Long-term bacteria-fungi-plant associations in permafrost soils inferred from palaeometagenomics

Barbara von Hippel

Univ.-Diss.

zur Erlangung des akademischen Grades

"doctor rerum naturalium"

(Dr. rer. nat.)

in der Wissenschaftsdisziplin „Ökologie“

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät

Institut für Biochemie und Biologie

der Universität Potsdam

angefertigt am

Alfred-Wegener-Institut

Helmholtz-Zentrum für Polar- und Meeresforschung

Datum der Einreichung: 29. Juni 2023

Hauptbetreuerin: Prof. Dr. Ulrike Herzschuh, Universität Potsdam

Weitere Gutachter: Prof. Dr. Susanne Liebner (Geoforschungszentrum Potsdam)
Prof. Dr. Georg Guggenberger (Universität Hannover)

Published online on the
Publication Server of the University of Potsdam:
https://doi.org/10.25932/publishup-63600
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-636009

“Essentially, all life depends upon the soil ... There can be no life without soil and no soil without life; they have evolved together.”

(Charles E. Kellogg, USDA Yearbook of Agriculture, 1938)

Summary

The arctic is warming 2 – 4 times faster than the global average, resulting in a strong feedback on northern ecosystems such as boreal forests, which cover a vast area of the high northern latitudes. With ongoing global warming, the treeline subsequently migrates northwards into tundra areas. The consequences of turning ecosystems are complex: on the one hand, boreal forests are storing large amounts of global terrestrial carbon and act as a carbon sink, dragging carbon dioxide out of the global carbon cycle, suggesting an enhanced carbon uptake with increased tree cover. On the other hand, with the establishment of trees, the albedo effect of tundra decreases, leading to enhanced soil warming. Meanwhile, permafrost thaws, releasing large amounts of previously stored carbon into the atmosphere. So far, mainly vegetation dynamics have been assessed when studying the impact of warming onto ecosystems. Most land plants are living in close symbiosis with bacterial and fungal communities, sustaining their growth in nutrient poor habitats. However, the impact of climate change on these subsoil communities alongside changing vegetation cover remains poorly understood. Therefore, a better understanding of soil community dynamics on multi millennial timescales is inevitable when addressing the development of entire ecosystems. Unravelling long-term cross-kingdom dependencies between plant, fungi, and bacteria is not only a milestone for the assessment of warming on boreal ecosystems. On top, it also is the basis for agriculture strategies to sustain society with sufficient food in a future warming world.

The first objective of this thesis was to assess ancient DNA as a proxy for reconstructing the soil microbiome (Manuscripts I, II, III, IV). Research findings across these projects enable a comprehensive new insight into the relationships of soil microorganisms to the surrounding vegetation. First, this was achieved by establishing (Manuscript I) and applying (Manuscript II) a primer pair for the selective amplification of ancient fungal DNA from lake sediment samples with the metabarcoding approach. To assess fungal and plant co-variation, the selected primer combination (ITS67, 5.8S) amplifying the ITS1 region was applied on samples from five boreal and arctic lakes. The obtained data showed that the establishment of fungal communities is impacted by warming as the functional ecological groups are shifting. Yeast and saprotroph dominance during the Late Glacial declined with warming, while the abundance of mycorrhizae and parasites increased with warming. The overall species richness was also alternating. The results were compared to shotgun sequencing data reconstructing fungi and bacteria (Manuscripts III, IV), yielding overall comparable results to the metabarcoding approach. Nonetheless, the comparison

also pointed out a bias in the metabarcoding, potentially due to varying ITS lengths or copy numbers per genome.

The second objective was to trace fungus-plant interaction changes over time (Manuscripts II, III). To address this, metabarcoding targeting the ITS1 region for fungi and the chloroplast P6 loop for plants for the selective DNA amplification was applied (Manuscript II). Further, shotgun sequencing data was compared to the metabarcoding results (Manuscript III). Overall, the results between the metabarcoding and the shotgun approaches were comparable, though a bias in the metabarcoding was assumed. We demonstrated that fungal shifts were coinciding with changes in the vegetation. Yeast and lichen were mainly dominant during the Late Glacial with tundra vegetation, while warming in the Holocene lead to the expansion of boreal forests with increasing mycorrhizae and parasite abundance. Aside, we highlighted that Pinaceae establishment is dependent on mycorrhizal fungi such as Suillineae, Inocybaceae, or *Hyaloscypha* species also on long-term scales.

The third objective of the thesis was to assess soil community development on a temporal gradient (Manuscripts III, IV). Shotgun sequencing was applied on sediment samples from the northern Siberian lake Lama and the soil microbial community dynamics compared to ecosystem turnover. Alongside, podzolization processes from basaltic bedrock were recovered (Manuscript III). Additionally, the recovered soil microbiome was compared to shotgun data from granite and sandstone catchments (Manuscript IV, Appendix). We assessed if the establishment of the soil microbiome is dependent on the plant taxon and as such comparable between multiple geographic locations or if the community establishment is driven by abiotic soil properties and as such the bedrock area. We showed that the development of soil communities is to a great extent driven by the vegetation changes and temperature variation, while time only plays a minor role. The analyses showed general ecological similarities especially between the granite and basalt locations, while the microbiome on species-level was rather site-specific. A greater number of correlated soil taxa was detected for deep-rooting boreal taxa in comparison to grasses with shallower roots. Additionally, differences between herbaceous taxa of the late Glacial compared to taxa of the Holocene were revealed.

With this thesis, I demonstrate the necessity to investigate subsoil community dynamics on millennial time scales as it enables further understanding of long-term ecosystem as well as soil development processes and such plant establishment. Further, I trace long-term processes leading to podzolization which supports the development of applied carbon capture strategies under future global warming.

Deutsche Zusammenfassung

Die Arktis erwärmt sich 2 – 4 mal schneller als der weltweite Durchschnitt, was einen starken Einfluss auf Ökosysteme wie die borealen Nadelwälder hat, die ein großes Gebiet der hohen nördlichen Breiten ausmachen. Durch die globale Erwärmung verschiebt sich die Baumgrenze nach Norden in Gebiete mit derzeitigem Tundra-Wachstum. Die Auswirkungen sind komplex: einerseits speichern die Wälder große Mengen an globalem terrestrischem Kohlenstoff und entziehen dem globalen Kohlenstoff-Kreislauf CO_2. Größere Baumbestände führen so vermutlich zu einer verstärkten Kohlenstoffaufnahme in Böden. Andererseits verringert sich der Albedo-Effekt von Tundra durch das Baumwachstum, was zu einer verstärkten Erwärmung des Bodens führt. Durch die Erwärmung taut der Permafrost, was große Mengen an gespeichertem Kohlenstoff freisetzt, der in die Atmosphäre gelangt. Bislang wurden vor allem Vegetations-Dynamiken unter Einfluss der Erderwärmung auf Ökosysteme untersucht. Die meisten Landpflanzen leben in Symbiose mit Bakterien und Pilzen, die deren Wachstum in nährstoffarmen Lebensräumen ermöglichen. Allerdings ist bislang wenig verstanden, wie diese Bodengemeinschaften durch Langzeit-Klimawandel und sich verändernde Vegetation beeinflusst werden. Deshalb ist es notwendig, bei der Rekonstruktion von Ökosystemdynamiken auch die Gemeinschaft von Bodenmikroorganismen auf Zeitskalen von Jahrzehntausenden zu betrachten. Das Verstehen von Langzeitabhängigkeiten zwischen Pflanzen, Pilzen und Bakterien ist nicht nur ein Meilenstein bei der Untersuchung von Klimawandel auf boreale Nadelwälder. Zusätzlich ermöglicht es, angepasste Strategien im Bereich der Landwirtschaft zu entwickeln, um die wachsende Bevölkerung auch in Zukunft mit ausreichend Nahrungsmitteln versorgen zu können.

Das erste Ziel der Arbeit war es, das Potential von *ancient* DNA für die Rekonstruktion des Bodenmikrobioms zu untersuchen (Manuskripte I, II, III, IV). Die Forschungsergebnisse aller Projekte ermöglichen einen umfassenden neuen Einblick in die Beziehungen von Bodenmikroorganismen zur umgebenden Vegetation. Zunächst wurde ein Primerpaar für die gezielte Amplifikation von Pilzen aus *ancient DNA* aus Seesedimenten etabliert (Manuskript I) und angewendet (Manuskript II). Um Zusammenhänge zwischen dem Auftreten von Pilzen und Pflanzen zu beurteilen, wurde mit der ausgewählten Primerkombination (ITS67, 5.8S), welche die ITS1-Region amplifiziert, die DNA von Proben aus fünf borealen und arktischen Seen vervielfältigt. Die Zusammensetzung der Pilzgemeinschaften wurde durch die Erwärmung beeinflusst. Während des Spätglazials dominierten die funktionellen ökologischen Gruppen von Hefen und Saprotrophen, während mit Erwärmung die Verbreitung von Mykorrhiza und parasitischen Pilzen zunahm. Auch der Artenreichtum war von der Erwärmung betroffen. Die Ergebnisse wurden mit

s*hotgun* Sequenzierdaten zur Rekonstruktion von Pilzen und Bakterien (Manuskripte III, IV) verglichen. Insgesamt waren die Ergebnisse der beiden Ansätze vergleichbar. Allerdings zeigte der Vergleich auch eine Verzerrung im *metabarcoding* auf, die möglicherweise auf unterschiedliche Längen der ITS-Region bei verschiedenen Arten oder variierende Anzahlen an ITS-Kopien pro Genom zurückzuführen ist.

Das zweite Ziel war es, Pilz-Pflanzen-Interaktionen im zeitlichen Wandel zu verfolgen (Manuskript II, III). Dafür wurde zunächst *metabarcoding* für die gezielte Amplifikation von Pilz- und Pflanzen-DNA angewendet (Manuskript II). Die *metabarcoding* Ergebnisse wurden anschließend mit *shotgun* Sequenzierdaten verglichen (Manuskript III). Alles in allem waren die Ergebisse beider Sequenzieransätze vergleichbar, wobei dennoch von einer Verzerrung im *metabarcoding* Ansatz ausgegangen wird. Ein Zusammenhang zwischen Dynamiken in den Pilzgesellschaften und der Vegetation wurde gezeigt. Mit der Tundravegetation während des Spätglazials dominierten Hefen und Flechten, während die Erwärmung im Holozän zur Ausbreitung borealer Wälder mit zunehmender Mykorrhiza- und Parasitenhäufigkeit führte. Zusätzlich wurde gezeigt, dass die Etablierung von Pinaceen auch langfristig von Mykorrhizapilzen wie Suillineae, Inocybaceae oder *Hyaloscypha*-Arten abhängt.

Das dritte Ziel der Arbeit war es, zeitliche Dynamiken in der Zusammensetzung von Bodenorganismen im Allgemeinen sowie die Entstehung von Böden im Speziellen auf Zeitskalen von Jahrzehntausenden zu untersuchen (Manuskript III, IV). Sedimentproben des nordsibirischen Lama-Sees wurden mittels *shotgun* Sequenzierung analysiert und die Pilz- und Bakteriengemeinschaften mit dem generellen Wandel des Ökosystems verglichen. Parallel dazu wurde die Entstehung von Podsol aus dem Basalt-Grundgestein rekonstruiert (Manuskript III). Das rekonstruierte Bodenmikrobiom wurde mit weiteren *shotgun* Datensätzen aus Granit und Sandstein Einzugsgebieten verglichen (Manuskript IV, Appendix). Es wurde untersucht, ob die Entstehung des Bodenmikrobioms abhängig von der Pflanzenart und somit vergleichbar zwischen geografischen Standorten ist oder ob abiotische Bodeneigenschaften und somit das Grundgestein die Zusammensetzung bestimmt (Manuskript IV, Appendix). Es wurde gezeigt, dass die Entwicklung der Bodengemeinschaften zu einem großen Teil von der Vegetation und Temperaturänderungen beeinflusst wurde, während die Zeit seit Gletscherrückgang eine eher geringe Rolle spielte. Die Analysen zeigten generelle ökologische Ähnlichkeiten vor allem an den Standorten mit Basalt sowie Granit als Grundgestein, während das Mikrobiom auf Art-Ebene eher ortsspezifisch war. Eine größere Anzahl von korrelierenden Bodenorganismen wurde in tieferwurzelnden Nadelbäumen im Vergleich zu flachwurzelnden Gräsern entdeckt, sowie Unterschiede zwischen Kräutern des Spätglazials im Vergleich zu Kräutern des Holozäns gezeigt.

Mit dieser Dissertation zeige ich die Notwendigkeit, die Dynamik von Bodenorganismen während mehrerer Jahrzehntausende zu untersuchen, da so ein besseres Verständnis von Langzeit-Ökosystem- sowie Bodenentstehungsprozessen ermöglicht wird. Zusätzlich wird durch ein besseres Verständnis von Podsolisierungsmechanismen die Entwicklung angewandter *carbon capture* Strategien vereinfacht.

Table of Contents

1 Introduction

1.1 Arctic ecosystems under global warming

The Arctic is warming 2-4 times faster than the global average (Solomon et al., 2007), a phenomenon commonly known as arctic amplification. The temperature increase results in Arctic greening with increasing plant biomass, height, cover, and abundance (Forbes et al., 2010; Myers-Smith et al., 2011; Elmendorf et al., 2012), while prolonged growing seasons lead to treeline shifts northwards and changing tundra vegetation (Zhang et al., 2013). Further impacts are permafrost thawing (Brown and Romanovsky, 2008) as well as more frequent and drastic forest fires (Flannigan et al., 2009).

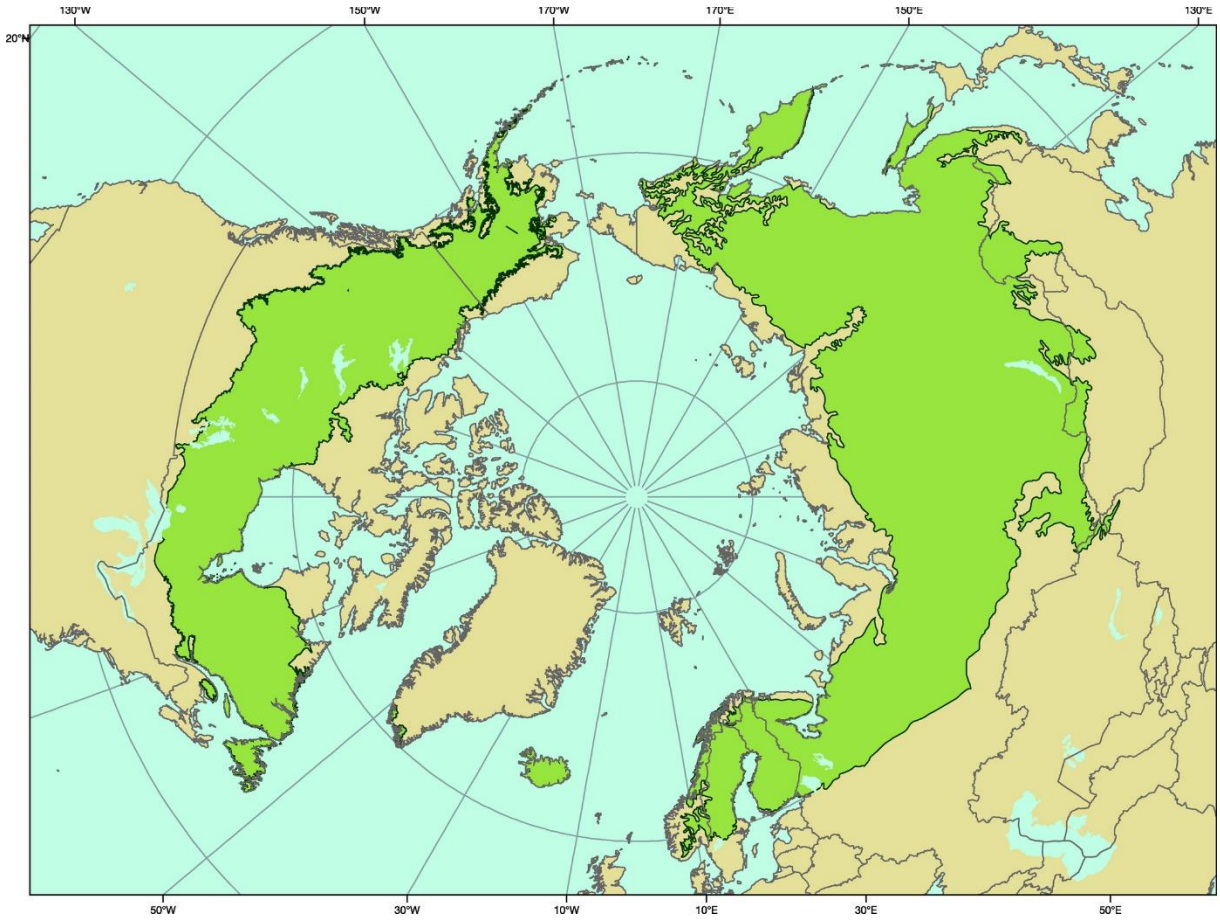


Figure 1: Distribution area of the circumboreal zone (Brandt et al., 2013).

Boreal forests stretch over vast areas in the high-latitudes of Alaska, Canada, and Russia (Potapov et al., 2008; Fig. 1), encompassing around 30 % of the global forest biome (Brandt et al., 2013), with approximately a third of the forests being underlain by permafrost (Zimov et al., 2006). These woodlands store more than 30 % of global terrestrial carbon (Kasischke, 2000), making them an important carbon sink (Pan et al., 2011; Bradshaw and Warkentin, 2015). Due to the typically harsh growing conditions in their distribution areas and decreasing diversity with increasing latitude, these

forests are usually low in species richness (Qian et al., 1998; Hart and Chen, 2006). Boreal forests are dominated by the coniferous taxa *Larix*, *Pinus*, *Picea*, and *Abies*. *Larix* and *Pinus sylvestris* dominated forests are forming the so-called light taiga, while forests dominated by *Abies* and *Picea* species, and *Pinus banksiana* are called dark taiga with all being evergreen conifers (Tautenhahn et al., 2016). The common soil type below these forests is podzol which is characterised by a low pH and the tendency for further acidification (Wiklander and Andersson, 1972). The release of aluminium and iron ions from rocks is induced by organic acids during the soil development, resulting in the formation of chelates with organic matter. Leaching of these complexes from the upper mineral horizons and subsequent deposition in the subsoil induces bleaching of the upper horizons, while leading to the characteristic reddish-brown colour of the subsoil (Lundström et al., 2000; Sauer et al., 2007). The transition zone from the boreal forest towards Arctic tundra vegetation is known as tundra-taiga ecotone. With ongoing warming, the tundra-taiga ecotone is expected to migrate further northwards into tundra areas (Grace et al., 2002; Shevtsova et al., 2020; Kruse and Herzschuh, 2022). This invasion of coniferous boreal taxa into tundra will drastically impact the climate system as carbon storage increases and albedo decreases (Chapin et al., 2005; Pearson et al., 2013). Also, nutrient cycling is altered (Van Cleve and Viereck, 1981).

