only approximately 120,000 species of fungi are thoroughly described and accepted (Hawksworth & Lücking, 2017).

One crucial problem in fungal systematics and taxonomy is that many fungi occur in morphs which differ drastically regarding their phenotypic appearance. Moreover, some fungal species can comprise multiple strains showing highly variable morphology, which has often led to their description as different species. Molecular approaches to assess biodiversity from environmental samples such as soils or sediments predominantly rely on metabarcoding and can be used on modern or ancient DNA (reviewed by Ruppert et al., 2019). This method is generally a powerful tool for assessing species richness in an ecosystem (Deiner et al., 2017); however, metabarcoding of fungi is notoriously challenging due to their overwhelming taxonomic diversity and, consequently, incomplete reference databases (James et al., 2020; Lücking et al., 2020; Nilsson et al., 2019). For fungi, the most common metabarcoding regions are the two internal transcribed spacers (ITS) ITS-1 and ITS-2 of the ribosomal RNA genes (Schoch et al., 2012; Stielow et al., 2015). Even though the resolution to species level based on ITS barcodes may be difficult, the current reference databases are the most comprehensive regarding this marker, compared to others (Lücking et al., 2020; Stielow et al., 2015), and it is also used as the standard barcode by the International Barcode of Life consortium. Databases for this marker have increased tremendously in the past years, facilitating more precise taxonomic assignments. Furthermore, the increase in available sequences potentially also offers the possibility to optimize metabarcoding assays in comparison to the primers that are currently in common use. However, fungal metabarcoding primers have not been re-evaluated in close to a decade (Bellemain et al., 2010, 2013; Epp et al., 2012), even though fungal metabarcoding may help reveal community turnovers and trace ecosystem changes on very long timescales (von Hippel et al., 2021). This is particularly critical regarding the high diversity of fungi and the resulting lack of truly universal but exclusive fungi metabarcoding primers (see UNITE primer notes, https://unite.ut.ee/primers.php). Moreover, for amplifying ancient DNA, amplicons should be particularly short so as to avoid bias due to DNA fragmentation, which is problematic as many common primer combinations produce amplicons exceeding 500 bp. Robust metabarcoding assays are of particular relevance for the analysis of potentially degraded environmental DNA, as recovered from ancient sedimentary deposits (Bellemain et al., 2013; Lydolph et al., 2005; Talas et al., 2021), thus we optimized metabarcoding primers for this purpose.

Analyses of sedimentary deposits can reveal fungal community alterations and associated ecosystem changes over long periods of time. Investigating past fungal biodiversity changes by reconstructing paleoecosystems may generate insights regarding basal structural developments that are to date mostly overlooked in classical palynological approaches due to reliance on microscopic remains (Chepstow-Lusty et al., 2019; Taylor & Osborn, 1996; Wood & Wilmshurst, 2013). While molecular genetic methods are commonly used at present to investigate modern ecosystems (Adamo et al., 2020; Heeger et al., 2018), far fewer studies have been conducted on fungal biodiversity in paleoecosystems. Early studies concentrated on samples from permafrost soils (Bellemain et al., 2013; Lydolph et al., 2005), in which DNA preservation is optimal and which was in general an early target for sedimentary ancient DNA (Haile et al., 2009; Willerslev et al., 2003, 2014). DNA from these deposits showed a high potential for the analysis of past communities, but with certain peculiarities regarding fungi. In particular, Bellemain et al. (2013) suggested that fungal DNA in permafrost potentially originated not only from ancient communities but also from organisms that were still alive. Moreover, sample material (such as sediment cores) is frequently stored for long periods of time before DNA isolation and may thus be prone to growth of fungi (e.g., molds), which could distort ancient community signals. Thus, such storage effects must be considered.

Regarding the ancient DNA of many other organismal groups, lake sediment cores have by now become the most commonly targeted environmental archive (Domaizon et al., 2017; Parducci et al., 2017). Lake sediments comprise organic and inorganic matter originating from the lake's catchment (and beyond), including intracellular and extracellular DNA of a vast spectrum of organisms which has been shed into the environment and was subsequently translocated to the sediment by various physical processes. Under adequate conditions (e.g., low temperatures and neutral to slightly basic

pH), integration into lake sediments may help preserve environmental DNA and shield it from degradation over considerable time spans (reviewed by Capo et al., 2021). Thus, sediment cores may contain a plethora of information with which past ecosystems and changes in biodiversity and community structures can be reconstructed. So far, explicit fungal metabarcoding from lake sedimentary DNA has been performed in only one study, which used a multiplex PCR approach (Talas et al., 2021), and showed that this approach can be used to assess past fungal communities and processes in lakes and the surrounding terrestrial environment.

Here, we developed metabarcoding primers to investigate past fungal biodiversity, and we assessed the specificities of fungal sedimentary ancient DNA (sedaDNA) extracted from lake sediment cores in Siberia (for details on the cores see von Hippel et al., 2021). To this end, we re-evaluated the existing metabarcoding assays for their use in paleoecology and updated metabarcoding primers for use with sedimentary ancient DNA. Using this assay, we examined taxonomic resolution and richness as well as replicability in paleo records of five Siberian lakes, four of which are located in the Arctic and one in boreal Southeast Siberia. All cores reach back through the Holocene and beyond and were stored for different time periods (2–22 years) in storage facilities. With this set of cores, we assessed potential biases due to sampling location, sample age, and storage time before DNA extraction, and we captured the taxonomic resolution of the marker and general fungal diversity on order level from lake sediments across a vast geographical area. For one core from the Taymyr Peninsula, Russia, which had the best temporal resolution, we explored trends of diversity in greater detail.

2 | MATERIALS AND METHODS

2.1 | Primer design and evaluation

2.1.1 | *In silico* analyses

Fungal ITS databases increased tremendously in the past years, for example, an *in silico* PCR using the established fungi primer pair ITS1F/ITS2 (Gardes & Bruns, 1993) produced 4658 sequences from EMBL release #102 (Epp et al., 2012) but 154,168 sequences from EMBL release #138, with the same settings. At the same time, the most common metabarcoding primers for fungi (see https://unite.ut.ee/primers.php) have been available for a much longer time (e.g., Gardes & Bruns, 1993; Vilgalys & Gonzalez, 1990; White et al., 1990). Thereby now, enlarged database requires re-evaluation of their specificity, universality, and taxonomic resolution and allows the design of additional primers. This may be particularly important for paleoecological studies of fungi, which rely on degraded DNA and thus should optimally amplify short DNA fragments (Pääbo, 1989). We, thus, tested a range of existing and newly designed PCR primers for the ITS-1 and ITS-2 regions and tested these *in silico* according to Bellemain et al., 2013 and Epp et al., 2012. Briefly, six databases were created by running an *in silico* PCR on the complete EMBL release #138 using ecoPCR software (Ficetola et al., 2010) and with the fungi barcoding primers of Epp et al. (2012). We allowed for three mismatches except for the last two nucleotide (nt) positions at the 3′-end of each primer and limited product size to 50–2000 nt, apart from one amplicon database which spanned the entire ITS-1–5.8S–ITS-2-region and was permitted a maximum product size of 4000 nt (primers ITS1 and ITS4). On each of these six amplicon databases, we used ecoPrimers software (Riaz et al., 2011) to identify potential PCR primers which had to exactly match 70% of the target sequences and match 90% of the target sequences with a maximum of three mismatches; product size was limited to 50–500 nt. From the output of each database, we selected primer pairs with a maximum difference in melting temperatures of 4°C. These primers were then tested in terms of universality within the largest fungal divisions, that is, Ascomycota, Basidiomycota, and other fungal taxa collated as "*incertae sedis*" using *in silico* PCRs; the produced amplicons were dereplicated using the command "obiuniq" of the OBITools software (Boyer et al., 2016). Primers that showed a strong bias against any of these three divisions were eliminated. Regarding non-target taxa, we examined amplification of Viridiplantae and Metazoa sequences; primers that produced a large number of these (i.e., more non-target sequences than sequences from any of the three fungi divisions) were eliminated. Sixteen primer pairs (eight targeting ITS-1 and eight targeting ITS-2) were tested *in silico* (Table 1, Table S1). To define optimal primers to universally and specifically target fungi in degraded DNA, primer combinations were excluded (1) if the mean amplicon size exceeded 250 bp, (2) if the subkingdom Dikarya as the largest phylum was represented by fewer than 50,000 unique amplicons, (3) if Basidiomycota and Fungi *incertae sedis* amplicons were more abundant than those of the largest division Ascomycota, and (4) if the number of off-target taxa (Viridiplantae and Metazoa, respectively) exceeded 10% of the number of Dikarya amplicons.

The following primer combinations targeting ITS-1 and ITS-2, respectively, were considered most promising: ITS67 (this study) and 5.8S_fungi (Epp et al., 2012) as well as ITS779 and ITS382r (this study). For the metabarcoding experiments, we selected the combination ITS67 and 5.8S_fungi as they produced slightly shorter amplicons and appeared to produce a somewhat more even ratio of ascomycetes and other divisions.