The typical tundra vegetation is comprised of dwarf shrubs, forbs, graminoids, and non-vascular plants (Walker et al., 2005). An increase in shrubby taxa with mid-Holocene warming has been reconstructed (Bigelow, 2003). This shrubification has a major impact on the carbon balance of tundra ecosystems by enhancing the carbon uptake (Walker et al., 2006; Myers-Smith et al., 2011) and additionally altering ecosystem respiration, subsequently affecting the nutrient cycling (DeMarco et al., 2014; Wang et al., 2018) and soil carbon stocks (Mcguire et al., 2010). Under warming, further expansion of shrubs also facilitates the invasion of tree taxa in tundra areas, leading to forest establishment (Edwards et al., 2005; Holmgren et al., 2015). The major ecological consequences of taiga invasion into tundra areas are far from being fully understood and an assessment of not only the vegetation cover, but other environmental factors such as their subsoil microbial communities, is necessary.

1.2 The plant-associated microbiome

The rhizosphere of plants is a dynamic micro-biosphere around the roots, containing a diverse range of microorganisms sustaining plant growth and survival (Hiltner, 1904; Sharma et al., 2020). Carbon release from the roots, termed as rhizodeposition, generates the so-called rhizosphere effect (Whipps and Lynch, 1985). Many different types of substrates are being secreted by the roots: water soluble exudates, such as amino acids, vitamins, or sugars derived from leaking without metabolic energy

involved; secretions of polymeric carbohydrates dependent on metabolic processes; lysates deriving from cell autolysis; gases such as CO_2 (Lynch and Whipps, 1990). The exuded compounds act as chemical attractants for soil microorganisms (Pandey et al., 2017) and, by changing the physico-chemical properties of soil, also determine the composition of the microbial communities (Yang et al., 2009). Plant-growth promoting bacteria facilitate nutrient and water uptake (Ahkami et al., 2017) and reduce the impact of abiotic stressors by - amongst other factors - increasing salt tolerance (Vacheron et al., 2013). Furthermore, mycorrhizal interactions between fungi and root tips enable plant-growth in nutrient poor habitats by providing the plant with mainly phosphate and nitrate, but also other organically bound nutrients (Read and Perez-Moreno, 2003). Based on the structure and function, a distinction is made between ecto, orchid, ericoid and arbuscular mycorrhizae (van der Heijden et al., 2015). Ectomycorrhizal fungal hyphae are not directly penetrating the plant cells, but forming a mantle-structure around the root tip, while hyphae in the cortex are forming the Hartig net around the plant cells, enabling the nutrient exchange (Smith and Read, 2010). In contrast, endomycorrhizae, such as orchid, ericoid, and arbuscular mycorrhizae, are directly penetrating the plant cells. In arbuscular mycorrhizal symbiosis, which are only formed by fungi of the Glomeromycota, hyphae are forming tree-shaped subcellular structures called arbuscules in the plant cells, predominantly enabling nutrient exchange (Parniske, 2008). Most land plants are dependent on mycorrhizal associations, with the majority forming arbuscular mycorrhizae (Brundrett, 2009).

Considering the well-known importance of subsoil microorganisms for plant establishment, relatively little research targeting natural long-term processes behind the establishment of a healthy root microbiome from a palaeo-perspective has been done. As a major question, it remains unclear how the population of root-associated microorganisms has changed in relation to long-term warming and plant establishment.

1.3 Drivers of soil development

The formation of soil is termed “pedogenesis” and dependent on the mineral composition of the parent material, time, climate, topography, as well as living organisms (Delgado-Baquerizo et al., 2020; Franzetti et al., 2020). Initial to pedogenesis, bare rocks are exposed to the surface by warming-induced glacier retreat and are subsequently being weathered and torn into soil. A variety of processes are leading to the weathering of the rocks.

Physical weathering describes the fragmentation of rocks without changes in their chemical composition caused by frost blasting, thermal stressors, penetrating plant roots or salt cristallization (Hack, 2020), resulting in initial rock-breakdown. In contrast, chemical weathering processes lead to

compositional changes of the rocks, converting primary minerals to secondary minerals through hydrolysis, oxidation or carbonation (Stonestrom et al., 1998; Viers et al., 2014). Chemical weathering can also lead to the complete dissolution of minerals, usually involving water, leading to the release of dissolved ions which can be taken up by living organisms or are lost in stream water (Schlesinger and Bernhardt, 2013).

In addition, multiple biogenic factors such as plant root exudates, fungi, and bacteria can drive weathering processes. Plants increase weathering by releasing organic acids into the soil as metabolic side products, which further react with mineral compounds and thus lead to the weakening and dissolving of the rocks (Kelly et al., 1998). Mycorrhizal fungi are spreading their hyphae inside cracks in stones (Hoffland et al., 2002) and release siderophores as well as low-molecular weight organic acids which further enhance weathering (Hoffland et al., 2004). The efflux of H^+ protons due to excess ammonium uptake results in soil acidification (Hoffland et al., 2004). Besides fungi and plants, bacteria also increase weathering by oxidation and reduction reactions. For example, the oxidation of iron-bearing minerals results in the release of sulfuric acid (Sasaki et al., 1998). As another example, the chelation of aluminium and iron ions with siderophores enhances the solubility and mobility of those ions in the environment (Welch et al., 2002; Uroz et al., 2009).

So far, the degree of contribution of the single processes to soil formation remains unclear. Also, it is still debated whether long-term soil development is driven mainly by the time passed by since initial deglaciation, yielding it a so-called trajectory. The importance of further environmental biotic and abiotic factors during pedogenesis is still under discussion.

1.4 Ancient DNA to unravel past ecosystems

1.4.1 Lake sediments as archives of past community changes

Looking into the past for gaining a better understanding of future warming-induced ecosystem changes is a promising emerging approach. One method to do so is by assessing sedimentary ancient DNA (sedaDNA). Lake sediments represent a unique archive for the analyses of such past community changes by incorporating and storing the information on their surrounding environment over long periods of time (e.g. Epp et al., 2015; Alsos et al., 2018). Terrestrial DNA derived from vertebrate urine and faeces (Andersen et al., 2012), soil organisms (Kisand et al., 2018), or plant remains (e.g. Parducci et al., 2013; Alsos et al., 2016) binds onto soil particles and, through erosion, gets continuously transported in the lake where it deposits at the bottom. Throughout time, the sediment accumulates, storing allochthonous signals derived from the catchment or beyond, also including substantial

amounts of terrestrial DNA, as well as autochthonous signals derived from the biological production within the lake or chemical precipitation (Giguet-Covex et al., 2019). The retrieval of sediment cores from the lake bottom enables the reconstruction of past terrestrial environmental changes assessing the sedaDNA (Parducci et al., 2017) or pollen grains (e.g. Herzschuh et al., 2004; Pedersen et al., 2013). This sedaDNA is relatively well preserved in the lakes as thermal stratification and the temperature of approximately 4 °C on the bottom of the lakes favour the development of anoxia (Parducci et al., 2017). Besides, further degradation is inhibited when extracellular DNA from lysed cells binds to charged minerogenic or organic particles, as immediately nuclease degradation gets prevented (Pietramellara et al., 2009).

By accelerator mass spectrometry (AMS) derived radiocarbon dating of the sediment cores or macrofossils, a determination of the age-depth relation of the cores is enabled (Andree et al., 1986; MacDonald et al., 1991). This allows the relation of ecosystem turnover to large-scale environmental changes such as the Bølling-Allerød warm period or the Pleistocene-Holocene transition. Advantageous over other archives such as permafrost is the continuous sedimentation of particles in the lakes (Melles et al., 2022), enabling the reconstruction of chronological community changes. To verify the ancient origin of the sedimentary DNA, methods such as mapDamage can be applied (Jónsson et al., 2013). This includes assessing typical damage patterns of the DNA, such as the deamination of cytosines (Briggs et al., 2007) and the increasing C to T substitution at the read ends (Meyer et al., 2012). In comparison to experimental warming, the understanding of multiple tens of thousands of years in ecology trends alongside natural warming conditions allows a more precise assessment of long-term ecosystem adaptation mechanisms and the prediction of future trends.

1.4.2 Metabarcoding for targeting specific communities

Until recently, sedimentary reconstructions of terrestrial ecosystem changes focused on the analysis of pollen grains, spores or macrofossils for assessing vegetation dynamics (e.g. Van Geel, 2001; Herzschuh et al., 2004; Pedersen et al., 2013). Fungi, however, possess only very limited fossil remains, yielding very little information deriving from non-pollen palynomorphs (Loughlin et al., 2018; Quamar and Stivrins, 2021). A more recent alternative to macrofossils, pollen, or non-pollen palynomorphs is high-throughput DNA sequencing, being advantageous as it often enables the identification of more species as well as the identification at higher taxonomic resolutions (Sønstebø et al., 2010; Parducci et al., 2015).

DNA metabarcoding is a method that can be used when addressing compositional changes in a specific taxonomic group. It is based on the PCR amplification of a short, highly variable region of the genome

which is flanked by highly conserved regions (Taberlet et al., 2007; Epp et al., 2012). The choice of the targeted DNA region depends on the selected taxonomic group, but requires to be relatively stable, conserved amongst the organisms within the taxonomic group, and short due to the high fragmentation and degradation of the DNA (Epp et al., 2012). So far, sedaDNA studies mostly targeted terrestrial vegetation changes using the trnL P6 loop marker (Taberlet et al., 2007; Parducci et al., 2017; Alsos et al., 2018; Liu et al., 2020). The advantage of this marker is that with its relatively short length of 10 – 143 bp, it is suitable to amplify highly degraded ancient DNA reads (Taberlet et al., 2007).

In contrast to vegetation studies on sedaDNA, the recovery of fungal communities so far is limited, mostly as from an assumption of 1.5 – 5 mio fungal species on earth, only around 200,000 species have been classified so far, yielding uncomplete databases, impeding the data interpretation. Nonetheless, samples derived from permafrost (Bellemain et al., 2013) or lake sediment (Talas et al., 2021) have recently been assessed. For the amplification of fungal DNA, the internal transcribed spacer (ITS) region is commonly used (White et al., 1990; Schoch et al., 2012). Located between the 18S (small subunit) and 28S (large subunit) rRNA genes, the variable spacer ITS1 and ITS2 and the intercalarly, highly conserved 5.8S form the entire ITS region (Schoch et al., 2012). The length of the ITS region is highly dependent on the fungal lineage with an average of 550 base-pairs (Feibelman et al., 1994; Schoch et al., 2014). As ancient DNA molecules are highly fragmented and relatively short, the amplification with most primer pairs for fungal metabarcoding is challenging, as sequence length of around 250-400 nucleotides are being targeted (Beeck et al., 2014; Tedersoo et al., 2015). A marker suitable for the short ancient DNA fragments is still lacking. Therefore, the need for a revised primer pair for the amplification of shorter DNA fragments emerged (Fig. 2).

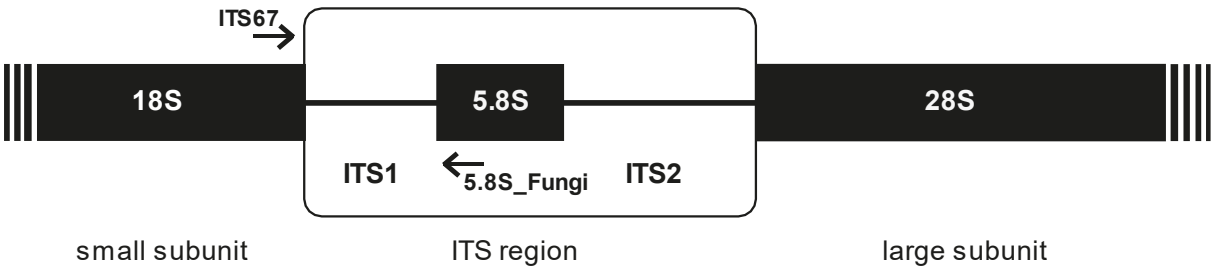


Figure 2: Map of the nuclear ribosomal genes and the intercalary ITS region. Respectively, the positions of the forward (ITS67) and reverse primers (5.8S_Fungi) which were established as part of this thesis are marked with arrows.

1.4.3 Shotgun sequencing for broader overview

Besides amplifying specific target regions and therefore assessing organismic groups, a new approach known as sedaDNA shotgun sequencing recently emerged. This method enables the reconstruction of complex ecosystems, taking a snapshot of the entire biome, as the whole DNA contained in a sample is being sequenced (Pedersen et al., 2016; Parducci et al., 2019). This became possible as genomic databases have been largely extended and improved and therefore a greater assignment of shotgun-sequenced reads has become possible. Moreover, general sequencing costs are declining. The method itself is not limited by sequence or molecule length, but bioinformatic processing of the sequenced data usually filters out sequences below a 30 basepair threshold (Pedersen et al., 2016). To assess the ancient origin of the reads, typical damage patterns of DNA which occur throughout degradation can be assessed, including DNA fragmentation or deamination (Ginolhac et al., 2011).

So far, the proxy has been applied on permafrost samples (Wang et al., 2021; Courtin et al., 2022; Kjær et al., 2022) or, rarely, lake sediment (Wang et al., 2021; Iwańska et al., 2022). Complex ecosystem dynamics can be assessed, so far focusing mainly on above-ground terrestrial ecosystems, including the dynamics between vegetation and mammals or birds. Studies targeting below-ground beneficial or parasitic associations amongst plants and soil microorganisms, including fungi and bacteria, are scarce (Courtin et al., 2022). However, understanding the temporal dynamics in the plant microbiome will further allow to unravel the complexity of plant-microbe interactions and help understand the impact of plants and further environmental factors on microbial community establishment. Vice-versa, tracing microbial community establishment helps understanding the functional roles of microorganisms in general plant health and ecosystem functioning.

1.5 Thesis objective

The main goal of this thesis was to contribute to the understanding of how root-associated microbial communities have changed over time in relation to vegetation and climate change.

1.) The first objective was to contribute to understanding of how soil communities changed over the past in relation to climate change. To do so, a primer pair suitable to specifically amplify short fungal ancient DNA was established. Then, the primer pair was applied to samples of multiple lakes to assess community dynamics across locations. For validation, the data were compared to fungal data derived from shotgun sequencing. Also, bacterial community changes in the shotgun data were assessed.
2.) The second objective was to trace fungus-plant covariation over time-scales covering multi millenia. A comparison of metabarcoding data with shotgun derived data has been done.

3.) The third objective of this thesis was to reconstruct weathering and soil development after deglaciation using lake sedimentary ancient DNA shotgun data. Aside, factors driving long-term pedogenesis have been assessed.

1.6 Thesis outline and author contributions

This thesis is designed as a cumulative dissertation, consisting of 7 chapters. The first chapter provides a general introduction to the research topic and presents the main objective and goal of the project. Chapters 2, 3, and 4 are comprised of publications which are either published or submitted to a peer-reviewed scientific journal. The respective status of the publication process is indicated in Table 1 below. In chapter 5, the overall outcome of the dissertation is being discussed. A list of references being used in the introduction and discussion part is provided in chapter 6. Chapter 7, the appendix, includes a fourth manuscript which is in preparation. The tables and figures in each chapter are numbered separately.

Table 1: Overview of the manuscripts included in this thesis

Manuscript I	**Evaluation of lake sedimentary ancient DNA metabarcoding to assess fungal biodiversity in Arctic paleoecosystems**
Authors	Peter A. Seeber, Barbara von Hippel, Håvard Kauserud, Ulrike Löber, Kathleen R. Stoof-Leichsenring, Ulrike Herzschuh, Laura S. Epp
Status	*Published* in *Environmental DNA*, 10.1002/edn3.315.
Summary	We re-evaluated primer combinations for the amplification of the ITS region for fungal metabarcoding of ancient sedimentary DNA using *in silico* and *in vitro* PCR. The analysis yielded the primer pair ITS67 and 5.8S most suitable for short fragments. PCR amplification was conducted on samples from a boreal and four arctic lakes, yielding a great biodiversity of terrestrial fungi.
Authors contributions	PAS: *in silico* and *in vitro* primer validation, data analyses, writing first draft of the manuscript; BvH: metabarcoding laboratory experiments, data analyses, input for interpretation of the metabarcoding data; HK, UL, KS-L, UH: input for the analysis and interpretation of the

	metabarcoding data, LSE: study design, data analyses, writing first draft of the manuscript
Manuscript II	**Long-term fungus-plant covariation from multi-site sedimentary ancient DNA metabarcoding**
Authors	Barbara von Hippel, Kathleen R. Stoof-Leichsenring, Luise Schulte, Peter Seeber, Laura S. Epp, Boris K. Biskaborn, Bernhard Diekmann, Martin Melles, Luidmila Pestryakova, Ulrike Herzschuh
Status	*Published* in *Quaternary Science Reviews*, 295: 107758.
Summary	We used sedimentary ancient DNA metabarcoding to analyse four arctic (Lama, CH12, Levinson Lessing, Kyutyunda) and one boreal lake (Bolshoe Toko) on their fungal and plant community composition changes during the last about 47 cal ka BP. We showed that fungal functional groups are shifting in relation to warming and their species richness is alternating. We provide evidence that fungal shifts coincide with vegetation changes, highlighting long-term dependencies of woody taxa towards mycorrhizal fungi.
Authors contributions	BvH: study design, metabarcoding laboratory experiments, data analysis, dating and age-depth modeling of Lama, writing first draft of the manuscript; KSL: study design, supervision laboratory part and bioinformatics; LS: sampling of the cores, supervision of DNA extractions; PS and LE: bioinformatic evaluation of the marker; BB: retrieval and dating including age-depth modelling of Bolshoe Toko, Kyutyunda and CH12; BD: project lead Kyutyunda and Bolshoe Toko; MM: retrieval of the sediment core of Lake Lama; UH: study design, supervision of writing of first manuscript version. All authors commented on the first and revised version of the manuscript.
Manuscript III	**Postglacial bioweathering, soil nutrient cycling, and podzolization from palaeometagenomics of plants, fungi, and bacteria**
Authors	Barbara von Hippel, Kathleen R. Stoof-Leichsenring, Martin Melles, Ulrike Herzschuh
Status	*Submitted* to *Nature Communications* (06-13-2023)
Summary	Here, we assessed shotgun sequencing data of lake Lama, northern-

	central Siberia to trace plant, fungal, and bacterial community changes and reconstructed soil development alongside. The data showed clear trends from lichen-dominated tundra communities with strong initial weathering in the Late Glacial to larch-spruce forests in the Holocene. On top, podzolization was reconstructed.
Authors contributions	BvH: study design, subsampling of Lama sediment, DNA extraction, built library, data analyses, writing first draft of the manuscript; KSL: supervision laboratory part and bioinformatics; MM: collection of the sediment record, XRF analyses; UH: study design, supervision of data analyses & supervision of writing of first manuscript version. All authors commented on the submitted version of the manuscript.
Manuscript IV	**Spatio-temporal microbial associations of plants in cold environments**
Authors	Barbara von Hippel, K.R. Stoof-Leichsenring, W. Shen, U. Herzschuh
Status	*in preparation*, Appendix
Summary	In this manuscript, we compared the soil communities of three shotgun datasets derived from lake sediment cores of two Siberian lakes (Lama and Bolshoe Toko) and one from the Tibetan Plateau (Ximencuo). The recovered soil communities were correlated to plant occurrences. We showed distinct rhizobial communities for each site as well as plant genus. Nonetheless, general overlaps between the sites and plant taxa were recovered.
Authors contributions	BvH: study design, subsampling of Lama sediment, DNA extraction, library built, data analysis, writing first draft; KSL: supervision laboratory part and bioinformatics; WS: extraction and library built Ximencuo; UH: study design, supervision of data analysis.