2.1.2 | PCR conditions

Primers ITS67 and 5.8S_Fungi were tested *in vitro* using DNA isolated from museum vouchers. Samples of the following species were obtained from the Natural History Museum of Oslo, Norway: *Sordaria alcina*, *Podospora fimicola*, *Podospora pyriformis*, *Cladonia rangiferina*, *Panaeolus fimicola*, *Coprinus sterquilinus*, *Coprinopsis cinerea*, and *Pilobolus chrystallinus*. DNA was isolated from these specimens using the NucleoSpin Plant II kit (Macherey Nagel, Düren, Germany) according to the manufacturer's instructions. The PCR reaction mix included 0.25-µl Platinum Taq DNA Polymerase High Fidelity (Thermo Fisher Scientific; 5 U/µL), 12.75-µl ultrapure water, 2.5 µl 10×buffer,

TABLE 1 Evaluation of candidate primers for fungal metabarcoding using *in silico* PCRs

	Dikarya		Ascomycota		Basidiomycota		Fungi incertae sedis		Viridiplantae	Metazoa	Percent	Percent
Primer combination	***Count***	***Length***	***Count***	***Length***	***Count***	***Length***	***Count***	***Length***	***Count***	***Count***	**on-target**	**off-target**
ITS1												
ITS67 + 5.8S_Fungi	219,484	191 (50–500)	157,048	186 (52–500)	62,165	206 (67–490)	7418	168 (65–437)	6078	386	97	3
NS7 + ITS2	603	464 (54–500)	115	431 (105–500)	480	476 (458–493)	2283	496 (387–500)	2982	733	44	56
ITSF-1 + ITS2	44,026	294 (59–497)	34,039	301 (64–499)	9268	268 (59–484)	8694	191 (63–489)	22	5603	90	10
ITS1 + ITS2	175,126	232 (50–498)	127,009	228 (50–500)	48,061	245 (98–500)	13,755	179 (65–484)	79,288	9782	68	32
ITS5 + 5.8S_Fungi	34,832	238 (66–500)	12,133	362 (73–496)	22,951	233 (98–493)	10,262	161 (150–497)	3316	109	93	7
ITS67 + ITS67r	209,739	203 (52–500)	147,084	152 (54–431)	62,465	218 (78–500)	14,084	152 (54–468)	68,142	40	77	23
NS70 + ITS70r	216,395	205 (54–499)	153,704	198 (54–500)	63,958	218 (75–499)	14,517	153 (50–432)	66,517	35	78	22
NS241 + ITS241r	205,084	207 (57–500)	143,562	202 (57–599)	60,922	221 (78–500)	14,308	155 (52–435)	62,895	43	78	22
ITS2												
ITS3 + LR3	19,728	366 (119–500)	15,457	381 (119–500)	3366	263 (144–481)	3	411 (411–413)	1662	897	89	11
ITS778 + ITS382r	193,800	206 (67–500)	140,473	193 (67–496)	53,512	239 (96–500)	7179	231 (132–442)	679	3438	97	3
ITS779 + ITS382r	245,072	203 (67–500)	194,556	194 (67–496)	50,500	239 (96–500)	6820	231 (132–442)	3182	3120	99	1
ITS779 + ITS4	35,622	235 (58–500)	27,244	230 (58–500)	8367	252 (131–471)	3405	244 (156–459)	2233	1715	91	9
ITS3 + ITS4	127,574	328 (56–499)	91,520	315 (56–498)	35,935	361 (97–499)	8800	355 (81–500)	35,320	8142	75	25
ITS3 + ITS382r	252,135	309 (82–500)	186,304	298 (116–500)	65,677	346 (82–500)	12,931	340 (66–499)	59,521	17,118	78	22
LROR + LR7	4	300 (271–387)	3	271 (271–271)	1	387	0	n/a	15	17	11	88
ITS3NGS1 + ITS4	130,220	330 (57–500)	91,600	316 (57–500)	38,582	362 (98–500)	3790	363 (261–490)	9224	8342	88	11

Notes: Targeted were the internal transcribed spacer (ITS)-1 (top section) and ITS2 (middle section), respectively, with EMBL release #142 as a reference database; shown are the counts of unique amplicons of the fungal subkingdom Dikarya, the divisions Ascomycota, Basidiomycota, and other divisions which are comprised at division rank as "Fungi *incertae sedis*", as well as the off-target clades Viridiplantae and Metazoa, and average amplicon lengths (with the respective minimum and maximum, in parentheses; for fungal taxa). The exclusion of primer combinations for comprehensive fungal metabarcoding according to the specified criteria is indicated by gray shading. Bold print indicates the primers selected for the metabarcoding experiments. Nucleotide sequences and references are shown in Table S1.

2.5-μl dNTPs (2.5 mM), 1-μl BSA (20 mg/mL; New England Biolabs, Frankfurt, Germany), 1-μl $MgSO_4$ (50 mM), and 1 μl of each primer (5 μM). The following thermocycling protocol was used: 94 °C for 5 min, followed by 40 cycles of 94 °C for 20 s, 56 °C for 20 s, 68 °C for 30 s, and final extension at 72 °C for 10 min.

2.2 | Evaluation of lake sediment core DNA for analyses of fungal paleoecology

2.2.1 | Sampling and DNA extraction

The laboratory works were performed at the Palaeogenetic Laboratory at AWI in Potsdam (von Hippel et al., 2021). We used sedaDNA isolated from sediment cores of five lakes in the Russian Federation, four of which are situated in the Arctic: a) a lake termed CH12 located in the Siberian Taymyr region (Khatanga, Russian Federation; 72.399°N, 102.289°E; 60 m a.s.l., collected in 2011; chronology of the core described previously [Klemm et al., 2016]; ages are shown as years before present [b.p.]), b) Lake Lama, Taymyr peninsula, 69.520°N, 89.948°E; 53 m a.s.l., collected in 1997, c) Lake Levinson-Lessing (74.512°N, 98.591°E; 47 m a.s.l.), Taymyr peninsula, collected in 2017, d) Lake Kyutyunda (69.630° N, 123.649°E; 66 m a.s.l.), Yakutia, northeastern Siberia, collected in 2010, and e) Lake Bolshoye Toko (56.265°N, 130.530°E, 903 m a.s.l.) southeastern Siberia (Neryungrinsky District, Sakha Republic), collected in 2013. The samples per core and the respective ages are shown in Table S2; for details on age-depth models, see von Hippel et al. (2021).

DNA was isolated from approximately 2–5 g sediment using the PowerMax Soil DNA Isolation kit (Qiagen), and DNA extracts were purified and normalized to 3 ng/μl using a GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific); for details, see von Hippel et al. (2021). Seventy samples were used (CH12: 28 samples; Lake Lama: 15 samples; Lake Levinson-Lessing: 9 samples; Lake Kyutuynda: 10 samples; Lake Bolshoye Toko: 8 samples; Table S2). Extraction blanks of each batch of DNA isolation were processed along with the DNA extracts. All extractions and metabarcoding PCR setup were performed in dedicated ancient-DNA laboratory facilities of the Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Potsdam, Germany.

2.2.2 | Metabarcoding PCRs and sequencing

sedaDNA metabarcoding PCRs of CH12 extracts were performed according to the established PCR conditions as specified above, using primers tagged with individual eight-bp tags preceded by three variable positions ('N') to improve cluster formation and using 9 ng DNA; PCRs on other cores are described by von Hippel et al. (2021). Six technical replicates of the PCR of each extract (N = 70) and extraction blank (N = 11) were used, with one non-template control per PCR batch (N = 30), resulting in 510 PCR reactions. MetaFast library preparation and paired-end sequencing at 2 × 250 bp on an Illumina MiSeq platform (Illumina, San Diego, CA, USA) were performed by a commercial service provider (Fasteris SA, Geneva, Switzerland).

2.2.3 | Data analyses

Data were processed and analyzed using OBITools software (Boyer et al., 2016) as described by von Hippel et al. (2021). Briefly, sequences were clustered using sumaclust (Mercier et al., 2013) at a score threshold of 0.97 and were annotated using the OBITools ecotag assignment with two databases—the latest UNITE release v. 8.2 (Abarenkov et al., 2010) and the EMBL Nucleotide Sequence Database (Kanz et al., 2005) release #142. Only reads containing both primers and both tags were kept in the dataset. The dataset was then filtered by best identity scores to count the number of unique taxa on order, family, genus levels in a range of 95%–100% best identity (Table 2). As we observed that the curves of additional numbers of taxa flattened at approximately 97% (Figure S1), we removed all OTUs with the best identity <97%. We retained only OTUs that a) occurred at a minimum of 10 reads per PCR replicate and b) occurred at a minimum of 100 reads across the entire dataset. To account for dysfunctional PCRs, replicates with <50 reads were discarded. The retained replicates showed highly variable read numbers, however, as we attempted no quantitative ecological interpretation but aimed for a full descriptive analysis, we opted against resampling so as to avoid exclusion of samples. For ecological interpretations based on normalizing the dataset by resampling, see von Hippel et al. (2021). To minimize bias due to varying PCR efficiency, we normalized the read counts by transforming each OTU's count in a replicate to a proportion of the sum of all reads in the respective sample (proportional data per replicate). To produce proportional data per sample, reads per replicate were summed, and proportions were calculated accordingly. Shannon–Wiener diversity indices of each sample were calculated per lake on the taxonomic levels genus, family, and order, using the function "ddply" of the R package plyr version 1.8.6 (Wickham, 2011).

Biases introduced through experimental procedures and/or through DNA degradation were examined in a number of ways: to assess whether we sampled the fungal diversity comprehensively and representatively, we examined accumulation curves for the combined PCR replicates through rarefaction in single PCRs. Specifically, to determine the number of OTUs per cumulative number of PCR replicates, we produced accumulation curves of each sample of core CH12 using the function "specaccum" of the R package vegan (Oksanen et al., 2020). To examine whether sequencing depth was sufficient, we rarefied the dataset of core CH12 using "rarefy" in vegan (Oksanen et al., 2020) and produced rarefaction curves. To test the replicability of the results, we investigated whether dissimilarities between samples were larger than those between replicates. For core CH12, which was examined at the highest temporal resolution and which has previously been analyzed for pollen and plant DNA (Epp et al., 2018; Klemm et al., 2016), we examined Bray–Curtis dissimilarities between PCR replicates of all samples to test whether

TABLE 2 Numbers of OTUs and percent of respective reads of fungi (UNITE database) and non-fungi taxa (EMBL database; lower section) per taxonomic order

Clade	Division	Class	Order	Number of OTUs	Assigned reads (%)		
					97% b.i.	90% b.i.	85% b.i.
Fungi	Ascomycota	Dothideomycetes	Pleosporales	7	7.29	5.42	7.63
			Dothideales	2	0.43	0.30	0.29
			Venturiales	2	0.19	2.27	1.87
			Cladosporiales	1	0.01	0.01	0.01
		Eurotiomycetes	Eurotiales	14	6.49	5.37	4.45
			Chaetothyriales	2	1.16	7.23	6.42
			Verrucariales	1	0.03	0.20	0.29
		Leotiomycetes	Helotiales	11	38.80	20.94	19.73
			Phacidiales	1	0.24	0.14	0.12
			Leotiales	2	0.15	0.28	4.33
			Erysiphales	1	0.06	0.03	0.02
		Lecanoromycetes	Peltigerales	4	0.72	0.73	0.60
			Lecanorales	2	0.06	0.39	0.45
		Saccharomycetes	Saccharomycetales	4	0.12	5.55	4.60
		Sordariomycetes	Hypocreales	1	0.45	3.33	2.74
	Basidiomycota	Agaricomycetes	Agaricales	23	3.51	5.40	5.01
			Polyporales	2	0.98	0.41	0.33
			Hymenochaetales	1	0.90	0.32	0.26
			Russulales	3	0.15	0.38	0.31
			Boletales	2	0.03	0.04	0.32
			Thelephorales	1	0.01	0.38	0.46
			Atheliales	1	0.01	0.13	0.11
		Tremellomycetes	Filobasidiales	10	3.18	1.14	0.95
			Tremellales	9	1.98	2.01	1.74
			Trichosporonales	1	0.04	0.07	0.06
			Cystofilobasidiales	1	0.02	0.54	0.45
		Microbotryomycetes	Sporidiobolales	5	2.17	0.77	0.63
			Leucosporidiales	2	0.47	0.16	0.15
	Mucoromycota	Mortierellomycetes	Mortierellales	9	19.81	21.90	22.01
		Mucoromycetes	Mucorales	1	0.06	0.02	0.10
		Glomeromycetes	Diversisporales	3	0.23	0.43	0.41
		Umbelopsidomycetes	Umbelopsidales	2	0.33	0.34	0.56
	Olpidiomycota	Olpidiomycetes	Olpidiales	1	0.06	0.10	0.09
Non-target taxa							
Viridiplantae	Chlorophyceae		Sphaeropleales	18	0.49		
			Chlamydomonadales	1	0.04		
	Trebouxiophyceae		Chlorellales	1	0.02		
	Chlorodendrophyceae		Chlorodendrales	1	0.01		
SAR supergroup	Eustigmatophyceae		Eustigmatales	1	0.01		
	Dinophyceae		Suessiales	1	<0.01		
Metazoa	Insecta		Diptera	1	<0.01		

Note: Percentages are based on the sum of assigned reads in the respective database, at three levels of best identity (b.i.).

within-sample variation is smaller than between-sample variation. We used the transformed proportional data of core CH12 for a multiple response permutation procedure (MRPP) with the function mrpp (vegan; Oksanen et al., 2020) and Bray–Curtis dissimilarity distances.

Using data from all cores, the potential effect of DNA degradation on the results was assessed both considering the GC content of recovered sequences and sequence length. To assess DNA degradation-induced bias toward amplicons with higher GC content (Dabney et al., 2013), we examined GC proportions over time as weighted means per sample. The effect of age on weighted mean GC content was tested by fitting a linear regression model. To determine differences in amplicon length between fungal divisions, we used an ANOVA and a Tukey's test post hoc; to assess the potential effects of age on amplicon length (i.e., whether longer fragments are less likely to be amplified in older samples), we fitted a linear regression of amplicon length and age for each core. Data analyses were performed using R software, version 3.6.0 (R Development Core Team, 2019).

3 | RESULTS

3.1 | Primer design and evaluation

We evaluated potential combinations of newly designed and previously established metabarcoding primers *in silico*, and three candidate primer pairs did not contravene any of the exclusion criteria (i.e., mean amplicon size ≤250 bp, ≥ 50,000 Dikarya amplicons, more Ascomycota than Basidiomycota and Fungi *incertae sedis* amplicons, and Viridiplantae or Metazoa at ≤10% of the number of Dikarya amplicons); the combination ITS67 (5′-ACC TGC GGA AGG ATC ATT-3′; this study) and 5.8S_Fungi (5′-CAA GAG ATC CGT TGT TGA AAG TT-3′; Epp et al., 2012) produced short amplicons (mean length of 183 bp), showed high specificity to fungi (91% of the amplicons assigned to fungi), and amplified a high number of target sequences *in silico* (N = 383,992). Most other primer combinations were excluded because they produced large numbers of off-target amplicons (mostly of plants) or amplicons which exceeded our chosen threshold of 250 bp (Table 1). For the selected primers, *in silico* amplicons of Ascomycota (mean length 186 bp; N = 157,058) were shorter than those of Basidiomycota (mean 207 bp; N = 62,108), and fungi *incertae sedis* (including Mucoromycota) amplicons were, on average, 168 bp long (N = 14,308). The taxonomic resolution (i.e., the proportion of unambiguously identified taxa) of the amplified marker was 53% at species, 67% at genus, and 60% at the family level, according to the ecotaxspecificity function of OBITools software (Boyer et al., 2016). Regarding off-target amplification, the selected primer combination produced 6078 Viridiplantae and 386 Metazoa sequences.

Primers ITS67 and 5.8S_Fungi were tested *in vitro* in PCRs, and the reaction conditions were optimized using template DNA extracted from Museum vouchers of eight fungal taxa, including the genera *Sordaria*, *Podospora*, *Cladonia*, *Panaeolus*, *Coprinus*, *Coprinopsis*, and *Pilobolus*. All reactions produced amplicons in the expected size range.

3.2 | Evaluation of lake sediment core DNA for fungal paleoecology

3.2.1 | Taxonomic resolution across the cores

After processing and filtering of the raw data including clustering at 97% (described in detail by von Hippel et al., 2021), the resulting 5411 cluster centroids were subjected to taxonomic assignments with each database and to subsequent filtering, as indicated above. Across the 70 samples of five cores and using the UNITE database, 135 operational taxonomic unit (OTU) cluster centroids were retained (Tables S3 and S4), which comprised 33 taxonomic orders (Table 2), 57 families, 79 genera, and 113 species. Regarding maximum taxonomic resolution, 121 (89%) OTUs were assigned to species level, 7 (5%) to genus level, 3 (3%) to family level, and 3 (3%) to order level. Using the EMBL database, 384 OTU cluster centroids were retained, comprising 51 orders, 100 families, 152 genera, and 188 species; 556 (50%) of the OTUs were assigned to species level, 275 (29%) to genus level, 66 (7%) to family level, and 13 (1%) to order level; 4% of the OTUs were assigned only to higher taxonomic levels. The overall weighted mean frequencies at class level were plotted across all samples of each lake (Figure 1), and changes in taxonomic resolution (i.e., assignment to species, genus, family, class, and order level) oversample age are shown in Figure 2.