2 Manuscript I

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

Status

Published in *Environmental DNA*, 10.1002/edn3.315.

Authors

Peter Seeber[1], Barbara von Hippel[2], Håvard Kauserud[3], Ulrike Löber[4,5], Kathleen Rosmarie Stoof-Leichsenring[2], Ulrike Herzschuh[2,6,7], Laura Saskia Epp[1*]

Affiliations

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

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

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

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

*corresponding author (laura.epp@uni-konstanz.de)

Keywords

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

2.1 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, there-fore, 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.

2.2 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.

2.3 Materials and Methods

2.3.1 Primer design and evaluation

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.

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

Primer combination	Dikarya		Ascomycota		Basidiomycota		Fungi incertae sedis		Viridiplantae	Metazoa	Percent on-target	Percent off-target
	Count	*Length*	*Count*	*Length*	*Count*	*Length*	*Count*	*Length*	*Count*	*Count*		
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

Primer combination	Dikarya		Ascomycota		Basidiomycota		Fungi incertae sedis		Viridiplantae	Metazoa	Percent on-target	Percent off-target
	Count	*Length*	*Count*	*Length*	*Count*	*Length*	*Count*	*Length*	*Count*	*Count*		
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.

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.

2.3.2 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, 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.

2.3.3 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.

2.3.4 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).

2.3.5 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).

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

Clade	Division	Class	Order	Number of OTUs	Assigned reads (%)		
					97% b.i.	90% b.i.	85% b.i.
			Chaetothyriales	2	1.16	7.23	6.42
			Verrucariales	1	0.03	0.20	0.29
			Helotiales	11	38.80	20.94	19.73
		Leotiomycetes	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
			Agaricales	23	3.51	5.40	5.01
			Polyporales	2	0.98	0.41	0.33
			Hymenochaetales	1	0.90	0.32	0.26
		Agaricomycetes	Russulales	3	0.15	0.38	0.31
			Boletales	2	0.03	0.04	0.32
			Thelephorales	1	0.01	0.38	0.46
	Basidiomycota		Atheliales	1	0.01	0.13	0.11
			Filobasidiales	10	3.18	1.14	0.95
		Tremellomycetes	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
		Mortierellomycetes	Mortierellales	9	19.81	21.90	22.01
		Mucoromycetes	Mucorales	1	0.06	0.02	0.10
	Mucoromycota	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							

Clade	Division	Class	Order	Number of OTUs	Assigned reads (%)		
					97% b.i.	90% b.i.	85% b.i.
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.).

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 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).

2.4 Results

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.

2.4.1 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.

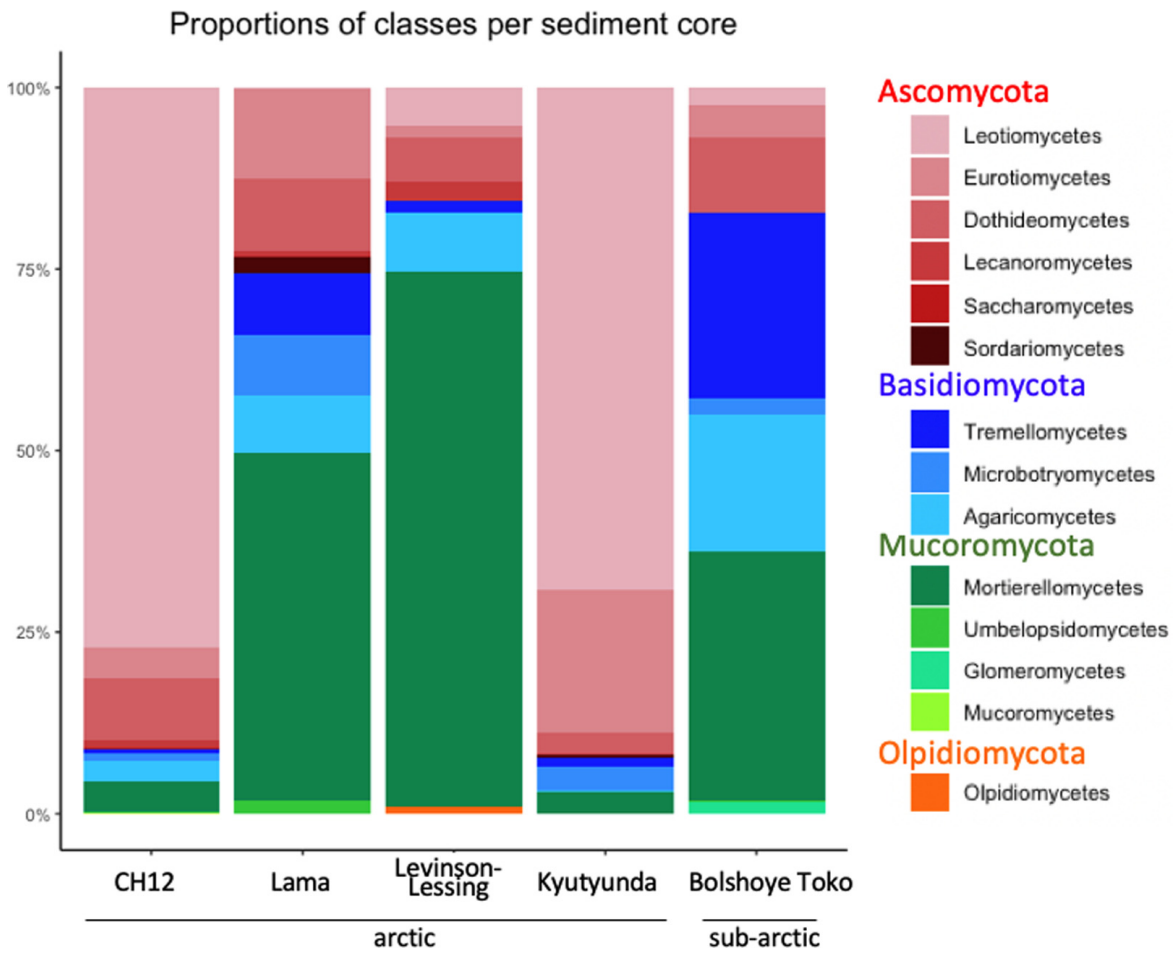


Figure 1: Proportions of fungal classes in lake sediments, averaged across all respective samples.

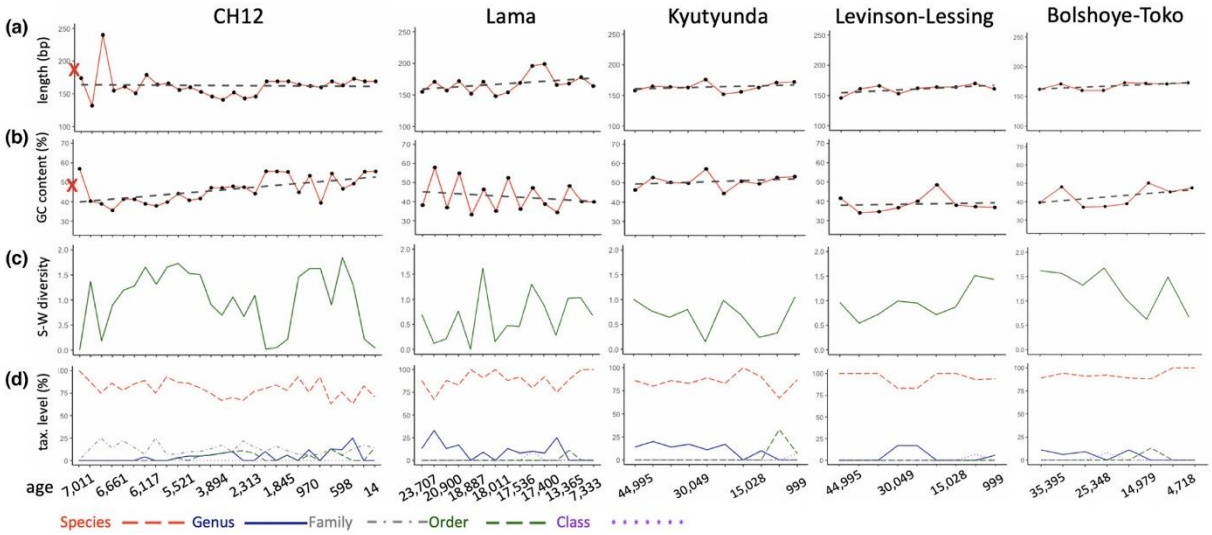


Figure 2: Effect of sample age on (a) mean amplicon lengths, (b) weighted mean GC content, (c) Shannon–Wiener (S-W) diversity indices of fungal taxa on order level, and (d) taxonomic (tax.) resolution, that is, the percentage of OTUs assigned on species (red), genus (blue), family (gray), order (green), 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).

2.4.2 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).

2.4.3 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 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%.

2.4.4 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).

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).

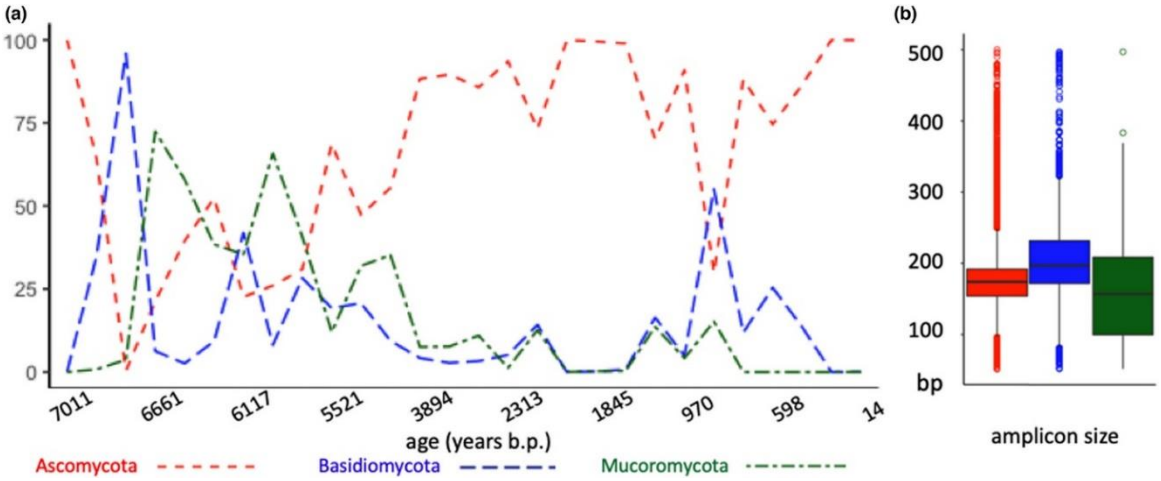


Figure 3: Changes in proportions of the three major divisions Ascomycota (red), Basidiomycota (blue), and Mucoromycota (green) over sample age of lake CH12 (a), and respective fragment lengths (b); boxes indicate 2nd and 3rd quartiles, center lines indicate median values, upper (and lower) whiskers extend 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.

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%.

2.5 Discussion

2.5.1 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 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.

2.5.2 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.

2.5.3 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.,

2.6 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.

2.7 Acknowledgements

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.

2.8 Conflict of interest

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

2.9 References

Abarenkov, K., Henrik Nilsson, R., Larsson, K.-H., *et al.*, 2010. The UNITE database for molecular identification of fungi - recent updates and future perspectives. *New Phytol* **186**: 281–285.

Adamo, M., Voyron, S., Chialva, M., *et al*., 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 Divers*.

Banchi, E., Stankovic, D., Fernández-Mendoza, F., *et al*., 2018. ITS2 metabarcoding analysis complements lichen mycobiome diversity data. *Mycol Prog* **17**: 1049–1066.

Bar-On, Y.M., Phillips, R., Milo, R., 2018. The biomass distribution on Earth. *Proc Natl Acad Sci* **115**: 6506–6511.

Bellemain, E., Carlsen, T., Brochmann, C., *et al*., 2010. ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. *BMC Microbiol* **10**: 189.

Bellemain, E., Davey, M.L., Kauserud, H., *et al*., 2013. Fungal palaeodiversity revealed using high-throughput metabarcoding of ancient DNA from arctic permafrost. *Environ Microbiol* **15**: 1176–1189.

Blaalid, R., Kumar, S., Nilsson, R.H., *et al*., 2013. ITS1 versus ITS2 as DNA metabarcodes for fungi. *Mol Ecol Resour* **13**: 218–224.

Blackwell, M., 2011. The Fungi: 1, 2, 3 … 5.1 million species? *Am J Bot* **98**: 426–438.

Booth, T., 2011. Taxonomic notes on coprophilous fungi of the Arctic: Churchill, Resolute Bay, and Devon Island. *Can J Bot* **60**: 1115–1125.

Botnen, S.S., Thoen, E., Eidesen, P.B., *et al*., 2020. Community composition of arctic root-associated fungi mirrors host plant phylogeny. *FEMS Microbiol Ecol* **96**:.

Boyer, F., Mercier, C., Bonin, A., *et al*., 2016. Obitools: a unix-inspired software package for DNA metabarcoding. *Mol Ecol Resour* **16**: 176–182.

Brundrett, M., 2004. Diversity and classification of mycorrhizal associations. *Biol Rev Camb Philos Soc* **79**: 473–495.

Capo, E., Giguet-covex, C., Rouillard, A., *et al*., 2020. Lake sedimentary DNA research on past terrestrial and aquatic biodiversity : Overview and recommendations.

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. *J Archaeol Sci* **102**: 61–70.

Clemmensen, K.E., Bahr, A., Ovaskainen, O., *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 Harb Perspect Biol* **5**: a012567–a012567.

Davis, O.K. , Shafer, D.S., 2006. Sporormiella fungal spores, a palynological means of detecting herbivore density. *Palaeogeogr Palaeoclimatol Palaeoecol* **237**: 40–50.

Deacon, J.W., Donaldson, S.J., Last, F.T., 1983. Sequences and interactions of mycorrhizal fungi on

birch. *Plant Soil* **71**: 257–262.

Deiner, K., Bik, H.M., Mächler, E., *et al*., 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Mol Ecol* **26**: 5872–5895.

Domaizon, I., Winegardner, A., Capo, E., *et al*., 2017. DNA-based methods in paleolimnology: new opportunities for investigating long-term dynamics of lacustrine biodiversity. *J Paleolimnol* **58**: 1–21.

Domsch, K.H., Gams, W., Anderson, T.H., 2007. Compendium of Soil Fungi, second edi. Eching: IHW-Verlag.

Epp, L.S., Boessenkool, S., Bellemain, E., *et al*., 2012. New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Mol Ecol* **21**: 1821–1833.

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

Ficetola, G., Coissac, E., Zundel, S., *et al*., 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. *J Exp Bot* **59**: 1115–1126.

Gardes, M., Bruns, T.D., 1993. ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. *Mol Ecol* **2**: 113–118.

Grau, O., Geml, J., Pérez-Haase, A., *et al*., 2017. Abrupt changes in the composition and function of fungal communities along an environmental gradient in the high Arctic. *Mol Ecol* **26**: 4798–4810.

Haile, J., Froese, D.G., MacPhee, R.D.E.E., *et al*., 2009. Ancient DNA reveals late survival of mammoth and horse in interior Alaska. *Proc Natl Acad Sci* **106**: 22352–22357.

Hawksworth, D.L., Lücking, R., 2017. Fungal Diversity Revisited: 2.2 to 3.8 Million Species. *Microbiol Spectr* **5**:.

Heeger, F., Bourne, E.C., Baschien, C., *et al*., 2018. Long-read DNA metabarcoding of ribosomal RNA in the analysis of fungi from aquatic environments. *Mol Ecol Resour* **18**: 1500–1514.

van der Heijden, M.G.A., Klironomos, J.N., Ursic, M., *et al*., 1998. Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396**: 69–72.

von Hippel, B., Stoof-Leichsenring, K.R., Schulte, L., *et al*., 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

James, T.Y., Stajich, J.E., Hittinger, C.T., Rokas, A., 2020. Toward a Fully Resolved Fungal Tree of Life. *Annu Rev Microbiol* **74**: 291–313.

Kanz, C., Aldebert, P., Althorpe, N., *et al*., 2005. The EMBL nucleotide sequence database. *Nucleic Acids Res* **33**: D29.

Khomich, M., Davey, M.L., Kauserud, H., *et al*., 2017. Fungal communities in Scandinavian lakes along a longitudinal gradient. *Fungal Ecol* **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. *Quat Sci Rev* **147**: 422–434.

Kumpula, J., 2001. Winter grazing of reindeer in woodland lichen pasture. *Small Rumin Res* **39**: 121–130.

Lammers, Y., Heintzman, P.D., Alsos, I.G., 2021. Environmental palaeogenomic reconstruction of an Ice Age algal population. *Commun Biol* **4**: 220.

Lücking, R., Aime, M.C., Robbertse, B., *et al*., 2020. Unambiguous identification of fungi: Where do we stand and how accurate and precise is fungal DNA barcoding? *IMA Fungus* **11**:.

Lydolph, M.C., Jacobsen, J., Arctander, P., *et al*., 2005. Beringian paleoecology inferred from permafrost-preserved fungal DNA. *Appl Environ Microbiol* **71**: 1012–1017.

Martin, F., 2014. The Ecological Genomics of Fungi - Wiley Online Library.

Mercier, C., Boyer, F., Bonin, A., Coissac, E., 2013. SUMATRA and SUMACLUST: fast and exact comparison and clustering of sequences. In *SeqBio 2013 workshop*. pp. 27–29.

Naranjo-Ortiz, M.A., Gabaldón, T., 2019. Fungal evolution: diversity, taxonomy and phylogeny of the Fungi. *Biol Rev* **94**: 2101–2137.

Nilsson, R.H., Anslan, S., Bahram, M., *et al*., 2019. Mycobiome diversity: high-throughput sequencing and identification of fungi Nature reviews | Microbiology. *Nat Rev Microbiol*. **17**.

Oksanen, A.J., Blanchet, F.G., Friendly, M., *et al*., 2020. Community Ecology Package vegan,. 2.5-7.

Pääbo, S., 1989. Ancient DNA: Extraction, characterization, molecular cloning, and enzymatic amplification. *Proc Natl Acad Sci U S A* **86**: 1939–1943.

Parducci, L., Bennett, K.D., Ficetola, G.F., *et al*., 2017. Ancient plant DNA in lake sediments. *New Phytol* **214**: 924–942.

Pedersen, M.W., Ruter, A., Schweger, C., *et al*., 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. *New Phytol* **220**: 1059–1075.

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

Raper, D., Bush, M., 2009. A test of Sporormiella representation as a predictor of megaherbivore presence and abundance. *Quat Res* **71**: 490–496.

Riaz, T., Shehzad, W., Viari, A., *et al*., 2011. ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Res* **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. *Glob Ecol Conserv* **17**: e00547.

Schoch, C.L., Seifert, K.A., Huhndorf, S., et al., 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc Natl Acad Sci U S A* **109**: 6241–6.