3.2.2 | Comprehensiveness: Rarefaction and accumulation curves

Rarefaction analysis was performed to estimate OTU richness as a function of sampling effort, that is, sequencing depth, based on the minimum number of observed sequence counts. Most rarefaction curves showed an asymptotic course, suggesting that sequencing depth was sufficient (Figure S2). To test whether the number of PCR replicates was sufficient to assess taxonomic richness per sample, we produced OTU accumulation curves of the core with the highest temporal resolution (from lake CH12; [Klemm et al., 2016]). These curves did not show saturation over the maximum of six replicates (Figure S3).

3.2.3 | Amplicon length and GC content to assess bias through degradation

Average amplicon length as a function of age per core was tested to assess the effect of age on DNA degradation. This suggested only negligible decreases in mean amplicon length over time (Figure 2a); in none of the five cores did the linear regression indicate a significant effect of age on amplicon length, regardless of taxonomic

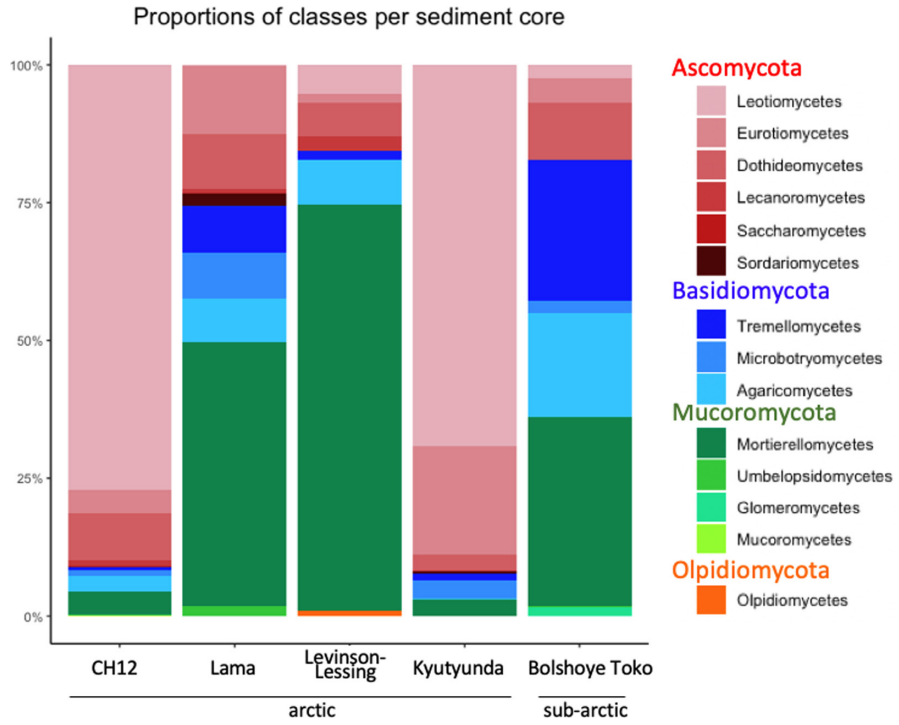

FIGURE 1 Proportions of fungal classes in lake sediments, averaged across all respective samples

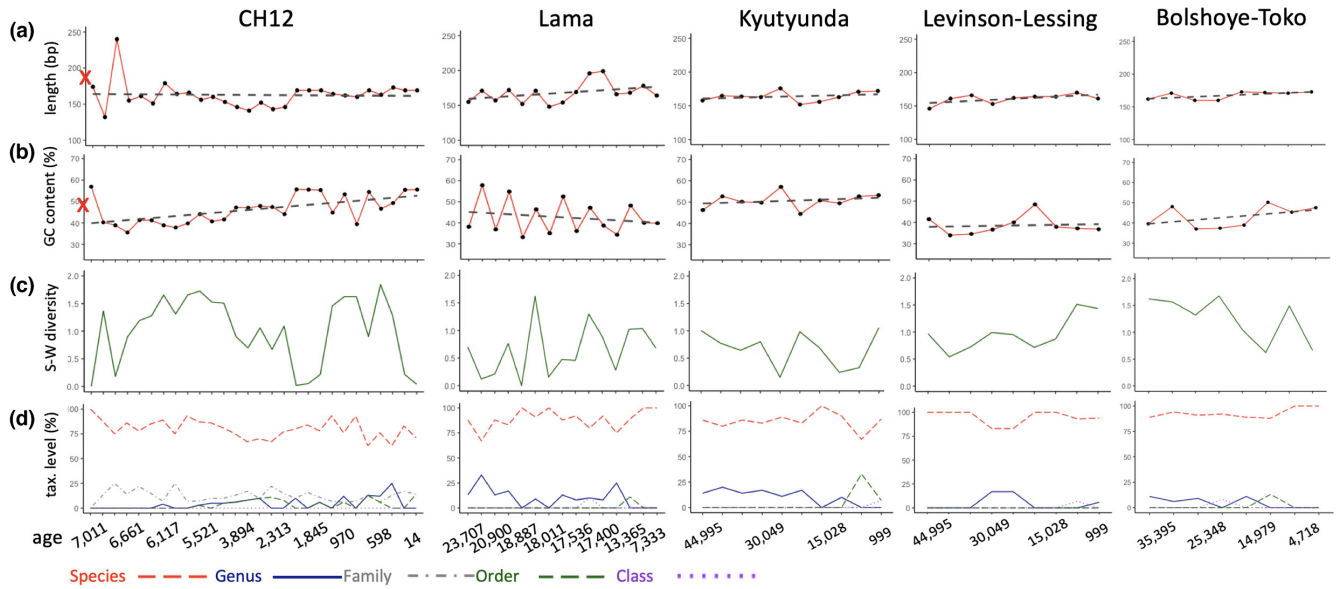

FIGURE 2 Effect of sample age on (a) mean amplicon lengths, (b) weighted mean GC content, (c) Shannon–Wiener (S-W) diversity indices of fungal taxa on order level, and (d) taxonomic (tax.) resolution, that is, the percentage of OTUs assigned on species (red), genus (blue), family (gray), order (green), and class (purple) level. Red crosses on the y-axes of panels a and B indicate the respective values of the *in silico* PCR output (183 bp length and 74% GC content)

assignment. Degradation processes linked to the age of DNA also lead to an increase in GC content (Dabney et al., 2013); in the current study, amplicon GC content ranged from 10% to 63% (mean 43% ± 10%), and we observed no bias toward higher GC content in older samples. In core CH12, GC content rather decreased significantly with age (est. -0.19; $t = -3.33$; $p = 0.003$), while no significant effect of age on GC content was observed in the cores of the other four lakes (Figure 2b). By comparison, the GC content in the *in silico*-PCR output was 47%.

3.2.4 | General taxonomic composition of fungi in Siberian lake sediment cores

The predominant taxa (Table 2 and Figure 1) in all lakes were terrestrial saprotrophs, mycorrhizal fungi, and other soil fungi, for example, OTUs assigned to the genera *Mortierella* (order Mortierellales) and *Inocybe* (Agaricales) (Domsch et al., 2007; Varma et al., 2017), respectively. We found very few sequences of fungi that were determined to be aquatic, that is, OTUs assigned to *Alatospora* sp. (Ascomycota).

These were restricted to four samples of Lake Levinson-Lessing and comprised less than 4000 reads in the total dataset.

The dataset also contained reads of off-target taxa (i.e., non-fungi) which were assigned using the EMBL database. Out of 936 EMBL-assigned OTUs, 24 OTUs at a total of 59,902 reads were assigned to non-fungi taxa, which belonged to the clades Viridiplantae (21 OTUs; 55,777 reads), the SAR supergroup (2 OTUs; 2788 reads), and Metazoa (1 OTU; 1337 reads; Table 2).

Reads produced from extraction blanks and non-template PCR controls were processed as described above and were assigned using the UNITE databases, which showed assignment of 141,746 reads to 34 OTUs. Three of the five most abundant taxa in the PCR and extraction controls (*Wickerhamomyces*, *Candida*, and *Pichia*) did not occur at all among assigned sample reads (Table S3), and only the genera *Aspergillus* and *Gryganskiella* occurred in total numbers of >10.000 reads in extraction blanks and non-template controls (Table S5).

3.3 | Diversity of fungal paleocommunities from lake CH12

We examined Bray–Curtis dissimilarities between PCR replicates of all samples of core CH12 to test whether the within-sample variation is smaller than between-sample variation, and MRPP analysis suggested significantly lower dissimilarity between PCR replicates of one sample (64.97%) than between those of different samples (88.46%; $p = 0.001$; within-group agreement = 0.26; observed delta 0.6503; expected delta 0.8774). In the graphical cluster visualization based on Bray–Curtis distances, however, most replicates clustered together regardless of samples, likely due to the predominance of specific taxa in most samples (Figure S4). We, therefore, repeated this analysis after excluding a) the most abundant taxa (> 50.000 reads in total) or b) the least abundant taxa (<50.000 reads in total); however, this analysis did also not visually differentiate samples based on similarities of the PCR replicates.

Across all cores, the fungal divisions Ascomycota, Basidiomycota, Mucoromycota, and Olpidiomycota were represented and accounted for approximately 62%, 15%, and 23%, and 0.06% of the fungal reads, respectively (Table 2). Amplicons of Basidiomycota were significantly longer (179 ± 44 bp) than those of Mucoromycota (153 ± 24 bp; ANOVA F = 3.383; $p = 0.037$; Tukey's test $p = 0.04$). Amplicon lengths of Ascomycota (169 ± 22 bp) did not differ significantly from those of Basidiomycota and Mucoromycota. The single OTU assigned to Olpidiomycota was 104 bp long. In the case of CH12, we found that Ascomycota was more abundant in younger samples and that the proportion of Mucoromycota (shortest amplicons) was higher in older samples; however, Basidiomycota (longest amplicons) also appeared to be more abundant in older samples (Figure 3). Shannon–Wiener diversity indices showed considerable variation in fungal communities over time (Figures 2c). In core CH12, OTUs were assigned to a total of 23 orders, and taxonomic richness ranged from 1 (in the oldest sample, dated 7011 years) to 14 (sample 5521 years; mean 8 ± 3) with particularly low diversity in samples dated 1976, 1845, and 14 years (diversity indices 0.02, 0,05, and 0.03, respectively), even though the respective OTUs were assigned to 7, 9, and 6 orders; this discrepancy between richness and diversity is due to the dominance of OTUs assigned to the order Helotiales in these samples (> 99%, each).

The cumulative overall proportions of the typical mold genera *Aspergillus*, *Cladosporium*, *Mucor*, and *Penicillium* in the cores were as follows: 1.37% in Levinson-Lessing (2017), 4.21% in Bolshoye Toko (2013), 3.72% in CH12 (2011), 15.66% in Kyutyunda (2010), and 5.76% in Lake Lama (1997). The respective proportions remained relatively stable over time and did not appear to increase in deeper layers (Figure S5). By comparison, the proportion of these genera *in silico* PCR output was 3.87%.

4 | DISCUSSION

4.1 | Preservation biases and potential contamination

Our results confirm that ancient DNA from lake sediment cores can be used for fungal metabarcoding from boreal to arctic sites, with samples aged up to 44,995 years. Fungal paleocommunities from lake sediments have also been recently reported from temperate conditions and spanning the Holocene (Talas et al., 2021). Our putatively authentic ancient fungal DNA, attributed primarily to taxa from the terrestrial surroundings of the lake, was in part recovered from cores that had been stored in a storage facility under non-frozen conditions for multiple years prior to sampling for DNA, underlining the value of sediment core collections also for work with sedimentary ancient DNA. For the ITS1-amplicon targeted by our primers, the average length of Ascomycota was shorter than that of Basidiomycota, both *in silico* using ecoPCR on the EMBL database and in the sequencing results (198 vs. 219 bp and 169 vs. 179 bp, respectively), which is in line with the results of a previous study (Bellemain et al., 2010). This difference was smaller in sequencing reads from sedaDNA than in *in silico* PCR products, which could possibly be attributed to the fragmentation of DNA in the environment. An increase in the proportion of shorter amplicons with sample age may suggest a bias against the amplification of longer DNA sequences, likely due to DNA fragmentation; however, our results suggest only minor, non-significant declines in amplicon lengths over time in each sediment core (Figure 2), indicating that the results are not impacted by DNA fragmentation.

A further source of bias linked to degradation can potentially change the GC content of the DNA. GC content can increase with the age of DNA due to degradation impacting primarily GC-poor fragments (Dabney et al., 2013). In the current study, GC content ranged from 10% to 63%, and we observed no bias toward higher GC content in older samples. Contrary to our expectations, we found a decrease in GC content with age in one lake (CH12). This may, however, be an incidental result considering that no significant effect

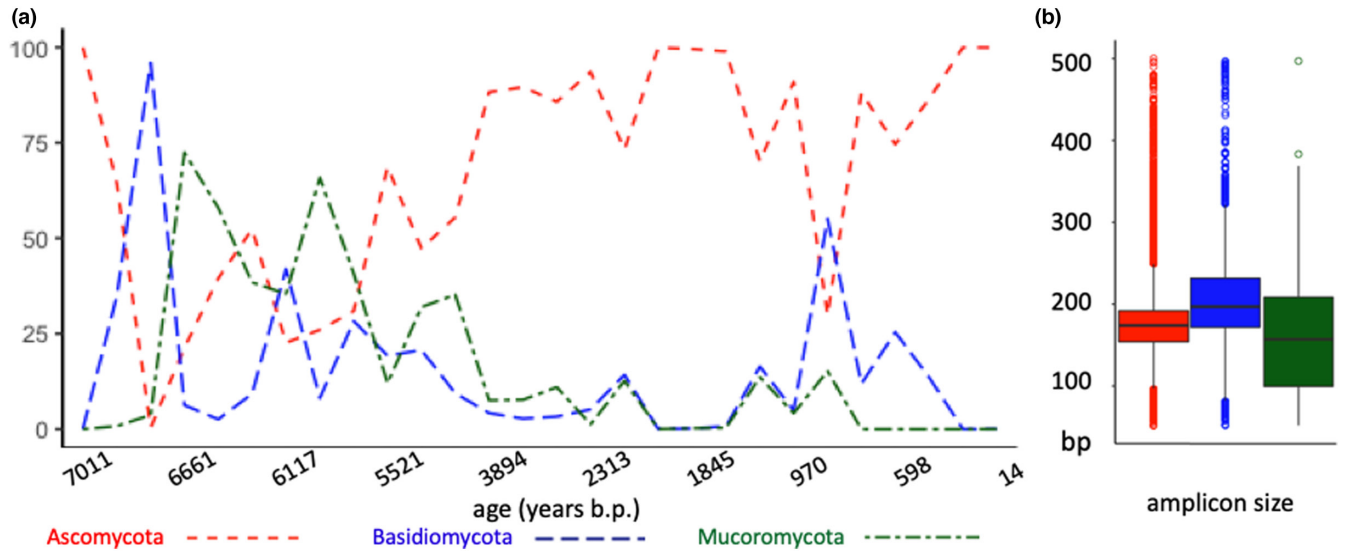

FIGURE 3 Changes in proportions of the three major divisions Ascomycota (red), Basidiomycota (blue), and Mucoromycota (green) over sample age of lake CH12 (a), and respective fragment lengths (b); boxes indicate 2nd and 3rd quartiles, center lines indicate median values, upper (and lower) whiskers extend to the highest (and lowest) value within 1.5-times the interquartile range. Data points beyond the end of the whiskers are shown as open dots

of age on GC content was observed in the cores of the other four lakes (Figure 2b). Taken together, the lack of GC bias corresponding to sample age suggests that degradation of arctic and subarctic *sed*aDNA may not be problematic for fungal metabarcoding. As we worked with cores that were not sampled for DNA soon after coring, but had in part been stored for 2–22 years prior to collecting and freezing DNA samples, we furthermore screened our results for the occurrence of taxa that point to recent fungal growth in or on the cores (e.g., dominance of molds). Growth of molds during storage of the cores or within the sediment would be expected to increase the proportion of these taxa; however, we found little indication for such a bias in our samples (Figure S4). Even though the second-oldest core (lake Kyutuynda; from 2010) showed the highest cumulative proportion of the four mold genera *Aspergillus*, *Cladosporium*, *Mucor*, and *Penicillium* (15.66%), this was only due to two samples with particularly high abundance, whereas in the core from 1997 (Lake Lama), a substantially lower proportion of these taxa was observed (5.76%). Thus, long storage times of sediment cores are not necessarily problematic for reproducing ancient fungal communities.

4.2 | Characteristics of the optimized sedaDNA ITS1 metabarcoding assay

A large part of the ancient DNA pool in sediment cores is typically reduced to fragments well below 50 bp (Lammers et al., 2021; Parducci et al., 2017; Pedersen et al., 2016) and will not be traceable at all through PCR; however, for metabarcoding, short amplicons are preferable. We, therefore, aimed for metabarcoding primers to amplify the shortest possible amplicons, so as to account for potential length bias due to fragmentation of DNA with age, while providing high specificity to fungi and high taxonomic resolution. Diversity estimates of fungal communities (also including lichenized fungi) based on metabarcoding may be considerably biased by the barcode locus (Banchi et al., 2018; Tedersoo et al., 2015; Tedersoo & Lindahl, 2016), and ITS-2 was suggested to be preferable over ITS-1, particularly with respect to taxonomic assignment of lichenized fungi and Basidiomycetes (Banchi et al., 2018; Tedersoo et al., 2015). Thus, a combined approach using both ITS loci would be most reliable. However, the larger fragment length of ITS-2 may introduce amplification bias when ancient DNA is concerned. Our results using the primer combination optimized *in silico* suggested high taxonomic diversity at relatively short amplicon lengths (mean length of 183 bp), without the discernible introduction of biases against any one division. Due to the structure of the ITS region and its marked length variation, we could not design a universal primer pair for all fungi with a shorter amplicon, but our results suggest no bias in results due to group-specific amplicon length differences. Compared to an assay used recently for fungal DNA from a lake sediment core (Talas et al., 2021), the amplicon we propose appears to be substantially shorter (Table 1); in the study of Talas et al. (2021), greater length did not seem to compromise the results, but this may not universally be the case, and a different study recovered largely similar communities using ITS-1 and ITS-2 markers (Blaalid et al., 2013).