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

Spatafora, J.W., Chang, Y., Benny, G.L., *et al*., 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., *et al*., 2015. One fungus, which genes? Development and assessment of universal primers for potential secondary fungal DNA barcodes. *Persoonia - Mol Phylogeny Evol Fungi* **35**: 242–263.

Talas, L., Stivrins, N., Veski, S., *et al*., 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. *Rev Palaeobot Palynol* **90**: 249–262.

Tedersoo, L., Anslan, S., Bahram, M., *et al*., 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. *Environ Microbiol Rep* **8**: 774–779.

Varma, A., Prasad, R., Tuteja, N. eds., 2017. Mycorrhiza - Function, Diversity, State of the Art. Cham: *Springer* International Publishing.

Vilgalys, R., Gonzalez, D., 1990. Organization of ribosomal DNA in the basidiomycete Thanatephorus praticola. *Curr Genet* **18**: 277–280.

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

Wickham, H., 2011. The Split-Apply-Combine Strategy for Data Analysis. *J Stat Softw* **40**:.

Willerslev, E., Davison, J., Moora, M., *et al*., 2014. Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature* **506**: 47–51.

Willerslev, E., Hansen, A.J., Binladen, J., *et al*., 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. *Quat Sci Rev* **77**: 1–3.

3 Manuscript II

Long-term fungus-plant covariation inferred from multi-site sedimentary ancient DNA metabarcoding

Status

Published in *Quaternary Science Reviews*, **295**, 107758 (2022).

Authors

Barbara von Hippel[1*], Kathleen R. Stoof-Leichsenring[1], Luise Schulte[1], Peter Seeber[2], Laura S. Epp[2], Boris K. Biskaborn[1], Bernhard Diekmann[1], Martin Melles[3], Luidmila Pestryakova[4], Ulrike Herzschuh[1,5,6*]

Affiliations

1 Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Polar Terrestrial Environmental Systems, Telegrafenberg A45, Potsdam, Germany

2 Limnological Institute, Department of Biology, University of Konstanz, Germany

3 Institute of Geology and Mineralogy, University of Cologne, Germany

4 Institute of Natural Sciences, North-Eastern Federal University of Yakutsk, Yakutsk, 677007, Russia

5 Institute of Environmental Science and Geography, University of Potsdam, Germany

6 Institute of Biochemistry and Biology, University of Potsdam, Germany

*corresponding authors (barbara.von.hippel@awi.de, ulrike.herzschuh@awi.de)

3.1 Abstract

Climate change has a major impact on arctic and boreal terrestrial ecosystems as warming leads to northward treeline shifts, inducing consequences for heterotrophic organisms associated with the plant taxa. To unravel ecological dependencies, we address how long-term climatic changes have shaped the co-occurrence of plants and fungi across selected sites in Siberia.

We investigated sedimentary ancient DNA from five lakes spanning the last 47,000 years, using the ITS1 marker for fungi and the chloroplast P6 loop marker for vegetation metabarcoding. We obtained 706 unique fungal operational taxonomic units (OTUs) and 243 taxa for the plants. We show higher OTU numbers in dry forest tundra as well as boreal forests compared to wet southern tundra. The most abundant fungal taxa in our dataset are Pseudeurotiaceae, *Mortierella*, Sordariomyceta, *Exophiala*, *Oidiodendron*, *Protoventuria*, *Candida vartiovaarae*, *Pseudeurotium*, *Gryganskiella fimbricystis*, and *Trichosporiella cerebriformis*. The overall fungal composition is explained by the plant

composition as revealed by redundancy analysis. The fungal functional groups show antagonistic relationships in their climate susceptibility. The advance of woody taxa in response to past warming led to an increase in the abundance of mycorrhizae, lichens, and parasites, while yeast and saprotroph distribution declined. We also show co-occurrences between Salicaceae, *Larix*, and *Alnus* and their associated pathogens and detect higher mycorrhizal fungus diversity with the presence of Pinaceae. Under future warming, we can expect feedbacks between fungus composition and plant diversity changes which will affect forest advance, species diversity, and ecosystem stability in arctic regions.

3.2 Introduction

Permafrost soils comprise almost half of the global soil organic carbon (Tarnocai et al., 2009). The permanently frozen ground enables only the establishment of vegetation with shallow roots (Blume-Werry et al., 2019), most usually, tundra plants. However, global warming is leading to increased shrub-growth on permafrost in arctic regions, particularly at the boundary between the High and Low Arctic (Myers-Smith et al., 2015). Changes in the vegetation cover will subsequently influence the carbon pool as tundra plants growing on permafrost regulate the uptake and release of carbon dioxide and methane (McGuire et al., 2009).

Boreal forests are the world's largest terrestrial biome, covering an area of around 9% of the total land mass between 45° and 70° north (Czimczik et al., 2005). The climate in their distribution area is defined by severe winters, warm summers, relatively little precipitation, and an overall short growing season (MacDonald et al., 2008), leading to generally low species diversity in the forests. Boreal ecosystems are highly impacted by global warming as increasing temperatures result in reduced snow cover (Kreyling et al., 2012), loss of permafrost (Fedorov et al., 2017), longer growing seasons (Jarvis and Linder, 2000) or severe droughts and wildfire (Flannigan et al., 2009). As a consequence, the distribution and abundance of boreal plants is altered (Zhang et al., 2011) and local wildlife is endangered (Bradshaw et al., 2009). Furthermore, in Siberia, the migration of evergreen coniferous tree taxa into *Larix* dominated areas after warming has been observed, possibly leading to decreased albedo and further temperature increases (Kharuk et al., 2007, 2009). The composition of underground microorganisms in boreal forests, for example fungi and bacteria, is also impacted as it is shaped by the tree species present (Urbanová et al., 2015), suggesting microbial communities will shift with changing vegetation.

Fungi represent a principal component of soils and sustain a broad variety of ecosystem functions (Frąc et al., 2018). Fungal functional types reflect the strong interactions between plants and fungi but, so

far, detailed processes involved in particular temporal changes in their covariation are not understood (Zobel et al., 2018). Grouping fungi according to their roles in the terrestrial ecosystem allows the assessment of compositional shifts. This, however, remains challenging as the ecological functions of many taxa are still not understood. Major ecological functional fungus groups in forest ecosystems are saprotrophs, mycorrhizae, and parasites. Saprotrophic species are decomposers in terrestrial ecosystems (Baldrian and Valášková, 2008). A plant's benefit from mycorrhizal fungi is the enabled acquisition of mineral nutrients in solution (e.g. phosphate), while the fungus in return receives carbohydrates from the plants (Finlay, 2008), making them indispensable for the plants' establishment and survival. Mycorrhizal fungi-plant associations include arbuscular mycorrhizae, ectomycorrhizae, ericoid mycorrhizae, and orchid mycorrhizae (Brundrett and Tedersoo, 2018). Parasitic fungi, such as *Heterobasidion,* infecting conifer tree taxa (Garbelotto and Gonthier, 2013), are important for eliminating weak trees to maintain the functioning of healthy forest ecosystems. While biotrophic plant parasites feed from living tissue, necrotrophic fungi penetrate the plant, destroy the tissue, and subsequently provoke plant death (Naranjo-Ortiz and Gabaldón, 2019).

Experimental warming led to an increase in ectomycorrhizal fungi and free-living filamentous fungi, while a decrease in yeast was observed (Treseder et al., 2016). It also showed an increase in the evenness of fungal tundra communities, including ectomycorrhizal fungi, in relation to rising temperature (Deslippe et al., 2012). Furthermore, under warming, saprotrophs shifted their metabolism from wood-decaying to self-maintenance and subsequent spore-production (Romero-Olivares et al., 2015, 2019). Warming also leads to an increase in specific plant pathogens (Otrosina and Cobb, 1989). Next generation sequencing studies of soil fungi in arctic Alaska revealed that warming does not affect overall fungus richness but leads to changes in the community composition. The results of these studies revealed that decreases in ectomycorrhizae, ericoid mycorrhizae, and lichens in the tundra accompany increases of saprotrophic, pathogenic, and root endophytic fungal richness (Geml et al., 2015; Mundra et al., 2016). Almost all information on fungal compositional turnover originates from short-term warming experiments (from one season up to a few years; e.g. Heinemeyer et al., 2003; Geml et al., 2015). Accordingly, time-series of compositional changes in fungal communities during past climate changes are highly desirable to assess potential shifts of ecosystem functioning in a rapidly warming world. Time-series are also an asset when testing whether lab experiments reflect real natural developments on long geological timescales.

Compositional changes of vegetation and associated fungal communities are slow. They occur on decadal, centennial, or even millennial time-scales not being covered by short observational time-series and thus the exploitation of palaeoecological archives is required. As large parts of Siberia were not covered by glaciers during the Last Glacial Maximum (LGM; Svendsen et al., 2004), lakes from this

region provide sedimentary archives which continuously cover the rather warm marine isotopic stage (MIS) 3 (50–30 thousand years (ka)), the cold MIS 2 (30–15.5 ka), and warm Holocene (MIS 1) (the last 11.6 ka) (Kreveld et al., 2000; Swann et al., 2005), thereby encompassing tremendous vegetation changes. Lake sediments represent natural archives of terrestrial environmental change (e.g. Epp et al., 2015; Alsos et al., 2018; Courtin et al., 2021). As northern Russia is warming faster than the global average (Biskaborn et al., 2019b), lake sediments can provide valuable information on associated terrestrial ecosystem changes. While many sedimentary pollen records focus on vegetation change, there is only limited information about fungi (e.g. from non-pollen palynomorphs; Van Geel, 2001) as their fossil remains are limited (Loughlin et al., 2018; Quamar and Stivrinsz, 2021).

Sedimentary ancient DNA metabarcoding (sedaDNA) is a promising palaeoecological proxy using specific genetic marker regions to study past biodiversity (Sønstebø et al., 2010). So far, many sedaDNA studies have investigated plant metabarcoding, mostly applying the trnL P6 loop marker (Parducci et al., 2017; Alsos et al., 2018; Liu et al., 2020), but only a few studies focus on fungal aDNA from sedimentary deposits (Lydolph et al., 2005; Bellemain et al., 2013; Talas et al., 2021), working on Siberian permafrost sediments (Bellemain et al., 2013) and lake sediments (Talas et al., 2021). For fungal metabarcoding, the internal transcribed spacer (ITS) region is the most commonly used DNA barcoding region (Seifert, 2009), but the primers used in early studies caused quite some amplification biases (Bellemain et al., 2010). Subsequently, primers specifically targeting ancient and degraded DNA were designed and used on permafrost deposits (Epp et al., 2012; Bellemain et al., 2013). These have been refined according to the current status of reference databases by Seeber et al. (2022), providing a primer pair that is highly suitable to amplify sedaDNA, targeting short amplicons of a mean length of 183 bp and highly specific towards fungi. This primer combination enables studies using sedaDNA to trace fungus-plant interactions over time and their adaptation mechanisms towards warming even up to species level. The advantage of targeting lake sediment samples as ancient DNA records over permafrost is that during the sedimentation processes, environmental DNA is continuously deposited in the lake, allowing a chronological community reconstruction. A recent investigation of lake sediments showed clear variations in fungal community compositions: while saprotrophs remained stable over time, host-specific fungi such as plankton parasites and mycorrhizae shifted in relation to human impact and changing climate (Talas et al., 2021).

This study analyses lake sediments from five sites in Siberia spanning MIS 3, 2, and 1 for their fungal composition using sedaDNA metabarcoding with the ITS1 marker as well as their vegetation composition applying the trnL P6 loop marker. We address the following questions: (1) How does fungal and plant alpha diversity change with varying climatic and environmental conditions? (2) How do fungal taxonomic composition and function change relative to vegetation transition? Based on the

answers, we draw conclusions on fungus-plant covariation under climatic changes over long timescales and its impact on biodiversity alteration in Arctic terrestrial environments. As exemplified by the response of fungus-plant interactions, our results help to predict more accurately the impact of future warming on biodiversity shifts.

3.3 Geographic setting and study sites

All study sites are located within Siberia, central eastern Russia (Fig. 1), and are characterised by permafrost soils (Brown et al., 1997; Tchebakova et al., 2009). The climate in the area is rather continental with hot summers and long, severe winters (Atlas Arktiki, 1985). The most prevalent vegetation is boreal forest with spruce, pine, fir, and larch in the western and southern parts and pure larch forest in the east. Arctic regions along the coast and on the Taymyr Peninsula are covered by tundra. The main characteristics of the sampling locations are displayed in Table 1. The climate data are taken from the Russian Institute of Hydrometeorological Information: World Data Center (2021) unless indicated differently.

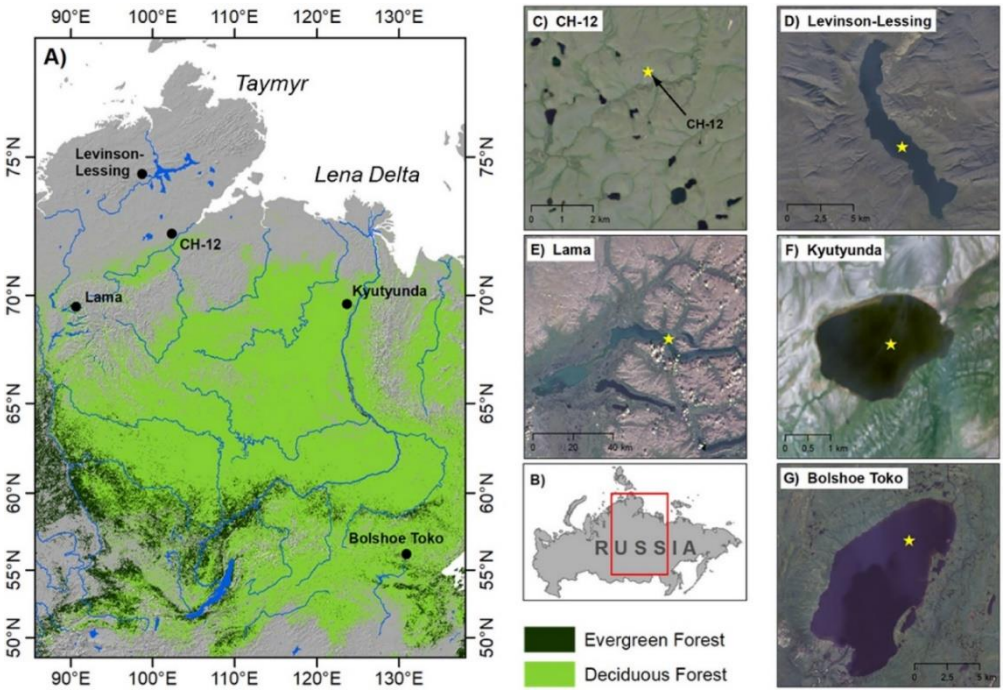


Figure 1: Map of central Russia showing the location of the study sites (A+B). (C) to (G) satellite images of the lakes with their surroundings. The locations of the cores are marked with a star. The "distribution of deciduous and evergreen forests" is taken from the ESA CCI Land Cover time-series v2.0.7 (1992–2015)- data set (https://www.esa-landcover-cci.org/). For the illustration, the land cover classes "70" ("Tree cover, needleleaved, evergreen") and "80" ("Tree cover, needleleaved, deciduous") were extracted.

Table 1. Main characteristics of the sampled lakes.

Lake	Coordinates	Type of vegetation	Mean temperature	Dimension	Coring
Levinson Lessing	74.27°N, 98.39°E; 48 m a.s.l. (Taymyr peninsula)	sparse lichen-herb, moss-forb, dry sedge-forb tundra with dominant *Dryas octopetala*, *Salix polaris,* and *Cassiope tetragona* (Anisimov and Pospelov, 1999)	July: 12.5 °C January: -31.5 °C Hatanga weather station; 71.98 °N, 102.47 °E; distance to the lake: 289 km	15 km × 2.5 km, maximum depth: 120 m (Lebas et al., 2019)	year: 2017 length: 46 m (Co1401; Fig. 1D) at a depth of 112 m age: 62 cal ka BP (Scheidt et al., 2021)
CH12	72.4°N, 102.29°E; 60 m a.s.l. (southern Taymyr Peninsula north of the Putorana Plateau)	shrub tundra dominated by *Sphagnum*, *Hylocomium, Aulacomnium, Dicranum,* and *Polytrichum* as well as *Empetrum nigrum, Betula nana,* and *Vaccinium uliginosum;* stands of *Larix gmelinii* (Klemm et al., 2016; Niemeyer et al., 2017)	July: 12.5 °C January: -31.5 °C Hatanga weather station; 71.98 °N, 102.47 °E; distance to the lake: 60 km	elliptically shaped, mean radius: 100 m, maximum depth: 14.3 m	year: 2011 length: 1.21 m (Fig. 1C) age: 7.1 cal ka BP (Stoof-Leichsenring et al., 2015; Klemm et al., 2016)
Kyutyunda	69.38 °N, 123.38°E; 66 m a.s.l. (northern Siberia on the central Siberian Plateau)	tundra-taiga transition zone, formed of a mosaic of *Larix* forest and shrub tundra with Poaceae, *Dryas,* and *Saxifraga* species	July: 13 °C January: -35.3 °C Kjusjur weather station; 70.68 °N, 127.4 °E; distance to the lake: 210 km	roughly circular at 2.2 km by 3 km, maximum depth: 3.5 m	year: 2010 length: 7 m (PG 2023; Fig. 1F) age: 38.8 cal ka BP (Supplement 1 and 2)
Lama	69.32°N, 90.12°E; 53 m a.s.l. (Putorana Plateau)	dense taiga with *Picea*, *Larix*, and *Betula*, shrubs such as *Alnus fruticosa*, *Salix,* and *Juniperus communis*, and dwarf shrubs (Andreev et al., 2004)	July: 13.8 °C January: -28.8 °C Volochanka weather station; 70.97 °N, 94.5 °E; distance to the lake: 247 km	area: 318 km^2; 80 km × 7 km; maximum depth: 254 m	year: 1997 length: 18.85 m (depth 66 m; PG1341, Fig. 1E) age: 23 kal ca BP (Supplement 3 and 4)
Bolshoe Toko	56.15° N, 130.30 °E; 903 m a.s.l. (northern slope of eastern Stanovoy Mountain Range in southern central Yakutia)	deciduous boreal forests formed by *Larix cajanderi* and *L. gmelinii* with occurrences of *Picea obovata*, *P. jezoensis,* and *Pinus sylvestris* (Konstantinov, 2000)	July: 34 °C January: -65 °C Toko weather station; 56.1 °N, 131.01 °E; distance to the lake: 44 km (Konstantinov, 2000)	area: 15.4 km × 7.5 km, maximum depth: 72.5 m	year: 2013 length: 3.8 m (at 26 m depth; PG2133; Fig. 1G) age: 33.8 cal ka BP (Courtin et al., 2021)

3.4 Materials and Methods

3.4.1 Sampling

After coring, all sediment was stored at 4 °C to remain cool until sampling and further processing to preserve the cores under conditions similar to those on the lakebed. A parallel study (Seeber et al., 2022) on the fungal metabarcoding marker addressed whether the time of the coring and the subsequent long-term storage conditions influenced the metabarcoding results, for example by promoting mould. The study demonstrated that there is no direct impact on the results. SedaDNA samples were taken from 1 m sub-core segments that were cut in half lengthwise. Subsampling was undertaken in the climate chamber of the Helmholtz Centre Potsdam – German Research Centre for Geosciences (GFZ) at 10 °C. The chamber is located in the cellar where no molecular genetic studies are conducted. Before subsampling, all surfaces were cleaned with DNA Exitus Plus™ (VWR, Germany) and demineralised water. All tools were cleaned according to the recommendations of Champlot et al. (2010) to avoid contamination with modern DNA and between the samples themselves. All materials were taken from the palaeogenetic DNA laboratory at the Alfred-Wegener-Institute (AWI) in Potsdam where they had been treated to remove DNA.