Despite its short length, the taxonomic resolution offered by our metabarcoding fragment was high not only *in silico*, but also in the sedaDNA results, with a high number of OTUs assigned to species or genus level. The exact number of assigned OTUs, and the taxonomic level reached, was highly dependent on the database used, underlining the importance of reference sequence collections. In general, and particularly for a taxonomic group as highly diverse as fungi, limited availability of reference sequences is one of the most

crucial factors curbing the potential of metabarcoding communities from bulk samples. Moreover, resolving taxonomic diversity in bulk environmental samples may also be confounded due to continuously developing taxonomies and phylogenies and subsequent disagreement between databases regarding taxonomic levels and assignments, for example, the EMBL database still uses the former division Zygomycota which as per the current standard was split into the phyla Mucoromycota and Zoopagomycota (Spatafora et al., 2016), as implemented in the UNITE database.

Despite these persistent, inherent challenges to fungal metabarcoding, the observed diversity regarding the proportions of fungal divisions resembles the generally known proportions. The numbers of described species within the Ascomycota and Basidiomycota as the two largest divisions of the fungi kingdom exceed 60,000 and 30,000, respectively, whereas all other divisions comprise fewer than 2000 known species, each (Naranjo-Ortiz & Gabaldón, 2019); we identified 55 OTUs of Ascomycota, 61 of Basidiomycota, and 15 of Mucoromycota, suggesting that the primer pair shows good universality to capture fungal diversity. We also screened the *in silico* PCR results for specific groups that have previously been targeted, or have been prominent, in sediment core studies. These are aquatic taxa, and in particular, Chytridiomycota, which constituted a major component in the recently published study of Talas et al. (2021), as well as coprophilous fungi, such as *Sporormiella* and *Sordaria*, which are used in palynological studies as proxies for the presence of mammals (Gill et al. 2009, 2013, Davies 2019). These taxa occurred in the output of the *in silico* PCR and should thus potentially be retrieved by our assay.

4.3 | Potential of lake sediment fungal DNA for paleoecology

Fungi which are specific regarding their associations with plants or regarding other environmental conditions such as temperature may be indicative of local plant communities and climate, respectively. Based on the ecological conditions in the environment of the sampled lakes and on previous morphological and molecular studies in various ecosystems, we expected fungal taxa belonging to different ecological functional groups such as terrestrial saprotrophs, mycorrhizal fungi, coprophilous fungi, and aquatic taxa (Booth, 2011; Botnen et al., 2020; Grau et al., 2017) which may indicate ecosystem structures and environmental changes. Some examples of such ecosystem indicators are provided here; however, ecological interpretations of these data are made elsewhere (von Hippel et al., 2021).

The main fungal divisions were Ascomycota, Basidiomycota, and Mucoromycota, and we found relatively few reads (0.26%) of taxa which may be assigned to Glomeromycota, according to previous taxonomic systems, which is in line with the results of a recent metabarcoding study on fungal diversity from sediment of a lake in Eastern Latvia at 56.76°N, 27.15°E (Talas et al., 2021) with respect to terrestrial taxa. The overall composition differed between cores, which suggests local effects on past communities in the catchment of the respective lake; however, each of them seemed to reflect terrestrial communities. The 15 most abundant taxa in each lake comprised terrestrial saprotrophs and mycorrhizal fungi (e.g., the genera *Mortierella* and *Inocybe*, respectively), which supports the use of lake sediment for reconstructing terrestrial fungal communities. OTUs assigned to the genus *Mortierella* were among the 15 most abundant genera in each core, and these fungi typically occur as saprotrophs in soil, on decaying leaves, and other organic material (Domsch et al., 2007). Among mycorrhizal fungi, the genus *Inocybe* (Deacon et al., 1983) dominated some of the samples of lake CH12 and also occurred sporadically in samples of other lakes (von Hippel et al., 2021).

Contrary to our expectations, and in contrast to one of the main uses of fungal remains in paleoecology using classical, microscopic techniques, we observed no taxa which may be considered obligate coprophilous fungi (e.g., *Sporormiella* sp., *Preussia* sp., and *Sordaria* sp.). Spores of these taxa are used in morphological studies as proxies for herbivore presence (Davis & Shafer, 2006; Feranec et al., 2011; Raper & Bush, 2009). These taxa were recovered previously by Bellemain et al. (2013), using fungal metabarcoding on ancient permafrost; however, these deposits were of purely terrestrial origin. A large proportion of samples from one of the analyzed sites (Main River) also yielded mammal DNA (Willerslev et al., 2014), indicating a particularly high abundance of megafauna at this site. However, our data do not support the use of fungal DNA from lake sediment cores as a proxy for the presence and abundance of mammals in a similar way to microscopic analyses of spores. Future comparative morphological and molecular investigation may help elucidate the discrepancy between the approaches. In terms of lichens, which constitute an important part of Arctic vegetation and are crucial as the sustenance of herbivores (Kumpula, 2001), we found some 16,000 reads assigned to *Peltigera* sp., which occurred only in cores of Arctic lakes. Other common lichen taxa such as *Cladonia* sp. or *Stereocaulon* sp. were not observed, which may be due to the absence of DNA of these taxa in the sediment cores, absence of these taxa in the respective ecosystems, or biases inherent to ITS-1-based metabarcoding of fungi (Banchi et al., 2018; Tedersoo et al., 2015).

While the lack of coprophilous taxa in the molecular lake sediment records may be viewed as a drawback compared to permafrost deposits, their overall use in paleoecology seems more straightforward as they seem to be less confounded by potentially living organisms. Permafrost deposits contained a relatively high proportion of psychrophilic taxa (Bellemain et al., 2013), which are not necessarily ancient and which give no further paleoecological information in an area of cold ground temperatures. In our set of lakes, we found two OTUs assigned to the genus *Leucosporidium* which may be considered psychrotolerant. These were among the most abundant taxa in several samples of lake CH12 and showed considerable variation through the core (0%–12% of the respective read proportions; mean 3% ± 4%). Here, these sequences, thus, may be indicative of climatic changes and show a potential value as paleoenvironmental proxy.

In comparison to the recently published dataset on the temperate lake from Eastern Latvia (Talas et al., 2021), one striking difference

is that we observed relatively few (predominantly) aquatic taxa, for example, *Alatospora* sp. (3227 reads in 4 samples). This could reflect the actual situation in the Arctic and Subarctic lakes target here, as little is known on aquatic fungi in this part of the world. A previous study discovered a substantial amount of aquatic fungi in Scandinavian lakes (Khomich et al., 2017), however, it is also possible that the aquatic taxa that do occur in the areas of the current study are underrepresented in the current reference databases, and their sequences were, therefore, not assigned.

This points to an aspect that cannot be fully resolved at this point: comparing taxonomic diversity and relative abundance between studies is not straightforward due to a) differences between primers (and inevitable PCR bias), b) discrepancies between analysis pathways (including data filtration assignment algorithms, and comprehensiveness of reference databases), and c) differences between ecosystems and sampling substrates (e.g., Arctic vs. non-Arctic systems and lake sediments vs. permafrost or soil cores). We acknowledge that the chosen best identity threshold of 97% may be rather conservative and may thus not reveal the entire diversity of the examined fungal communities, which would bias our interpretations to some extent. However, due to the overwhelming diversity of fungi and the lack of complete reference databases, we argue that particularly for ecological conclusions which are based on relative abundances, a more stringent approach is likely more reliable than a more relaxed approach that would allow a larger proportion of false-positive assignments.

Taken together, our results support the use of sedimentary ancient DNA from lake sediments for reconstructing past fungal communities. The observed community changes may hold valuable information on ecosystem changes regarding the abundance of host plants of mycorrhizal or pathogenic fungi, climatic changes, and other ecological functions exerted by specific functional groups of fungi (for details, see von Hippel et al., 2021). Ecological interpretations, however, should be made with caution due to currently limited knowledge on the ecology and ecosystem functions of numerous fungal taxa. Moreover, comprehensive global reference databases must be established to reduce bias in the interpretation of such data. Further research is, thus, needed to link alterations in past and recent fungal communities and with changes on the scale of ecosystems, which may require further elucidating the ecology of crucial fungal taxa.

AUTHOR CONTRIBUTIONS

LSE conceived the study and PAS performed *in silico* and *in vitro* primer validation. BvH performed metabarcoding laboratory experiments. PAS and LSE conducted the analyses and wrote the manuscript. BvH, UL, HK, KS-L, and UH provided input for the analysis and interpretation of the metabarcoding data.

ACKNOWLEDGMENTS

This research was funded through the 2017–2018 Belmont Forum and BiodivERsA joint call for research proposals, under the BiodivScen ERA-Net COFUND program, and with the funding organizations Deutsche Forschungsgemeinschaft (DFG grant EP-98/3-1 to L.S. Epp), Agence Nationale de la Recherche (ANR), Research Council of Norway (NFR), the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (Formas), Academy of Finland, National Science Foundation (NSF), and the Natural Sciences and Engineering Research Council of Canada (NSERC-CRSNG). This study was supported by the ERC consolidator grant Glacial Legacy to U. Herzschuh (no. 772852). We thank the Natural History Museum of Oslo, Norway, particularly Gunnhild Marthinsen and Katriina Bendiksen, for providing vouchers for DNA extraction for the specificity tests and optimization of the primers. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

None of the authors have any conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings originate from von Hippel et al. 2021, and will be made accessible in a public repository, which is accessible through DOI: 10.1101/2021.11.05.465756.

ORCID

Peter Andreas Seeber https://orcid.org/0000-0003-4401-8248
Ulrike Löber https://orcid.org/0000-0001-7468-9531

REFERENCES

Abarenkov, K., Henrik Nilsson, R., Larsson, K.-H., Alexander, I. J., Eberhardt, U., Erland, S., Høiland, K., Kjøller, R., Larsson, E., Pennanen, T., Sen, R., Taylor, A. F. S., Tedersoo, L., Ursing, B. M., Vrålstad, T., Liimatainen, K., Peintner, U., & Kõljalg, U. (2010). The UNITE database for molecular identification of fungi - recent updates and future perspectives. *The New Phytologist*, *186*, 281–285.

Adamo, M., Voyron, S., Chialva, M., Marmeisse, R., & Girlanda, M. (2020). Metabarcoding on both environmental DNA and RNA highlights differences between fungal communities sampled in different habitats. *PLoS One*, *15*, e0244682.

Baldrian, P., Větrovský, T., Lepinay, C., & Kohout, P. (2021). High-throughput sequencing view on the magnitude of global fungal diversity. *Fungal Diversity*, 1–19.

Banchi, E., Stankovic, D., Fernández-Mendoza, F., Gionechetti, F., Pallavicini, A., & Muggia, L. (2018). ITS2 metabarcoding analysis complements lichen mycobiome diversity data. *Mycological Progress*, *17*, 1049–1066.

Bar-On, Y. M., Phillips, R., & Milo, R. (2018). The biomass distribution on earth. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 6506–6511.

Bellemain, E., Carlsen, T., Brochmann, C., Coissac, E., Taberlet, P., & Kauserud, H. (2010). ITS as an environmental DNA barcode for fungi: An in silico approach reveals potential PCR biases. *BMC Microbiology*, *10*, 189.

Bellemain, E., Davey, M. L., Kauserud, H., Epp, L. S., Boessenkool, S., Coissac, E., Geml, J., Edwards, M., Willerslev, E., Gussarova, G., Taberlet, P., & Brochmann, C. (2013). Fungal palaeodiversity revealed using high-throughput metabarcoding of ancient DNA from arctic permafrost. *Environmental Microbiology*, *15*, 1176–1189.

Blaalid, R., Kumar, S., Nilsson, R. H., Abarenkov, K., Kirk, P. M., & Kauserud, H. (2013). ITS1 versus ITS2 as DNA metabarcodes for fungi. *Molecular Ecology Resources*, *13*, 218–224.

Blackwell, M. (2011). The fungi: 1, 2, 3 … 5.1 million species? *American Journal of Botany*, *98*, 426–438.

Booth, T. (2011). Taxonomic notes on coprophilous fungi of the Arctic: Churchill, Resolute Bay, and Devon Island. *Canadian Journal of Botany*, *60*, 1115–1125.

Botnen, S. S., Thoen, E., Eidesen, P. B., Krabberød, A. K., & Kauserud, H. (2020). Community composition of arctic root-associated fungi mirrors host plant phylogeny. *FEMS Microbiology Ecology*, *96*, fiaa185.

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). Obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, *16*, 176–182.

Brundrett, M. (2004). Diversity and classification of mycorrhizal associations. *Biological Reviews of the Cambridge Philosophical Society*, *79*, 473–495.

Capo, E., Giguet-Covex, C., Rouillard, A., Nota, K., Heintzman, P. D., Vuillemin, A., Ariztegui, D., Arnaud, F., Belle, S., Bertilsson, S., & Bigler, C. (2021). Lake sedimentary DNA research on past terrestrial and aquatic biodiversity: Overview and recommendations. *Quaternary*, *4*, 6.

Chepstow-Lusty, A. J., Frogley, M. R., & Baker, A. S. (2019). Comparison of Sporormiella dung fungal spores and oribatid mites as indicators of large herbivore presence: Evidence from the Cuzco region of Peru. *Journal of Archaeological Science*, *102*, 61–70.

Clemmensen, K. E., Bahr, A., Ovaskainen, O., Dahlberg, A., Ekblad, A., Wallander, H., et al. (2013). Roots and associated fungi drive long-term carbon sequestration in boreal Forest. *Science (80-)*, *339*, 1615–1618.

Dabney, J., Meyer, M., & Paabo, S. (2013). Ancient DNA damage. *Cold Spring Harbor Perspectives in Biology*, *5*, a012567.

Davis, O. K., & Shafer, D. S. (2006). Sporormiella fungal spores, a palynological means of detecting herbivore density. *Palaeogeography Palaeoclimatology Palaeoecology*, *237*, 40–50.

Deacon, J. W., Donaldson, S. J., & Last, F. T. (1983). Sequences and interactions of mycorrhizal fungi on birch. *Plant and Soil*, *71*, 257–262.

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., et al. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*, 5872–5895.

Domaizon, I., Winegardner, A., Capo, E., Gauthier, J., & Gregory-Eaves, I. (2017). DNA-based methods in paleolimnology: New opportunities for investigating long-term dynamics of lacustrine biodiversity. *Journal of Paleolimnology*, *58*, 1–21.

Domsch, K. H., Gams, W., & Anderson, T. H. (2007). *Compendium of soil fungi, second edi*. IHW-Verlag.

Epp, L. S., Boessenkool, S., Bellemain, E., Haile, J., Esposito, A., Riaz, T., Erséus, C., Gusarov, V., Edwards, M. E., Johnsen, A., Stenøien, H. K., Hassel, K., Kauserud, H., Yoccoz, N. G., Bråthen, K. A., Willerslev, E., Taberlet, P., Coissac, E., & Brochmann, C. (2012). New environmental metabarcodes for analysing soil DNA: Potential for studying past and present ecosystems. *Molecular Ecology*, *21*, 1821–1833.

Epp, L. S., Kruse, S., Kath, N. J., Stoof-Leichsenring, K. R., Tiedemann, R., Pestryakova, L. A., & Herzschuh, U. (2018). Temporal and spatial patterns of mitochondrial haplotype and species distributions in Siberian larches inferred from ancient environmental DNA and modeling. *Scientific Reports*, *8*, 17436.

Feranec, R. S., Miller, N. G., Lothrop, J. C., & Graham, R. W. (2011). The Sporormiella proxy and end-Pleistocene megafaunal extinction: A perspective. *Quaternary International*, *245*, 333–338.

Ficetola, G., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., Taberlet, P., & Pompanon, F. (2010). An in silico approach for the evaluation of DNA barcodes. *BMC Genomics*, *11*, 434.

Finlay, R. D. (2008). Ecological aspects of mycorrhizal symbiosis: With special emphasis on the functional diversity of interactions involving the extraradical mycelium. *Journal of Experimental Botany*, *59*, 1115–1126.

Gardes, M., & Bruns, T. D. (1993). ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. *Molecular Ecology*, *2*, 113–118.

Grau, O., Geml, J., Pérez-Haase, A., Ninot, J. M., Semenova-Nelsen, T. A., & Peñuelas, J. (2017). Abrupt changes in the composition and function of fungal communities along an environmental gradient in the high Arctic. *Molecular Ecology*, *26*, 4798–4810.

Grigoriev, I. V. (2013). A changing landscape of fungal genomics. In F. Martin (Ed.), *The ecological genomics of fungi* (pp. 1–20). John Wiley & Sons.

Haile, J., Froese, D. G., MacPhee, R. D. E. E., Roberts, R. G., Arnold, L. J., Reyes, A. V., Rasmussen, M., Nielsen, R., Brook, B. W., Robinson, S., Demuro, M., Gilbert, M. T. P., Munch, K., Austin, J. J., Cooper, A., Barnes, I., Möller, P., & Willerslev, E. (2009). Ancient DNA reveals late survival of mammoth and horse in interior Alaska. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 22352–22357.

Hawksworth, D. L., & Lücking, R. (2017). Fungal diversity revisited: 2.2 to 3.8 million species. *Microbiology Spectrum*, *5*, 5–4.

Heeger, F., Bourne, E. C., Baschien, C., Yurkov, A., Bunk, B., Spröer, C., Overmann, J., Mazzoni, C. J., & Monaghan, M. T. (2018). Long-read DNA metabarcoding of ribosomal RNA in the analysis of fungi from aquatic environments. *Molecular Ecology Resources*, *18*, 1500–1514.