During sampling, protective clothing as well as face masks were worn. The surfaces of the core halves were scraped off twice with sterile scalpel blades and samples were taken using four knives and then placed in sterile 8 mL Sarstedt tubes. All samples were taken under the same conditions. The core from Levinson Lessing was similarly sampled in the laboratories of the Institute of Geology and Mineralogy at the University of Cologne. After sampling, the aDNA samples were frozen at −20 °C until DNA extraction and amplification.

Samples for DNA analyses were taken according to their estimated ages, at intervals of about 5 cal kyr (calibrated kiloyears), leading to 15 samples from Lama, 9 samples from Levinson Lessing, 10 samples from Kyutyunda, and 8 samples from Bolshoe Toko. For CH12, 28 samples were taken, at a higher temporal resolution of intervals of about every 100–250 years.

3.4.2 DNA extraction and amplification

SedaDNA was extracted using the DNeasy PowerMax Soil DNA Isolation Kit (Qiagen, Germany) according to the manufacturer's instructions. Before adding 3–7 g of wet sediment material for each sample, the PowerBead solution was mixed with buffer C1 and additionally Proteinase K (2 mg mL^{-1})

and DTT (5 M) to break up remaining small pieces of tissue and yield higher DNA concentration. The Proteinase K was added to the bead beating tubes before vortexing to reduce the risk of cross-contamination. We placed the tubes on a vortexer for 10 min and included an additional incubation step at 56 °C in a rotation oven overnight. All further steps were conducted according to the manufacturer's instructions. The final elution was conducted using 2 mL of solution C6. Each extraction batch was processed on a different day to avoid contamination between batches. 0.5 mL of the CH12 extracts were purified and concentrated to 50 µL with a GeneJET PCR purification Kit (Thermo Fisher Scientific, Germany). For all other lakes, 1 mL DNA extract was used for the purification. Afterwards, the concentration was measured with a Qubit Fluorometer (Qubit dsDNA BR assay kit, Qubit 4.0 Fluorometer, Thermo Fisher Scientific, USA) and the DNA diluted to 3 ng μL^{-1} which balanced out the differently processed volumes. Small aliquots were prepared to avoid freeze-thaw cycles. DNA extraction blanks were not concentrated, but used directly for subsequent PCR analyses.

For the amplification of fungal DNA, the tagged forward primer ITS67 and reverse primer 5.8 S were used (Seeber et al., 2022). The amplified region has a size of approximately 183 bp (without the primers). The use of tagged primers is essential to enable the assignment of the DNA sequences to original samples after next generation sequencing. For each batch, six replicates were conducted independently from each other.

For the reconstruction of the palaeovegetation, we used the chloroplast trnL P6 loop marker region with the tagged primer trnL g as the forward and trnL h as the reverse primer (Taberlet et al., 2007). For each batch, three replicates were conducted independently from each other.

A PCR reaction contained in total 25 µL consisting of 3 µL DNA at a concentration of 3 ng μL^{-1}, 0.2 µM of each primer, 10× HiFi buffer, 2 mM MgSO4, 0.1 mM dNTPs, 0.8 mg mL^{-1} BSA, and 1.25 U Platinum Taq High Fidelity DNA Polymerase (Invitrogen, United States), which can replicate through Uracil and is thus suitable for PCRs on damaged DNA (Rasmussen et al., 2010). Each PCR batch also contained 3 µL of the corresponding DNA extraction blank and a PCR negative control with 3 µL of DEPC-water instead of the DNA sample. All steps were conducted in the palaeogenetic laboratories at AWI Potsdam.

The PCR reaction itself was conducted in the Post-PCR laboratories at AWI Potsdam, which are located in a separate building to avoid contamination of ancient DNA samples with amplified DNA. Initially, lake CH12 was conducted as a separate project itself for the establishment of the metabarcoding primers (Seeber et al., 2022). To support the study of the marker establishment as well as this study, the results were merged after sequencing which led to small differences in the PCR protocol as

described as follows. The fungal ITS marker amplification for the CH12 samples was conducted in a thermocycler (Biometra, Germany) following the protocol for voucher samples (Seeber et al., 2022) while the other fungal samples were amplified using the following protocol: initial denaturation at 94 °C for 2 min, 40 cycles of 30 s denaturation at 94 °C, 30 s annealing at 54 °C, and 30 s elongation at 72 °C, final elongation of 10 min at 72 °C. The thermocycler protocol for plant trnL P6 loop amplification followed the protocol of Epp et al. (2018).

All PCR products were checked by gel electrophoresis (2% agarose gels). Only products showing expected gene bands were used for purification and subsequent sequencing. Purification was done with the MinElute PCR Purification Kit (Qiagen, Germany) with the elution in 50 µL of the elution buffer. Each PCR product was used entirely for purification and treated independently. DNA concentration was measured with a Qubit 4.0 Fluorometer, measuring dsDNA using the broad-sensitivity kit. For sequencing, 40 ng of each purified PCR product were pooled. If the concentration was not measurable, the total purified volume was added to the pool. All PCR replicates were used for the final pools (6 for the ITS1 metabarcoding, 3 for the chloroplast P6 loop metabarcoding). For extraction blanks and PCR no-template controls, 5 µL of each PCR product was added to avoid diluting the concentration of the final sequencing pool too much. The final pool was purified again with MinElute and adjusted to a final concentration of 33 ng μL^{-1} in 30 µL. Three fungal sequencing pools with 175–187 samples each (Pool 1: 187 samples (3 replicates of each lake besides CH12 and 18 samples of a different project), Pool 2: 187 samples (all CH-12 samples), Pool 3: 175 samples (3 more replicates for the other lakes)) were sent to Fasteris SA sequencing service (Switzerland). The service included library preparation using a specified protocol (Metafast library; a PCR-free library preparation method), quality control and sequencing on an Illumina MiSeq platform (2 × 250 bp, V3 chemistry with an expected output of 20 million paired-end reads).

The plant PCR products were treated equally to the fungal PCR products. We sequenced two pools for the plant metabarcoding. These pools were sequenced on an Illumina NextSeq500 device (2 × 150 bp, 120 million paired-end reads). In addition, plant trnL P6 loop data from the lake CH12 were used from Epp et al. (2018).

3.4.3 Bioinformatic analysis

For the quality-check, filtering and taxonomic assignment of the sequencing results, we used the open source OBITools pipeline (Boyer et al., 2016). First, *illuminapairedend* was conducted to pair sequence

ends, followed by *obigrep* which filters out all unpaired reads. Afterwards, *ngsfilter* was used to demultiplex the file into unique samples and *obiuniq* was applied to dereplicate sequence reads. All sequences shorter than 10 bp and with fewer than 10 reads were deleted applying *obigrep*. A detailed description of all filtering steps is attached (Supplement 5).

After filtering, *ecotag* was applied to the vegetation dataset to perform taxonomic classification of the sequences against the sequence database. For the taxonomic assignments in the metabarcoding community, different approaches can be used. One approach is to work with each assigned sequence variant (ASV) present in the sample (after filtering out sequencing errors) and compare them to a reference database. In the case of the vegetation dataset, we are working on the level of ASVs. For taxonomic classification of the vegetation dataset, we used the ArctBorBryo database based on quality-checked and curated Arctic and Boreal vascular plant and bryophyte reference libraries (Sønstebø et al., 2010; Willerslev et al., 2014; Soininen et al., 2015). Only those ASVs that have a 100% match to the database were kept in the plant dataset. The taxonomic names (either family, genus or species level) of the plant ASVs were checked on https://www.gbif.org/ for their occurrence in the study area. To simplify the dataset, all reads assigned to the ASVs with the same scientific name are merged into one taxon.

A different approach to analyse metabarcoding data is to work with operational taxonomic units (OTUs). When working on the OTU level, the sequence types are clustered together according to a specific threshold of sequence identity. For fungal metabarcoding, working on ASV level instead of OTUs might lead to an overestimation of the richness of common fungal species due to their haplotype variation (Estensmo et al., 2021; Tedersoo et al., 2022), but might also result in an underestimation of the richness of rare species. As the common fungi drive the main composition of the datasets, we therefore chose to work on an OTU level for this study. Additionally, clustering fungal OTUs is, in general, commonly used by the community, following the guideline of Tedersoo et al. (2022), and makes a comparison to other studies easier. For the fungal dataset, the open source *sumaclust* algorithm (Mercier et al., 2013) was applied to cluster sequences with an identity threshold of 0.97, generating operational taxonomic units (OTUs) before applying *ecotag*.

An analysis comparing ASVs and OTUs for the fungal data showed only small differences in the results, which did not change the overall pattern. Only absolute numbers of assigned sequences differed (5684 ASVs and 5411 OTUs). Nonetheless, we tested the distribution of the fungal taxa based on OTUs as well as on ASVs in a PCA. A fairly similar distribution of the samples is found (Supplements 6 and 7). The richness analysis revealed very similar trends when using ASVs or OTUs. We compared the ordinations by applying the functions procrustes () and protest () in the R package vegan (Oksanen

et al., 2020). The Procrustes comparison of the first two PCA axis scores of the datasets with OTUs and ASVs yielded a sum of squares of 1.44 while the Protest comparison with 999 permutations showed the sum of squares to be 0.7213 with a significance level of 0.001.

After filtering, *ecotag* was applied to perform taxonomic classification of the OTUs against the embl142 (based on the EMBL nucleotide sequence database, release 142; Kanz et al., 2005) and the UNITE database release for fungal metabarcoding (Nilsson et al., 2019). The UNITE database is a curated fungus database where detection of false positive reads might be lower than in the broader EMBL release. Using only the UNITE database for the assignment, however, might preclude identification of certain taxa. Therefore, the final assignment for each fungal OTU is based on the assignment from the database with the higher identity. When both databases match the same identity but differ in their specific species assignment, the UNITE database is used for final taxonomic classification.

All databases were built to be applicable for the *ecotag* algorithm as following: the sequences of the databases and NCBI taxonomy files were formatted in the ecoPCR format and ecoPCR was run to simulate *in silico* amplification of database sequences with the subsequent primers (allowing 5 mismatches in each primer sequence). The putatively amplified sequences were used as the reference databases and taxonomy information was added.

Fungal OTUs with identity levels equal to or higher than 98% were used for further analyses to keep only well-annotated sequences. All contaminants (non-fungal reads; OTUs occurring only in no-template controls/extraction blanks) and aquatic fungi as well as OTUs with total read counts lower than 10 have been excluded from further analysis. Seeber et al. (2022) describe these reads in more detail and show that they make up only a small part of the dataset, validating the reliability of the primer pair. For the vegetation data, the identity cut-off was at 100%. The taxa were checked on https://www.gbif.org/ for their occurrence in the study area. Taxa which do not occur in the area were filtered out from the dataset. Further excluded ASVs are algae which are also amplified by the marker, but are not part of the terrestrial vegetation being assessed in this study. We resampled both datasets to normalise the count data following the script of Kruse (2020; https://github.com/StefanKruse/R_Rarefaction) while using the lowest overall count of a sample as the base count. The vegetation data were resampled to a base count of 12,489, resulting in an exclusion of the samples from 9.9 cal ka BP (calibrated kiloanni before 1950 CE) of Kyutyunda and 7 cal ka BP from CH12 due to too low counts. The fungus data were resampled to a base count of 5,284, resulting in the exclusion of the sample from 5 cal ka BP from Kyutyunda and the sample from 18.8 cal ka BP from Lama.

3.4.4 Assessment of negative controls and contamination

For the plant dataset, we ran in total 29 extraction blanks (EBs) and 29 no-template controls (NTCs) along with the 189 samples. 100% of EBs and 90% of NTCs are clean and show no or a negligible proportion of contamination (lower than 0.01% of total reads). In 10% of NTCs we detected between 0.014 and 0.035% of the total reads (Supplement 8). We also checked the blanks for their contained ASVs and the percentage of the reads in the blanks vs. the samples. We identified 12 different ASVs which are present at more than 10% of their abundance (samples + blanks) in the blanks (Supplement 9).

We ran a total of 58 EBs and 45 NTCs along with the 384 samples for the fungal metabarcoding. 81% of EBs and 82.2% of NTCs are clean and show no or a negligible proportion (lower than 0.01%) of total reads. In 19% of EBs we detected 0.01–0.49% of the total reads, and in 17.8% of NTCs we detected 0.01–0.44% of the total reads (Supplement 10). We identified 13 different OTUs which are present at more than 10% of their abundance (samples + blanks) in the blanks (Supplement 11). We excluded these OTUs from the analysis and ran the RDA again. The RDA with the excluded OTUs is displayed in Supplements 13 and 14. The results are very similar to the RDA in Fig. 4 (RDA1: 11.95% and RDA2: 2.83%; with excluded OTUs: RDA 1: 12.15% and RDA2: 2.84%). Also, the distribution of the samples and the taxa is robust. Those OTUs which we found in the controls are mostly highly dominant in the samples which therefore can easily lead to cross-contamination during the laboratory work. As this happened in only a very few controls, we kept the OTUs in the dataset. Nonetheless, we cannot rule out entirely that the OTUs which do occur in the controls are partially contamination as these are mostly taxa which can be found ubiquitously (e.g. *Malassezia* can be found in soil but also on human skin).

3.4.5 Statistical analysis and visualization

We filtered the fungus dataset following Schiro et al. (2019) and assigned identified taxa to functional types according to their most probable role in the ecosystem (Schulze and Mooney, 2012). Mycorrhizal fungi include arbuscular mycorrhizae, ectomycorrhiza, and ericoid mycorrhizae. Other groups are saprotrophs, parasites, lichens, yeasts, and symbionts. A large number remained as “unknown” if their role in the ecosystem is not well understood. Identified plant taxa were assigned to either woody or herbaceous taxa.

For the statistical analysis, all data have been double-square rooted to better account for low-abundant taxa. All statistical analyses were implemented on percentage data using R, version 4.0.3 (R Core Team, 2020). Taxa were plotted colour-coded after their assigned functions. Plotting was done using the tidyverse package and ggplot2 (Wickham, 2016). To analyse differences in species diversity amongst the samples and locations, we calculated the alpha diversity using specnumber () of each sample from the resampled fungus and plant dataset.

To investigate relationships between fungi and vegetation, we used the functions cor.test () and cor (). First, we assessed whether there is a correlation between fungal OTU and ASV richness and plant taxon richness. Second, we related the fungal richness to the most significant vegetation PCA axis scores which were extracted from the PCA performed on the vegetation dataset. Finally, we applied the significant vegetation PCA axes as constraining variables in a redundancy analysis (RDA) performed on fungal compositional data using the function rda (). The scores of the vegetation PC axes were combined in a data frame using the function as.data.frame () to be used as the explanatory variable. For each axis, only taxa making up most of the separation of the axes were plotted with their names in the final RDA to not overload the plot. The significance of the vegetation PC axes was identified using PCAsignificance (). We used 10 samples from CH12 which are evenly distributed over the sediment record for the RDA to balance the weight of all lakes in the ordination. All ordination analyses were performed on double square-rooted data.

3.5 Results

3.5.1 Fungi: sedaDNA sequencing results and overall patterns of alpha diversity and taxonomic composition

In total, we obtained 52, 213, 129 paired read counts in the fungal dataset. After applying the OBItools pipeline, we retained 25,751 unique sequences with 32, 027, 606 counts. Clustering at a similarity threshold of 97% with *sumaclust* resulted in 5411 OTUs. Excluding OTUs with a similarity lower than 98% against the databases led to 716 remaining OTUs for the embl142 database, whereas the UNITE database returned 268 OTUs. After resampling to a base count of 5284 and subsequent filtering steps, 118 OTUs remained, covering 95.25% of the initial reads obtained after applying the OBItools pipeline. The filtered OTUs are regarded as “rare” and are not further assessed.

The highest OTU numbers before subsequent filtering of taxa are detected for CH12 (209 OTUs). This is followed by forested Bolshoe Toko (146 OTUs), Levinson Lessing (137 OTUs), and Lama (135 OTUs). The lowest OTU number is detected for the northern lake Kyutyunda (78 OTUs). The OTU richness of

single samples ranges from 3 OTUs (Kyutyunda, 30 cal ka BP) to 82 OTUs (CH12, 5.5 cal ka BP) with a mean of 23.53 OTUs. Samples from the Holocene show higher richness in comparison to samples from MIS2 and MIS3 in most lakes, while for Bolshoe Toko the overall OTU richness follows a decreasing trend.

The 10 most dominant taxa, summing up to 71% of the entire fungal dataset, are Pseudeurotiaceae (20%; 30 samples), *Mortierella* (13%; 63 samples), Sordariomyceta (11%; 26 samples), *Exophiala* (5.8%; 6 samples), *Oidiodendron* (5.6%; 10 samples), *Protoventuria* (5.5%; 14 samples), *Candida vartiovaarae* (3.1%; 7 samples), *Pseudeurotium* (2.7%; 9 samples), *Gryganskiella fimbricystis* (2.6%; 32 samples), and *Trichosporiella cerebriformis* (2.4%; 11 samples).

The most dominant functional type in the dataset are saprotrophs (40%; 38 OTUs), while yeasts are present at 10% (23 OTUs). Parasites (9.05%; 13 OTUs) and mycorrhizae (4.5%; 14 OTUs) are relatively rare. Least abundant are other symbionts (1.07%; 5 OTUs), lichens (0.2%; 4 OTUs), and mould (0.2%; 1 OTU). Fungi of unknown function comprise 24.2% (21 OTUs) of the dataset.

3.5.2 Vegetation: sedaDNA sequencing results and overall patterns of alpha diversity and taxonomic composition

In total, we obtained 48, 939, 032 reads for the vegetation data. Assembling of paired-end reads, demultiplexing into samples, and cleaning resulted in 152,194 unique sequence types with 20, 063, 932 counts. A total of 243 distinct taxa were obtained with 100% similarity to the database (Sønstebø et al., 2010; Willerslev et al., 2014; Soininen et al., 2015).