James, T. Y., Stajich, J. E., Hittinger, C. T., & Rokas, A. (2020). Toward a fully resolved fungal tree of life. *Annual Review of Microbiology*, *74*, 291–313.

Kanz, C., Aldebert, P., Althorpe, N., Baker, W., Baldwin, A., Bates, K., et al. (2005). The EMBL nucleotide sequence database. *Nucleic Acids Research*, *33*, D29.

Khomich, M., Davey, M. L., Kauserud, H., Rasconi, S., & Andersen, T. (2017). Fungal communities in Scandinavian lakes along a longitudinal gradient. *Fungal Ecology*, *27*, 36–46.

Klemm, J., Herzschuh, U., & Pestryakova, L. A. (2016). Vegetation, climate and lake changes over the last 7000 years at the boreal treeline in north-Central Siberia. *Quaternary Science Reviews*, *147*, 422–434.

Kumpula, J. (2001). Winter grazing of reindeer in woodland lichen pasture. *Small Ruminant Research*, *39*, 121–130.

Lammers, Y., Heintzman, P. D., & Alsos, I. G. (2021). Environmental palaeogenomic reconstruction of an ice age algal population. *Communications Biology*, *4*, 220.

Lücking, R., Aime, M. C., Robbertse, B., Miller, A. N., Ariyawansa, H. A., Aoki, T., et al. (2020). Unambiguous identification of fungi: Where do we stand and how accurate and precise is fungal DNA barcoding? *IMA Fungus*, *11*, 1–32.

Lydolph, M. C., Jacobsen, J., Arctander, P., Gilbert, M. T. P., Gilichinsky, D. A., Hansen, A. J., Willerslev, E., & Lange, L. (2005). Beringian paleoecology inferred from permafrost-preserved fungal DNA. *Applied and Environmental Microbiology*, *71*, 1012–1017.

Mercier, C., Boyer, F., Bonin, A., & Coissac, E. (2013). SUMATRA and SUMACLUST: Fast and exact comparison and clustering of sequences. In *Programs and abstracts of the seqBio 2013 workshop. Abstract* (pp. 27–29).

Naranjo-Ortiz, M. A., & Gabaldón, T. (2019). Fungal evolution: Diversity, taxonomy and phylogeny of the fungi. *Biological Reviews*, *94*, 2101–2137.

Nilsson, R. H., Anslan, S., Bahram, M., Wurzbacher, C., Baldrian, P., & Tedersoo, L. (2019). Mycobiome diversity: High-throughput sequencing and identification of fungi. *Nature Reviews Microbiology*, *17*(2), 95–109.

Oksanen, A. J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P. R., Hara, R. B. O., Simpson, G. L., Solymos, P., Stevens, M. H. H., & Szoecs, E. (2020) Community Ecology Package vegan, 2.5–7.

Pääbo, S. (1989). Ancient DNA: Extraction, characterization, molecular cloning, and enzymatic amplification. *Proceedings of the National Academy of Sciences of the United States of America*, *86*, 1939–1943.

Parducci, L., Bennett, K. D., Ficetola, G. F., Alsos, I. G., Suyama, Y., Wood, J. R., & Pedersen, M. W. (2017). Ancient plant DNA in lake sediments. *The New Phytologist*, *214*, 924–942.

Pedersen, M. W., Ruter, A., Schweger, C., Friebe, H., Staff, R. A., Kjeldsen, K. K., Mendoza, M. L. Z., Beaudoin, A. B., Zutter, C., Larsen, N. K., Potter, B. A., Nielsen, R., Rainville, R. A., Orlando, L., Meltzer, D. J., Kjær, K. H., & Willerslev, E. (2016). Postglacial viability and colonization in North America's ice-free corridor. *Nature*, *537*, 45–49.

Powell, J. R., & Rillig, M. C. (2018). Biodiversity of arbuscular mycorrhizal fungi and ecosystem function. *The New Phytologist*, *220*, 1059–1075.

R Development Core Team. (2019) *A language and environment for statistical computing*, R Foundation for Statistical Computing (ed) : R Foundation for Statistical Computing.

Raper, D., & Bush, M. (2009). A test of Sporormiella representation as a predictor of megaherbivore presence and abundance. *Quaternary Research*, *71*, 490–496.

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., & Coissac, E. (2011). ecoPrimers: Inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, *39*, e145–e145.

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

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., Chen, W. & Fungal Barcoding Consortium. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for fungi. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 6241–6246.

Smith, S., & Read, D. (2008). *Mycorrhizal symbiosis* (3rd ed.). Elsevier.

Spatafora, J. W., Chang, Y., Benny, G. L., Lazarus, K., Smith, M. E., Berbee, M. L., Bonito, G., Corradi, N., Grigoriev, I., Gryganskyi, A., James, T. Y., O'Donnell, K., Roberson, R. W., Taylor, T. N., Uehling, J., Vilgalys, R., White, M. M., & Stajich, J. E. (2016). A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia*, *108*, 1028–1046.

Stielow, J. B., Lévesque, C. A., Seifert, K. A., Meyer, W., Irinyi, L., Smits, D., Renfurm, R., Verkley, G. J. M., Groenewald, M., Chaduli, D., Lomascolo, A., Welti, S., Lesage-Meessen, L., Favel, A., Al-Hatmi, A. M. S., Damm, U., Yilmaz, N., Houbraken, J., Lombard, L., … Robert, V. (2015). One fungus, which genes? Development and assessment of universal primers for potential secondary fungal DNA barcodes. *Persoonia - Molecular Phylogeny and Evolution of Fungi*, *35*, 242–263.

Talas, L., Stivrins, N., Veski, S., Tedersoo, L., & Kisand, V. (2021). Sedimentary ancient DNA (sedaDNA) reveals fungal diversity and environmental drivers of community changes throughout the Holocene in the present boreal Lake Lielais Svetinu (eastern Latvia). *Microorganisms*, *9*, 719.

Taylor, T. N., & Osborn, J. M. (1996). The importance of fungi in shaping the paleoecosystem. *Review of Palaeobotany and Palynology*, *90*, 249–262.

Tedersoo, L., Anslan, S., Bahram, M., Põlme, S., Riit, T., Liiv, I., Kõljalg, U., Kisand, V., Nilsson, H., Hildebrand, F., Bork, P., & Abarenkov, K. (2015). Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi. *MycoKeys*, *10*, 1–43.

Tedersoo, L., & Lindahl, B. (2016). Fungal identification biases in microbiome projects. *Environmental Microbiology Reports*, *8*, 774–779.

van der Heijden, M. G. A., Klironomos, J. N., Ursic, M., Moutoglis, P., Streitwolf-Engel, R., Boller, T., Wiemken, A., & Sanders, I. R. (1998). Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature*, *396*, 69–72.

Varma, A., Prasad, R., & Tuteja, N. (Eds.). (2017). *Mycorrhiza – function, diversity, state of the art*. Springer International Publishing.

Vilgalys, R., & Gonzalez, D. (1990). Organization of ribosomal DNA in the basidiomycete Thanatephorus praticola. *Current Genetics*, *18*, 277–280.

von Hippel, B., Stoof-Leichsenring, K. R., Schulte, L., Seeber, P. A., Epp, L. S., Biskaborn, B. K., Diekmann, B., Melles, M., Pestryakova, L. A., & Herzschuh, U. (2021). Long-term fungus–plant co-variation from multi-site sedimentary ancient DNA metabarcoding in Siberia. *bioRxiv*. https://doi.org/10.1101/2021.11.05.465756

White, T. J., Bruns, T., Lee, S., & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In M. Innis & D. Gelfand (Eds.), *PCR Protocols* (pp. 315–322). Elsevier.

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, *40*, 1–29.

Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M. E., Lorenzen, E. D., Vestergård, M., Gussarova, G., Haile, J., Craine, J., Gielly, L., Boessenkool, S., Epp, L. S., Pearman, P. B., Cheddadi, R., Murray, D., Bråthen, K. A., Yoccoz, N., … Taberlet, P. (2014). Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature*, *506*, 47–51.

Willerslev, E., Hansen, A. J., Binladen, J., Brand, T. B., Gilbert, M. T. P., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D. A., & Cooper, A. (2003). Diverse plant and animal genetic records from holocene and pleistocene sediments. *Science (80-)*, *300*, 791–795.

Wood, J. R., & Wilmshurst, J. M. (2013). Accumulation rates or percentages? How to quantify Sporormiella and other coprophilous fungal spores to detect late quaternary megafaunal extinction events. *Quaternary Science Reviews*, *77*, 1–3.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Seeber, P. A., von Hippel, B., Kauserud, H., Löber, U., Stoof-Leichsenring, K. R., Herzschuh, U., & Epp, L. S. (2022). Evaluation of lake sedimentary ancient DNA metabarcoding to assess fungal biodiversity in Arctic paleoecosystems. *Environmental DNA*, *4*, 1150–1163. https://doi.org/10.1002/edn3.315