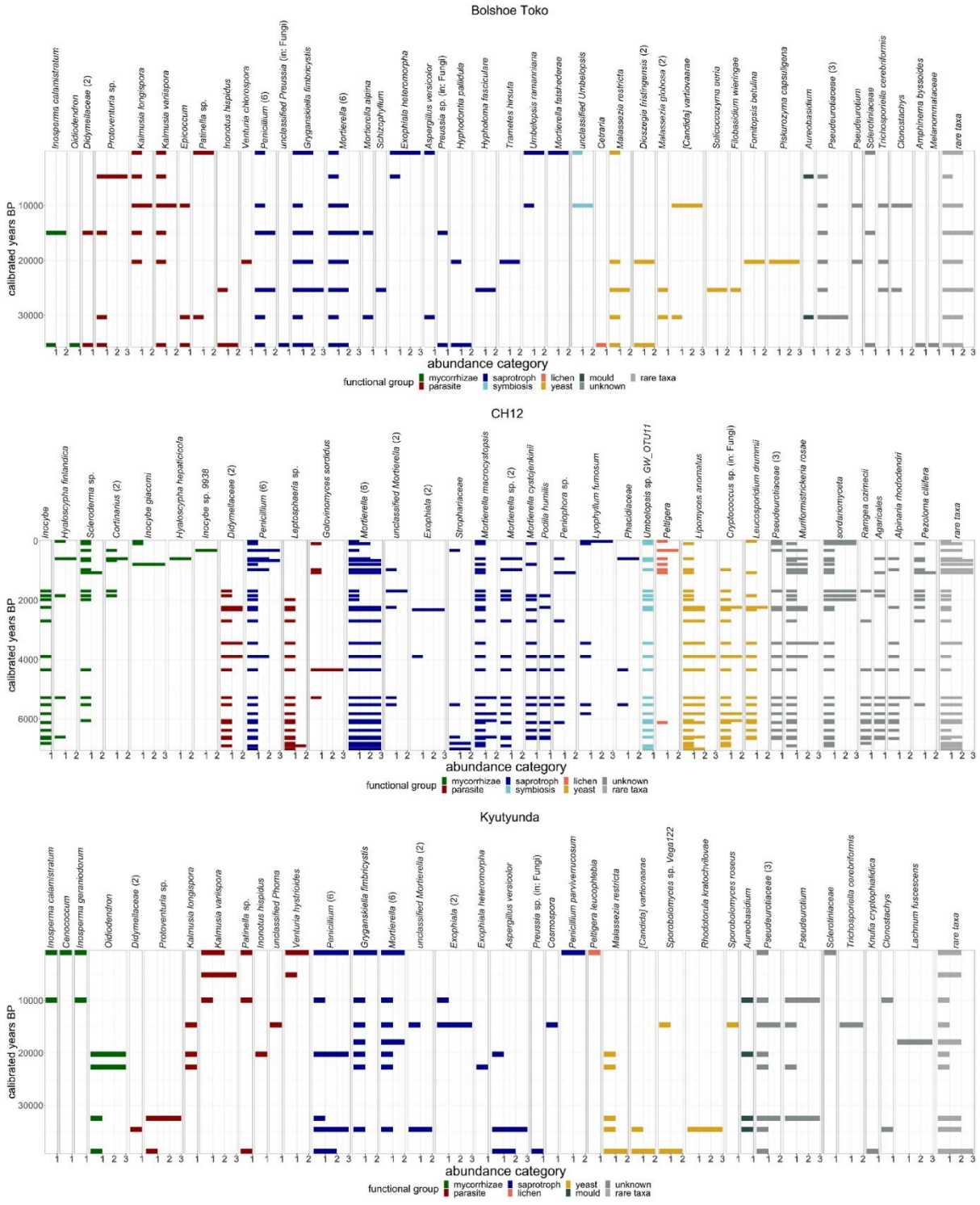
The comparison of taxa richness between the lakes reveals the highest number for Lama (163), followed by Bolshoe Toko (152) and Levinson Lessing (146). CH12 (138) and Kyutyunda (133) have the lowest numbers. The taxon richness of single samples varies between 9 (7 cal ka BP, CH12) and 112 (35 cal ka BP, Bolshoe Toko).

The most common plant taxa are Salicaceae (37.4%; 69 samples), *Dryas* (20.4%; 69 samples), *Larix* (5.94%; 44 samples), *Alnus alnobetula* (5.88%; 67 samples), *Papaver* (3.86%; 59 samples), *Menyanthes trifoliata* (3.83%; 45 samples), *Bistorta vivipara* (2.72%, 63 samples), Asteraceae (2.43%; 66 samples), *Betula* (1.6%; 67 samples), and *Anemone patens* (1.4%; 18 samples). These taxa constitute 85.5% of the dataset.

3.5.3 Site-specific plant-fungus covariation

3.5.3.1 Fungus and plant covariation in arctic Siberia from MIS3 to the Holocene

In Levinson Lessing (northern Taymyr Peninsula, tundra, 40–0 cal ka BP), the Pseudeurotiaceae (unknown function) as well as *Mortierella* and *Gryganskiella fimbricystis* (both saprotrophs) are highly abundant during MIS3 (Fig. 2). Around 38 cal ka BP, the Didymellaceae (parasites) also occur. At the end of MIS3, *Thamnolia vemicularis* (lichen) occurs. The most abundant plant taxa at this time are Salicaceae, *Dryas,* and *Papaver*. The most abundant fungus taxa in MIS2 are also Pseudeurotiaceae and *Mortierella*, but *Trichosporiella cerebriformis* (unknown function) also occurs often. For plants, the most dominant taxa are Salicaceae and *Papaver*, followed by *Dryas* at the end of MIS2 (Fig. 3). During the Holocene, *Mortierella* remains the most frequent fungal taxon but mycorrhizal OTUs (*Inosperma calamistratum, I. geraniodorum, Mallocybe fuscomarginata, Oidiodendron*) and parasites (Didymellaceae, *Kalmusia variispora*) become abundant as well. In the Holocene, there is a drastic decline in *Papaver* while *Alnus alnobetula* becomes highly abundant. *Dryas* as well as Salicaceae remain mostly abundant.


Bolshoe Toko
Inosperma calamistratum
Oidiodendron
Didymellaceae (2)
Protoventuria sp.
Kalmusia longispora
Kalmusia variispora
Epicoccum
Patinella sp.
Inonotus hispidus
Venturia chlorospora
Penicillium (6)
unclassified Preussia (in: Fungi)
Grygansiella fimbricystis
Mortierella (6)
Mortierella alpina
Schizophyllum
Exophiala heteromorpha
Aspergillus versicolor
Preussia sp. (in: Fungi)
Hyphodontia pallidula
Hypholoma fasciculare
Trametes hirsuta
Umbelopsis ramanniana
Mortierella fatshederae
unclassified Umbelopsis
Cetraria
Malassezia restricta
Dioszegia fristingensis (2)
Malassezia globosa (2)
[Candida] vartiovaarae
Solicoccozyma aeria
Filobasidium wieringae
Fomitopsis betulina
Piskurozyma capsuligena
Aureobasidium
Pseudeurotiaceae (3)
Pseudeurotium
Sclerotiniaceae
Trichosporiella cerebriformis
Clonostachys
Amphinema byssoides
Melanommataceae
rare taxa
calibrated years BP
10000
20000
30000
abundance category
functional group
mycorrhizae
parasite
saprotroph
symbiosis
lichen
yeast
mould
unknown
rare taxa
CH12
Inocybe
Hyaloscypha finlandica
Scleroderma sp.
Cortinarius (2)
Inocybe glaccmi
Hyaloscypha hepaticicola
Inocybe sp. 9938
Didymellaceae (2)
Penicillium (6)
Leptosphaeria sp.
Golovinomyces sordidus
Mortierella (5)
unclassified Mortierella (2)
Exophiala (2)
Strophariaceae
Mortierella macrocystopsis
Mortierella sp. (2)
Mortierella cystojenkinii
Podila humilis
Peniophora sp.
Lyophyllum fumosum
Phacidiaceae
Umbelopsis sp. GW_OTU11
Peltigera
Lipomyces anomalus
Cryptococcus sp. (in: Fungi)
Leucosporidium drummii
Pseudeurotiaceae (3)
Muriformistrickeria rosae
sordariomyceta
Ramgea ozimecii
Agaricales
Alpinaria rhododendri
Pezoloma ciliifera
rare taxa
0
2000
4000
6000
Kyutyunda
Inosperma calamistratum
Cenococcum
Inosperma geraniodorum
Oidiodendron
Didymellaceae (2)
Protoventuria sp.
Kalmusia longispora
Kalmusia variispora
Patinella sp.
Inonotus hispidus
unclassified Phoma
Venturia hystrioides
Penicillium (6)
Grygansiella fimbricystis
Mortierella (6)
unclassified Mortierella (2)
Exophiala (2)
Exophiala heteromorpha
Aspergillus versicolor
Preussia sp. (in: Fungi)
Cosmospora
Penicillium parvivenucosum
Peltigera leucophlebia
Malassezia restricta
[Candida] vartiovaarae
Sporobolomyces sp. Vega122
Rhodotorula kratochvilovae
Sporobolomyces roseus
Aureobasidium
Pseudeurotiaceae (3)
Pseudeurotium
Sclerotiniaceae
Trichosporiella cerebriformis
Knufia cryptophialidica
Clonostachys
Lachnum fuscescens
rare taxa

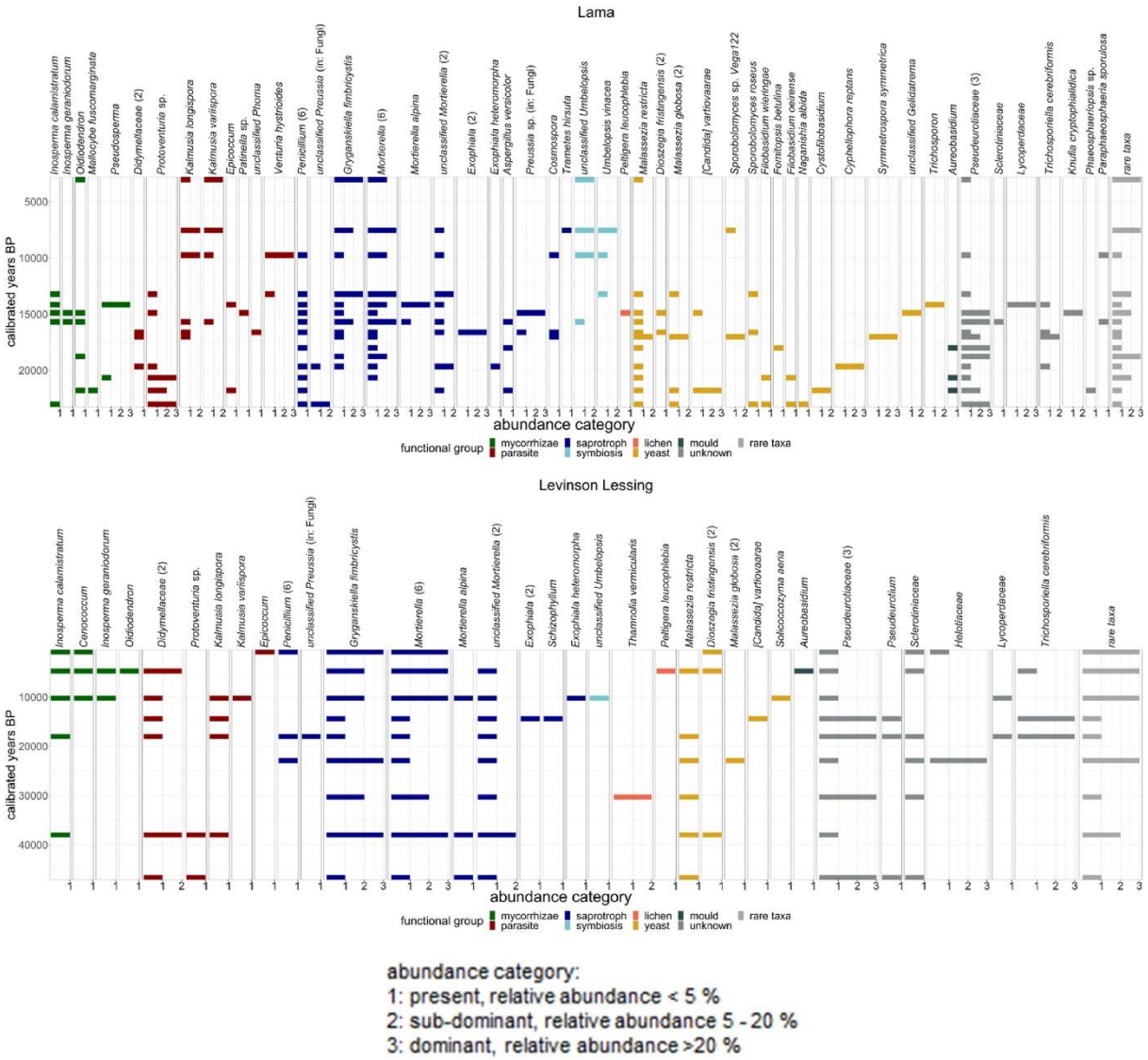


Figure 2: Site-specific fungus abundance displayed in abundance categories according to relative percentages. Fungi of the same functional type are colour-coded. The numbers in brackets give the OTUs detected for the specific taxon.

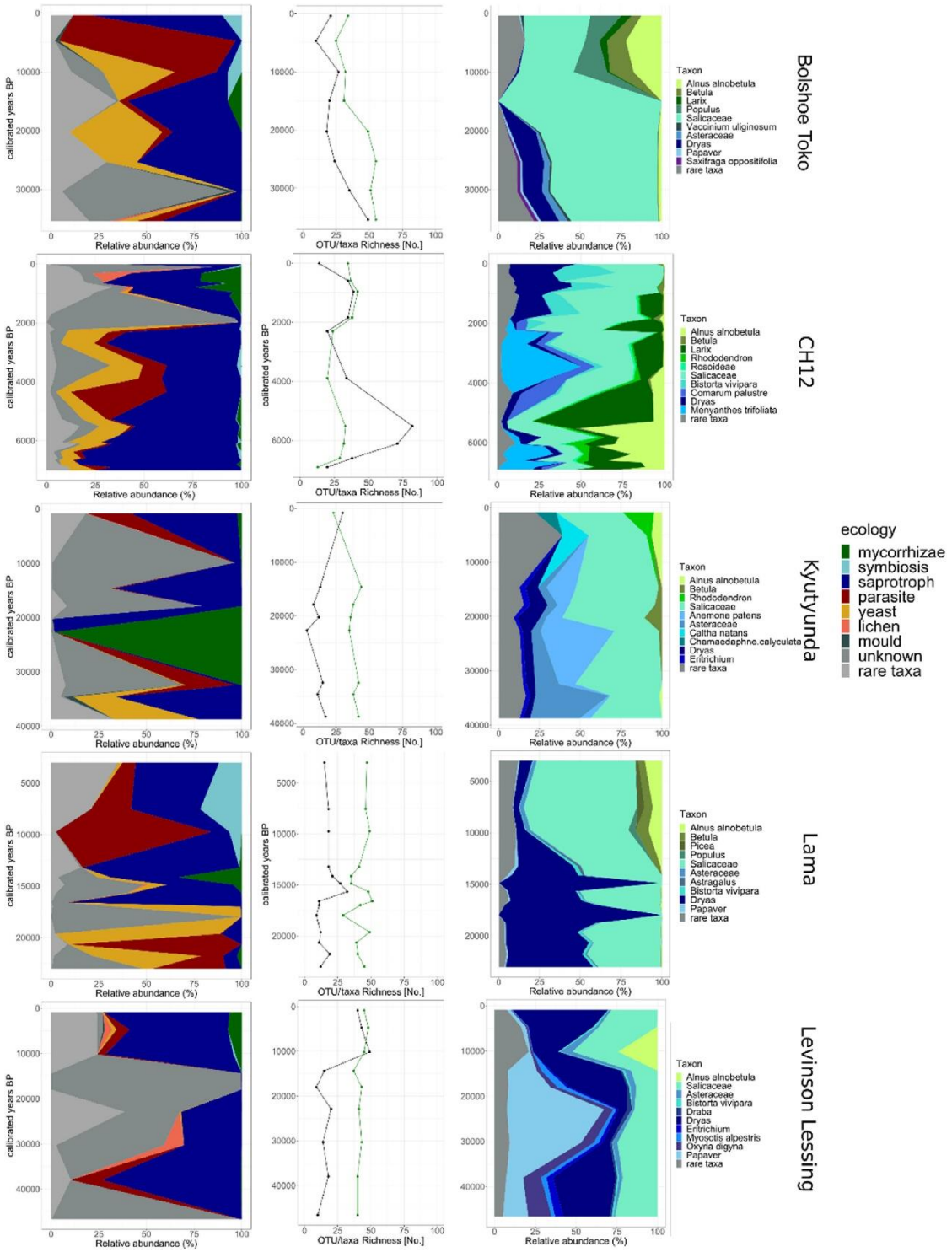


Figure 3: Fungal functional types in relation to fungal OTU richness and dominant plant taxa. Left column: distribution of fungal functional types for each lake. Middle column: fungal OTU richness of each lake (total OTU numbers), with the black line representing the fungal taxa while the green line marks the vegetation taxa for comparison. Right column: ten most dominant plant taxa of each lake.

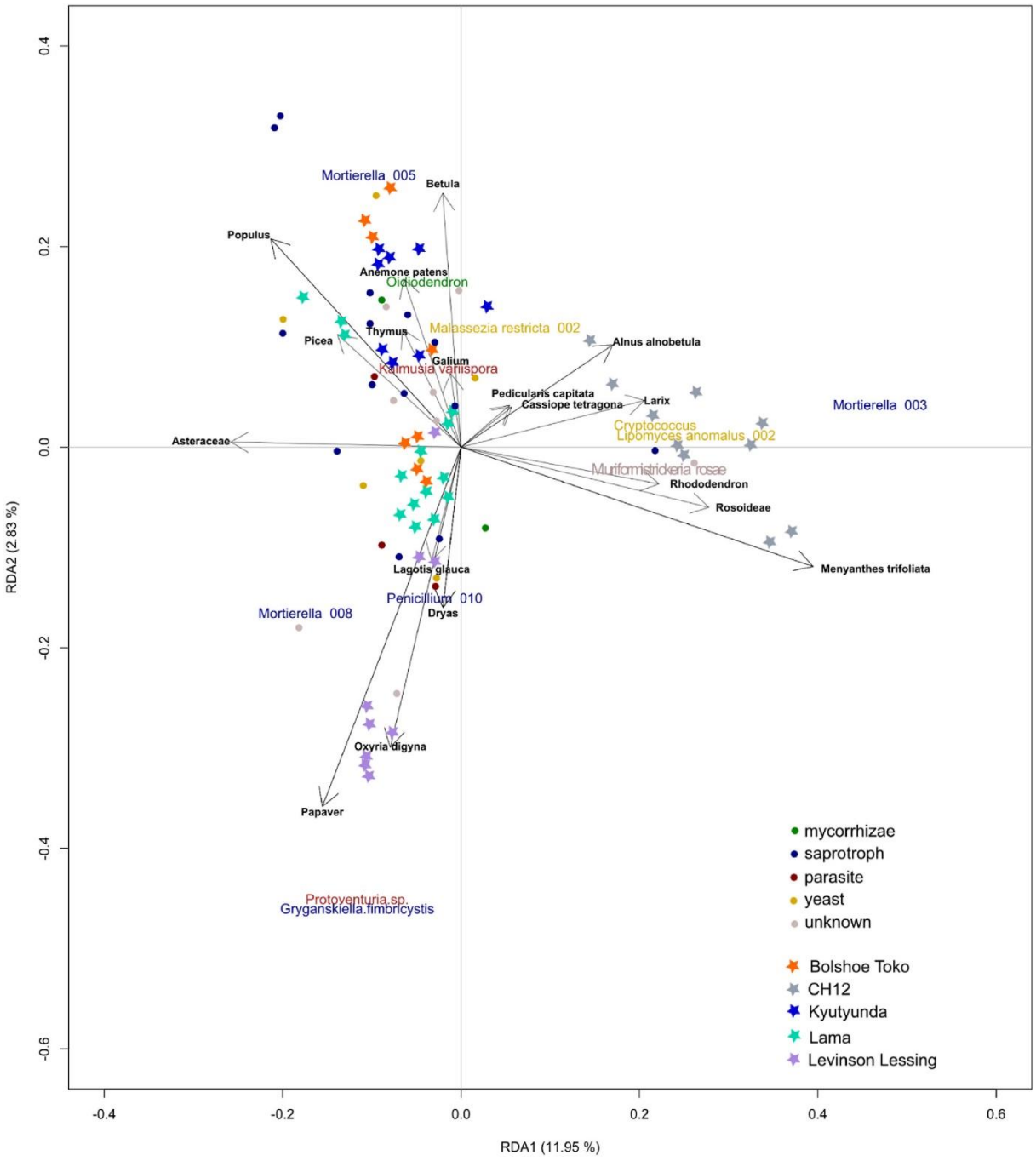


Figure 4: Fungal and plant co-variation displayed in a redundancy analysis (RDA). The most relevant principal component axes of the vegetation were determined, the scores extracted and then integrated into the RDA. The fungal taxa are displayed either with their names or as a dot colour-coded according to their functional group (see Fig. 2, Fig. 3). The plant taxa are marked with black arrows. The numbers after the taxa names indicate the specific OTU. The samples are displayed as stars and colour-coded according to their lake origin. The vegetation explains 20% of the fungus distribution.

CH12 (southern Taymyr Peninsula, tundra, 7–0 cal ka BP) only spans the mid to late Holocene. Around 7 cal ka BP, *Inocybe* (mycorrhizae) as well as *Golovinomyces sordidus* and Didymellaceae (parasites) are highly abundant. *Mortierella* is present throughout the whole record but shows strong declines when mycorrhizae and parasites are abundant around 5 cal ka BP (Fig. 2). Until 5.5 cal ka BP, *Alnus alnobetula* and Salicaceae are highly abundant. Woody taxa such as *Alnus alnobetula*, *Larix*, *Betula,* and *Rhododendron* have their highest abundances around 5 cal ka BP (Fig. 3). After 5 cal ka BP, an increase in yeast taxa (e.g. *Lipomyces anomalus, Cryptococcus*) is detected. This coincides with a decline in the aforementioned woody taxa. The lichen genus *Peltigera* is abundant in more recent times when the variety of mycorrhizal taxa also increases and *Inocybe*, *Hyaloscypha finlandica, Scleroderma, Cortinarius, Inocybe giacomi,* and *Hyaloscypha hepaticola* occur. Saprotrophic taxa such as *Mortierella* species, *Lyophyllum fumosum*, *Penicillium,* and *Exophiala* are present throughout the whole record.

Lama (northern Siberia, tundra-taiga transition zones, 24–0 cal ka BP) covers MIS2 and the Holocene. The most abundant fungal taxa during MIS2 are Pseudeurotiaceae, *Protoventuria* (parasite), *Mortierella*, and *Cyphellophora reptans* (yeast) (Fig. 2). *Dryas* as well as Salicaceae dominate the vegetation. Around the beginning of the Bølling/Allerød (15 cal ka BP), *Pseudosperma* and *Inosperma* species (mycorrhizae) become abundant. A little later, *Venturia hystrioides* and *Kalmusia* species (all parasites) start to occur. Salicaceae is still the most dominant plant taxon, but *Alnus alnobetula*, *Picea*, *Betula,* and *Populus* also frequently occur after 15 cal ka BP. Additionally, a drastic decline in *Dryas* took place after 15 cal ka BP (Fig. 3).

Kyutyunda covers the late MIS3 to the Holocene (northern Siberia, tundra-taiga transition zones, 38.8–0 cal ka BP). The most dominant fungal taxa during the late MIS3 are *Oidiodendron*, *Pseudeurotium* (unknown function), and *Penicillium* (saprotroph) (Fig. 2). During this time, Salicaceae and Asteraceae are the most abundant plant taxa but *Alnus alnobetula* also occurs (Fig. 3). *Oidiodendron* is mainly present at the end of MIS3. Shortly after, a large increase in *Betula* is detectable. In MIS2, the fungal taxa *Oidiodendron* and *Penicillium* are still highly prevalent and the taxon *Lachnum fuscescens* (unknown function) becomes common (Fig. 2). Salicaceae remains the most dominant plant taxon and *Betula* starts to occur more frequently. High abundance of *Dryas* as well as the first instances of *Alnus alnobetula* are detectable in the late MIS2 (Fig. 3). During the Holocene, *Pseudeurotium* (unknown function) and *Kalmusia variispora* became the most abundant fungal taxa. Salicaceae maintained its broad distribution while other woody taxa such as *Alnus alnobetula* and *Rhododendron* increased in their abundances.

Bolshoe Toko also spans the late MIS3 to the Holocene (central Yakutia, taiga, 35–0 cal ka BP). During the late MIS3, Pseudeurotiaceae are the most abundant fungal family but parasitic species (e.g. *Kalmusia* species, *Inonotus hispidus*) and saprotrophs (e.g. *Mortierella, Grykanskiella fimbricystis*) also occur (Fig. 2). At this time, Salicaceae is the most abundant plant taxon with *Dryas* occurring frequently. *Alnus alnobetula* and *Betula* are also present but at low abundance (Fig. 3). In MIS2, *Gryganskiella fimbricystis* and *Mortierella* are highly abundant fungi and a few yeast taxa (e.g. *Dioszegia fristringensis, Piskurozyma capsuligena*) start to occur. In late MIS2, *Inosperma calamistratum* (mycorrhizae) also occurs. Vegetation is still dominated by Salicaceae until the end of MIS2 with scarce abundances of *Alnus alnobetula* and *Betula*. In the Holocene, *Protoventuria* (parasite) is the most abundant fungal taxon but also *Kalmusia* species (parasite), *Exophiala heteromorpha* (saprotroph), and *Candida vartiovaarea* (yeast) are commonly found. A large increase in more diverse woody taxa is detected with more occurrences of Salicaceae as well as *Alnus alnobetula*, *Betula*, *Larix,* and *Populus*.

3.5.3.2 Quantitative relationships between fungi and plant richness and composition

In all records, we found only a weak borderline-significant correlation between fungi OTU and plant taxa richness (*r* 0.2394, *p*-value 0.098). For the fungal ASVs and plant taxa richness, the correlation is similar (*r* 0.2351, *p*-value 0.1039) Fungal richness is positively correlated to the sample scores of the first plant PCA axis (PC1: *r* 0.3863, *p*-value 0.006; ASVs: PC1 *r* 0.387, *p*-value 0.006) and negatively correlated to the sample scores of the second plant PCA axis (PC2: *r* −0.41, *p*-value 0.003; ASVs: PC2 *r* −0.424, *p*-value 0.002). The first axis reflects the differences between samples characterised by woody taxa including *Larix* and *Alnus alnobetula* and typical tundra taxa. On the second axis, we detected herbaceous plant taxa such as *Anemone patens* and *Thymus* positively correlating alongside other taxa preferring wetter habitats. Taxa such as *Oxyria digyna* and *Dryas*, which are associated with rather dry sites, show a negative correlation.

Sample scores of plant PCA axes 1–5 explain 20% of fungi composition (*p* 0.001) as revealed by RDA (Fig. 4 and Supplement 12). Woody taxa such as *Alnus alnobetula*, *Larix,* and *Rhododendron* appear in the upper right quadrant of the RDA plot together with the fungal taxa *Mortierella*_003, *Cryptococcus*, and *Muriformistrickeria rosae* (unknown function) (Fig. 3) and samples from CH12 aged 5.5 and 1.8 cal ka BP. The RDA also shows that parasitic fungi, such as Didymellaceae, and yeast, such as *Lipomyces anomalus* and *Cryptococcus*, tend to occur in the presence of woody taxa. Lichens occur predominantly in samples of Holocene age. *Papaver* and *Dryas* together with the fungal taxa Pseudeurotiaceae,

Grykanskiella fimbricystis (saprotroph), and *Mortierella* species occur in the lower left quadrant together with all samples from Levinson Lessing. *Populus* and *Ranunculus* in the upper left quadrant appear together with *Lipomyces anomalus* (yeast), *Cryptococcus* (yeast), and *Mortierella*. Samples here mostly originate from Bolshoe Toko although there are some from Lama (around the Bølling/Allerød period) and Kyutyunda (Holocene). Samples from the Holocene all occur in the upper half of the RDA where woody plant taxa are found and a broader fungal species richness is detected. In general, Lake CH12 shows a unique fungal composition in comparison to the other lakes. The samples can be found in the right quadrants of the RDA while the samples of the other lakes are located in the left quadrants or centred (Fig. 4 and Supplement 6).

3.6 Discussion

3.6.1 Fungus and plant diversity along a spatiotemporal gradient in Siberia

Assessing the species richness in an environment enables the determination of community diversity and can be an indicator for ecosystem turnover (Hillebrand et al., 2018). After applying metabarcoding on 70 samples from five lakes across Siberia, we detected high fungal richness (706 OTUs) while the analyses of plant richness yielded 243 distinct taxa. In comparison to Liu et al. (2020) who investigated plant species richness in lake sediments in north-eastern Siberia, we detected slightly higher plant diversity (their study: 90 to 120 taxa in a single lake, this study: 133 to 163 taxa in a single lake) which might be explained by the sediments from the present study covering a longer time span and therefore more diverse climate scenarios. To assess fungal richness, we used OTU clusters instead of sequence variants, which might lead to under- or overestimation in comparison to species assessment (Frøslev et al., 2017). Underestimation may also originate from missing reference material of local taxa in databases (Goodwin et al., 2016; Quince et al., 2009). Comparably, a modern species assessment from the western Ural yielded 376 observed fungal species (Palamarchuk and Kirillov, 2019), which supports the conclusion of Seeber et al. (2022) that their marker is suitable to assess diversity even on long time-scales. Nonetheless, higher OTU richness (1125 OTUs in 55 samples) was obtained by Talas et al. (2021) in their study of a Holocene lake sediment core from eastern Latvia. Their discovery of higher overall diversity is explained by inclusion and detection of aquatic fungi (23%, terrestrial 40%), which are mostly lacking in our data. Additionally, they kept very short reads and included reads with fewer counts (4 instead of 10 in our study).

Bolshoe Toko (146 OTUs) and Lama (135 OTUs) are in forested areas and show higher fungal OTU richness compared to Kyutyunda (78 OTUs) from the northern tundra (Fig. 3). A relationship between fungal richness and vegetation composition has been shown by multiple studies (e.g. Tedersoo et al., 2013; Geml et al., 2017), however data from the Siberian treeline are lacking. Our data concur that ectomycorrhizal fungal richness is highest with forest cover (Geml et al., 2017). Spatial fungal richness is confirmed by the temporal relationship: we observed co-occurrences of high fungal richness and woody vegetation. Levinson Lessing shows a large increase in fungal OTU richness and woody taxa dominance during the warm Holocene compared to the late Glacial although experimental warming did not result in higher fungal diversity (Geml et al., 2015; Mundra et al., 2016). Talas et al. (2021) show high richness as well as community turnover with increases in plankton parasitic species and mycorrhiza after 4 cal ka BP, suggesting that fungi with more specific hosts or substrates (e.g. ectomycorrhizae) are more susceptible to ecosystem changes than taxa with wide preferences. CH12 shows higher OTU richness than the other lakes, even when considering similar sample numbers, supporting the hypothesis that fungal communities from the warmer Holocene might be more species rich. Potentially, warming-induced vegetation responses rather than direct warming shape the diversity in fungal communities, suggesting a broadening diversity of fungi species alongside future treeline migration.

Metabarcoding on arctic tundra communities reveals that each specific tundra type has a unique community of associated fungi (Wallenstein et al., 2007; Geml et al., 2021). This might explain the overall highest fungal OTU richness originating from CH12 (dry forest tundra), while Kyutyunda (wet southern tundra) shows a rather low richness. Furthermore, our analysis shows a negative correlation between fungal richness and the second vegetation PC axis, covering a wetness gradient from species related to drier areas (high PC scores) to species rather related to wetter areas.

A modern spatial study on the Tibetan plateau showed that fungal richness is positively correlated with plant richness (Yang et al., 2017). Interestingly, through statistical analyses, we find only a weak positive correlation between fungal and plant richness. Most probably, this indicates a quite complex relationship between plant richness and vegetation composition. It is also known that other biotic or abiotic factors such as the bacterial composition or soil N and C content influence the fungal and plant communities (Singh et al., 2009). Besides, the shorter amplicon length of the plant compared to the fungal marker can cause biases when amplifying highly fragmented DNA, resulting in a weak statistical relationship. Incomplete databases for arctic fungi might also lead to underestimation of taxa richness. For plants, it is known that the catchment influences the record quality if the plants are growing closer to the soil surface (Giguet-Covex et al., 2019). Potentially, some fungi are also more likely to end up in the sediment of the lake if they are growing in the upper soil horizons with their DNA transported to

the lake either via animals or rainfall. At the broad scale, plant richness decreases with latitude (Kerkhoff et al., 2014). However, a modern study from Kamchatka, Russian Far East, reports highest plant species richness in alpine tundra and snowbed communities (Doležal et al., 2013). Recent sedaDNA studies from the treeline in Chukotka, Russian Far East, and from Bolshoe Toko showed highest terrestrial plant richness in the late Pleistocene in steppe-tundra areas and lowest in the forested Holocene (Huang et al., 2020; Courtin et al., 2021). This indicates that a high correlation between fungus and plant alpha diversity cannot be expected.

3.6.2 Changes in ecosystem functioning over a spatiotemporal gradient

To date, molecular analyses of ecosystem functioning that trace fungus-plant covariation have been addressed by multiple modern studies. These studies mainly focus on the organisms required for modern plant establishment, that is, over spatial (Merges et al., 2018) and short temporal gradients (Zhang et al., 2016) or under varying growing conditions (Mohamed and Martiny, 2011). In palaeo-research, plant communities have been subject to molecular analysis such as metabarcoding studies (Liu et al., 2020), shotgun sequencing (Parducci et al., 2019) or target capture (Murchie et al., 2021). So far, the turnover of entire ecosystems tracing not only the plant constituent but also their associated fungal symbionts has not yet been studied. Our data form not only one of the first molecular biological studies revealing fungus community changes over a large temporal gradient but also allow conclusions to be drawn on their long-term impact on forest establishment. The following examples highlight the impact of vegetation changes alongside climate change on fungus ecological functionality and subsequent whole ecosystem turnover.

3.6.2.1 Long-term mutualism in arctic environments inevitable for plant establishment

Most mycorrhizal taxa detected are from the families Cortinariaceae and Inocybaceae and some from Myxotrichaceae and Hyaloscyphaceae (Fig. 2), in agreement with previous metabarcoding studies (Nilsson et al., 2005; McGuire et al., 2013; Botnen et al., 2014). We retrieved *Inocybe* (including the subgenus *Inosperma* (Matheny et al., 2020)) and *Cortinarius,* both known from high latitudes (Timling et al., 2012). They represent ectomycorrhizal associates of arctic tundra and shrubs including *Salix* and *Dryas integrifolia* (Ryberg et al., 2009; Botnen et al., 2014), both being common taxa in our plant data. Our data also support previous studies from boreal forests (McGuire et al., 2013), including a study from the Russian Far East detecting *Cortinarius*, in addition to *Lactarius* and *Russula*, as important

ectomycorrhizal of *Larix gmelinii* (Miyamoto et al., 2021). *Cortinarius* also associates with shrubs of Salicaceae and Rosaceae as well as herbaceous Cyperaceae (Garnica et al., 2005), which frequently occur in our plant dataset, indicating a broad variety of host taxa for these fungi, from which we can infer high adaptability towards warming and changing overall environmental conditions.

After the Last Glacial, vegetation species richness decreased as well as arbuscular mycorrhizal taxa while ectomycorrhiza associated with woody taxa and non-mycorrhizal fungi increased (Zobel et al., 2018). This resulted in changes in the mutualist trait structure after the Last Glacial Maximum, making mycorrhizal associations important factors when predicting ecosystem responses to changing environmental conditions. We also observe increasing fungal richness in Holocene samples (Fig. 3). This underlines the suitability of sedaDNA fungal metabarcoding studies for appropriate ecosystem reconstructions and when considering adaptation mechanisms alongside ecosystem turnover.

Interestingly, we observed highest values of Pinaceae only after the presence of mycorrhizal taxa (e.g. *Cortinarius*, *Inosperma calamistratum*) (Fig. 2, Fig. 3), although this might be due to low sample numbers. Without mycorrhizal fungi, Pinaceae growth is restrained or establishment is inhibited as nutrient uptake is impossible (Marschner and Dell, 1994). Studies from Japan (Ishida et al., 2007) and temperate areas in the Himalaya (Pande et al., 2004) revealed Cortinariaceae as the main ectomycorrhizal associates of Pinaceae, strengthening the precision of our dataset and its ability to correctly recover fungal-plant covariation over long time scales and its possibility to assess ecosystem dynamics. Our analysis also highlights the longevity of the dependency of Pinaceae on these particular fungi.

3.6.2.2 Wood-decaying species highly impacted by warming

We found *Mortierella*, *Penicillium,* and *Exophiala* as the main biomass-decaying taxa (Fig. 2). These are common soil fungi in high-latitude ecosystems (Treseder et al., 2007; Allison et al., 2009) and are reported amongst the main soil fungi in arctic tundra soils (Kurek et al., 2007; Zhang et al., 2016) due to their cold tolerance. *Mortierella* associates with *Vaccinium uliginosum*, *Betula nana*, *Salix glauca*, *Empetrum nigrum*, and *Cassiope tetragona* (Voříšková et al., 2019), which are typical taxa in our study. Rhizosphere samples from *Larix sibirica* and *Betula pendula* from Krasnoyarsk, Siberia revealed *Penicillium* as one of the main constituents (Boyandin et al., 2012). *Larix* forests growing on permafrost show broad host spectra towards saprotrophs (Leski and Rudawska, 2012) as a response to changing environment, for example after wildfires (Miyamoto et al., 2021), which explains the overall broad distribution of saprotrophs after warming in the area.

Saprotrophs are generally highly abundant throughout all records. Their significant decrease around 10 cal ka BP (Fig. 2, Fig. 3) demonstrates that the climate change during the Pleistocene/Holocene transition (Biskaborn et al., 2016, 2019a) also affected soil communities alongside vegetation. This finding agrees with results from experimental warming studies in boreal ecosystems, indicating that relative saprotroph abundance declines with warming, while the abundance of mycorrhizal fungi and lichens increases, underlining long-scale ecosystem turnover as a response to warming (Deslippe et al., 2012; Geml et al., 2015; Mundra et al., 2016).

3.6.2.3 Host-specific parasites show strong co-occurrence with woody taxa

We detected parasitic OTUs mostly in samples from the warm Holocene (Fig. 2), confirming early findings that experimental warming leads to increases in parasitic and virulent fungi (Geml et al., 2015) along with woody taxa expansion. The most abundant parasitic species from our dataset are *Protoventuria*, *Kalmusia variispora, K. longispora,* and Didymellaceae which co-occur with Salicaceae, *Larix,* and *Alnus alnobetula* (Fig. 2, Fig. 3). In shrubby tundra in Greenland with *Salix* occurrences, *Venturia* species are amongst the highest abundant fungi (Voříšková et al., 2019), indicating a strong covariation between these taxa. Interestingly, we observe a decline in Salicaceae after *Protoventuria* abundance around 20 cal ka BP (Fig. 2, Fig. 3), supporting previously noted fungal parasite abundances in permafrost during the LGM (Lydolph et al., 2005). Venturiaceae has been assigned to Salicaceae as pathogens in northern latitudes (Hosseini-Nasabnia et al., 2016), while *Kalmusia* has been detected in *Alnus* forests (Iznova and Rukšėnienė, 2012). Didymellaceae co-occurs with a broad range of host plants such as *Larix decidua* (Chen et al., 2017). The RDA reveals that *Kalmusia* species preferentially occur in forested areas alongside saprotrophic and mycorrhizal species (Fig. 4), supporting the value of our data and the feasibility of co-occurrence analysis in sedaDNA studies and their potential when assessing ecosystem dynamics up to species level. Plant-parasite interplay in relation to climate change is not fully understood (Burdon and Zhan, 2020) but it is assumed that parasitic fungi are more specific in their hosts than mycorrhizal taxa, making them a great target when assessing ecosystem dynamics and turnover (Põlme et al., 2018).

3.6.2.4 Lichens influence soil carbon dynamics and local fauna

The recovered lichen OTUs belong to 16 families with the highest richness in Peltigeraceae and Parmeliaceae. The most abundant genera are *Thamnolia*, *Peltigera,* and *Cetraria,* all of them being

common in northern Siberian communities (Zhurbenko and Yakovchenko, 2014) and permafrost (Lydolph et al., 2005). *Thamnolia* species often occur in arctic tundra (Sheard, 1977), showing low specificity concerning their photobiont while associating with various *Trebouxia* species (Nelsen and Gargas, 2009). *Peltigera* preferentially grows in temperate regions on soils and among mosses over rocks, but also on tree trunks (Nash, 2002) and in boreal forests (Asplund and Wardle, 2015), explaining their abundance in the forested Holocene in CH12 (Fig. 2).

Our analyses are among the few palaeoecological studies detecting lichens (Fig. 2). Lichens are commonly missing from fossil records (Taylor and Osborn, 1996) despite being an important component of boreal forest and tundra biomass (Asplund and Wardle, 2017; Shevtsova et al., 2020). However, we could only detect a few reads, belonging to 48 OTUs (<1% of the whole dataset).

Unexpectedly, most lichens are recorded from warm periods with well-developed vegetation (late MIS3 and Holocene). Lichens are a prominent feature of arctic landscapes and short-term experimental warming in the Canadian arctic led to their decline (Fraser et al., 2014). For Siberia, lichens have been recovered along a broad latitudinal gradient with high diversity and biomass (Safronova and Yurkovsksya, 2019). Lichen cover on permafrost produces a cooling effect, making lichens of great importance when considering thawing effects (Porada et al., 2016). Reduced or no lichen cover during the glacial might be relevant for past soil carbon dynamics. Lichens tolerate high percentages of CO_2 (Badger et al., 1993), but studies about the impact of low CO_2 supply are missing. Possibly, they suffer more than other fungi from reduced atmospheric CO_2 content as lichens also have to supply their algal or cyanobacterial symbionts.

Lichen distribution also impacts the occurrence of animals such as *Moschus moschiferus*, which preferentially settle in lichen-rich habitats for their food supply (Slaght et al., 2019). Increased lichen coverage during the Holocene may have supported the compositional turnover in the megaherbivore fauna. Reindeer mostly feed on lichen but changing environmental conditions might impact their distribution and diet to include less lichen (Drucker et al., 2011) or to vary seasonally (Bocherens et al., 2015), giving them higher survival advantage. Changing fungus communities will thus not only impact the boreal forest, but also its fauna.

3.6.2.5 Habitat-loss of fungal species due to warming resulting in feedback on whole ecosystem

The most abundant yeast taxa in our dataset are *Candida vartiovaarae*, *Malassezia restricta*, *Cyphellophora reptans*, *Cryptococcus*, and *Lipomyces anomalus* (Fig. 2), which are widely distributed

in Siberian soils (Polyakova and Chernov, 2001). *Candida vartiovaarae* is broadly present in forest as well as in grassland soils (Yurkov et al., 2012), while *Cryptococcus* is associated with peatland (Thormann, 2006) and boreal swamps (Kachalkin and Yurkov, 2012). A correlation between *Malassezia* species and nematodes in central European forests was discovered, suggesting that nematodes act as vectors for the fungi (Renker et al., 2003). To investigate these zoophilic relationships and their contribution to ecosystem stability, further metabarcoding data on small soil organisms could be an asset.

Whenever yeasts are highly abundant in the records, especially in colder time periods like the LGM, mycorrhizae decrease (Fig. 2, Fig. 3). Most yeasts show adaptive responses when temperatures drop to maintain their survival (Kandror et al., 2004). Experimental warming also shows yeast decline with rising temperatures (Treseder et al., 2016), indicating that some species will lose their habitats with ongoing warming, resulting in a major feedback to the ecosystem, potentially leading to shifts in the entire ecosystem and subsequent turnover from, for example, tundra to taiga.

From our data, it is not possible to determine the role of yeast in soil. Generally, they serve both as biotoxins (Santos et al., 2004; Compant et al., 2005) or growth promoters for plants (Nassar et al., 2005; El-Tarabily and Sivasithamparam, 2006). In Siberia, yeasts might either function as plant parasites (Hernández-Fernández et al., 2021) or as biodegraders, as after a period of high yeast abundance, we detect decreasing woody taxa. Further research on modern mutualistic and parasitic interactions in the area will help to solve this research gap and to understand yeast impact on long-term ecosystem stability.

3.6.3 Implications of our results for ecosystem functioning and future research avenues

The interplay between climate, vegetation, fungi, and microorganisms in the boreal forest ecosystem is not yet understood. As fungi are a key component of ecosystem functioning, a major impact on future ecosystem-climate feedback is expected alongside compositional change and varying soil microbiome (McCalley et al., 2014). To our knowledge, we conducted the first study on fungus-plant interactions and co-occurrences in the palaeo context, assessing community shifts in boreal forests as well as tundra ecosystems. However, our results are only a first proxy on future community changes as the magnitude of warming differs strongly between our samples and present warming and any relationship may incorporate lagged responses over large time-scales (Biskaborn et al., 2021).

To our knowledge, this is the first long-term dataset showing antagonistic relationships among fungal functional types as well as warming-related vegetation change related to fungus diversity and composition changes. By analogy to the past, future woody taxa advance into arctic regions might result in higher fungus diversity and a relative increase in mycorrhizae, parasites, and potentially lichens at the cost of saprotroph and yeast abundance.

Our study design does not allow a definite conclusion to be drawn on whether future treeline advances will rely on the presence of specific fungal communities. As ectomycorrhizal communities in sub-arctic tundra are generally species-rich and do not show high host preferences (Ryberg et al., 2009, 2011), major changes may not ensue. However, the investigated soils in the sub-arctic already have long histories of soil development, unlike the northern tundra sites and upper mountain areas. Temperature wise, these are potential habitats for forest establishment but might not be favourable for diverse soil fungus composition due to a lack of nutrients.

Lichens do not generally suffer from warming but are affected by the vegetation. The observed decline in lichens with denser canopy cover (Cornelissen et al., 2001) may only be relevant to the more southerly forests. As our study only returned a few lichen OTUs, it is not possible to draw a robust conclusion here. High CO_2-concentrations during experimental darkening leads to generally quick CO_2 uptake by the genus *Peltigera* and subsequently relatively slow release (Badger et al., 1993), making lichens potentially valuable for the storage of future warming-induced CO_2 from soil. Further research into lichens is promising to delve into mechanisms supporting ecosystem adaptation towards changing environments.

Besides the limitation in temporal resolution, our study suffers from limited taxonomic resolution and complex abundance patterns. SedaDNA metabarcoding is highly susceptible to damage and degradation, leading to biases in PCR products as taxa might be dismissed due to short lengths (Coissac et al., 2012; Taberlet et al., 2012). Sometimes, reference genomes are missing and identification at the relevant taxonomic resolution is not possible (Sønstebø et al., 2010). Also, different taxa possess varying amounts of genome copy numbers per cell which might lead to overrepresentation of taxa with high copy numbers while rare taxa can be missed (Behnke et al., 2011). To strengthen the metabarcoding data, a further comparison to target capture similar to Murchie et al. (2021) but on the fungal DNA of the same samples would be an asset to validate the recovered abundances and diversity.

3.7 Conclusions

This is the first study showing spatial and temporal changes in palaeo fungus-plant covariation. Knowing which fungi influenced the growth of specific plant communities in the past will help to predict future community turnover due to varying climate. To understand palaeo community turnover in more detail, it is necessary to consider a plant's associated heterotrophic organisms in present times. This will help to place knowledge gained in this study into a better context. Additionally, our data are a great asset to existing knowledge about boreal forests as they help to shed light on adaptation mechanisms of plants towards warming and their subsequent northward migration. Nevertheless, there are still many ecological interactions that are unknown which need to be addressed in future research, such as which organisms contribute to the rhizospheres of specific plants and whether or how these associations change with varying climate or how the fauna is impacted by a changing habitat and food source. This might help the development of future afforestation and silviculture strategies. Despite this, our findings will already help the assessment of future tipping points in boreal forest stability.

Funding

This research has been funded by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (Grant Agreement No. 772852, ERC Consolidator Grant “Glacial Legacy”) and the Initiative and Networking Fund of the Helmholtz Association. The record from lake Levinson Lessing was recovered and initially processed with funds of the BMBF project ‘PLOT - Paleolimnological Transect’ (grant no. 03F0830A).

Availability of data and material

The data are available under Dryad and Pangaea and will be publicly available after the acceptance of the manuscript.

Pangaea: Metadata of the cores and links to existing Pangaea entries. doi: 10.1594/PANGAEA.948180

Dryad: Fungal and Plant DNA Datasets (Raw data and scripts). doi:10.5061/dryad.05qfttf3x

Author contribution

The study was designed by BvH, KSL, UH; KSL supervised and BvH conducted the experimental lab work; BvH analysed the data under supervision of UH and KSL; LS sampled the cores and supervised the DNA extractions; PS and LE performed the bioinformatic evaluation of the marker; MM retrieved the sediment core of Lake Lama; BD led the projects on Kyutyunda and Bolshoe Toko; BB retrieved and dated the sediment cores including age depth modelling of Bolshoe Toko, Kyutyunda and CH12; BvH dated Lama and performed the age modelling; BvH supervised by UH wrote a first version of the manuscript; all authors commented on the first and revised version of the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are very grateful for the funding of the research through the project grant “Glacial Legacy”, funded by the European Research Council (ERC-2017-COG; project reference 772852). The work is supported by the Russian Ministry of Education and Science No. FSRG 2020-0019. We would like to thank everyone involved in the several field campaigns retrieving the lake sediment cores, especially Luidmila Pestryakova and the team of NEFU Yakutsk for the organisation of the field work. We thank Svetlana Karachurina and Sarah Olischläger for their help conducting the laboratory part of the DNA extractions and the vegetation metabarcoding and Gregor Pfalz for fruitful discussions about the age model. We are also grateful to Håvard Kauserud for the revision of the ecological interpretation of the fungal taxa. The radiocarbon dating has been done by MICADAS (AWI Bremerhaven). We thank Cathy Jenks for the English proof-reading of the manuscript.

3.8 References

Allison, S.D., LeBauer, D.S., Ofrecio, M.R., *et al*., 2009. Low levels of nitrogen addition stimulate decomposition by boreal forest fungi. *Soil Biol. Biochem*. **41**, 293–302. https://doi.org/10.1016/j.soilbio.2008.10.032

Alsos, I.G., Lammers, Y., Yoccoz, N.G., *et al*., 2018. Plant DNA metabarcoding of lake sediments: How does it represent the contemporary vegetation. *PLOS ONE* **13**, e0195403. https://doi.org/10.1371/journal.pone.0195403

Andreev, A.A., Tarasov, P.E., Klimanov, V.A., *et al.,* 2004. Vegetation and climate changes around the Lama lake, Taymyr Peninsula, Russia during the Late Pleistocene and Holocene. *Quat. Int.* **122**, 69–84. https://doi.org/10.1016/j.quaint.2004.01.032

Anisimov, M.A., Pospelov, I.N., 1999. The landscape and geobotanical characteristics of the Levinson-Lessing lake basin, Byrranga mountains, central Taimyr, in: Kassens, H., Bauch, H.A., Dmitrenko, I.A., Eicken, H., Hubberten, H.-W., Melles, M., Thiede, J., Timokhov, L.A. (Eds.), Land-Ocean Systems in the Siberian Arctic: Dynamics and History. *Springer*, Berlin, Heidelberg, pp. 307–327. https://doi.org/10.1007/978-3-642-60134-7_27

Asplund, J., Wardle, D.A., 2017. How lichens impact on terrestrial community and ecosystem properties. *Biol. Rev*. **92**, 1720–1738. https://doi.org/10.1111/brv.12305

Asplund, J., Wardle, D.A., 2015. Changes in functional traits of the terricolous lichen *Peltigera* aphthosa across a retrogressive boreal forest chronosequence. *The Lichenologist* **47**, 187–195. https://doi.org/10.1017/S0024282915000092

Atlas Arktiki, 1985. Glav. Upravl. Geodezii i Kartografii pri Sovete Ministrov SSSR.

Badger, M.R., Pfanz, H., Büdel, B., *et al*., 1993. Evidence for the functioning of photosynthetic CO_2-concentrating mechanisms in lichens containing green algal and cyanobacterial photobionts. *Planta* **191**, 57–70. https://doi.org/10.1007/BF00240896

Baldrian, P., Valášková, V., 2008. Degradation of cellulose by basidiomycetous fungi. *FEMS Microbiol. Rev.* **32**, 501–521. https://doi.org/10.1111/j.1574-6976.2008.00106.x

Behnke, A., Engel, M., Christen, R., *et al*., 2011. Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions. *Environ. Microbiol*. **13**, 340–349. https://doi.org/10.1111/j.1462-2920.2010.02332.x

Bellemain, E., Carlsen, T., Brochmann, C., *et al*., 2010. ITS as an environmental DNA barcode for fungi: an *in silico* approach reveals potential PCR biases. *BMC Microbiol.* **10**, 189. https://doi.org/10.1186/1471-2180-10-189

Bellemain, E., Davey, M.L., Kauserud, H., *et al*., 2013. Fungal palaeodiversity revealed using high-throughput metabarcoding of ancient DNA from arctic permafrost. *Environ. Microbiol*. **15**, 1176–1189. https://doi.org/10.1111/1462-2920.12020

Biskaborn, B.K., Nazarova, L., Kröger, T., *et al*., 2021. Late Quaternary Climate Reconstruction and Lead-Lag Relationships of Biotic and Sediment-Geochemical Indicators at Lake Bolshoe Toko, Siberia. *Front. Earth Sci*. **9**, 703. https://doi.org/10.3389/feart.2021.737353

Biskaborn, B.K., Nazarova, L., Pestryakova, L.A., *et al*., 2019a. Spatial distribution of environmental indicators in surface sediments of Lake Bolshoe Toko, Yakutia, Russia. *Biogeosciences* **16**, 4023–4049. https://doi.org/10.5194/bg-16-4023-2019

Biskaborn, B.K., Smith, S.L., Noetzli, J., *et al*., 2019b. Permafrost is warming at a global scale. *Nat. Commun.* **10**, 264. https://doi.org/10.1038/s41467-018-08240-4

Biskaborn, B.K., Subetto, D.A., Savelieva, L.A., *et al*., 2016. Late Quaternary vegetation and lake system dynamics in north-eastern Siberia: Implications for seasonal climate variability. *Quat. Sci. Rev*. **147**, 406–421. https://doi.org/10.1016/j.quascirev.2015.08.014

Blaauw, M., Christen, J.A., 2011. Flexible paleoclimate age-depth models using an autoregressive gamma process. *Bayesian Anal*. **6**, 457–474. https://doi.org/10.1214/11-BA618

Blume-Werry, G., Milbau, A., Teuber, L.M., *et al*., 2019. Dwelling in the deep – strongly increased root growth and rooting depth enhance plant interactions with thawing permafrost soil. *New Phytol.* **223**: 1328-1339. https://doi.org/10.1111/nph.15903

Bocherens, H., Hofman-Kamińska, E., Drucker, D.G., *et al*., 2015. European bison as a refugee species? Evidence from isotopic data on early Holocene bison and other large herbivores in northern Europe. *PLOS ONE* **10**, e0115090. https://doi.org/10.1371/journal.pone.0115090

Botnen, S., Vik, U., Carlsen, T., *et al*., 2014. Low host specificity of root-associated fungi at an Arctic site. *Mol. Ecol*. **23**, 975–985. https://doi.org/10.1111/mec.12646

Boyandin, A.N., Rudnev, V.P., Ivonin, V.N., *et al*., 2012. Biodegradation of polyhydroxyalkanoate films in natural environments. *Macromol. Symp*. **320**, 38–42. https://doi.org/10.1002/masy.201251004

Boyer, F., Mercier, C., Bonin, A., *et al*., 2016. obitools: a unix-inspired software package for DNA metabarcoding. *Mol. Ecol. Resour*. **16**, 176–182. https://doi.org/10.1111/1755-0998.12428

Bradshaw, C.J.A., Warkentin, I.G., Sodhi, N.S., 2009. Urgent preservation of boreal carbon stocks and biodiversity. *Trends Ecol. Evol*. **24**, 541–548. https://doi.org/10.1016/j.tree.2009.03.019

Brown, J., Jr, O.J.F., Heginbottom, J.A., Melnikov, E.S., 1997. Circum-Arctic map of permafrost and ground-ice conditions. *Circum-Pac. Map*. https://doi.org/10.3133/cp45

Brundrett, M.C., Tedersoo, L., 2018. Evolutionary history of mycorrhizal symbioses and global host plant diversity. *New Phytol*. **220**, 1108–1115. https://doi.org/10.1111/nph.14976

Burdon, J.J., Zhan, J., 2020. Climate change and disease in plant communities. *PLOS Biol*. **18**, e3000949. https://doi.org/10.1371/journal.pbio.3000949

Champlot, S., Berthelot, C., Pruvost, M., *et al*., 2010. An efficient multistrategy DNA decontamination procedure of PCR reagents for hypersensitive PCR applications. *PLOS ONE* **5**, e13042. https://doi.org/10.1371/journal.pone.0013042

Chen, Q., Hou, L., Duan, W., *et al*., 2017. Didymellaceae revisited. *Stud. Mycol*. **87**. https://doi.org/10.1016/j.simyco.2017.06.002

Coissac, E., Riaz, T., Puillandre, N., 2012. Bioinformatic challenges for DNA metabarcoding of plants and animals. *Mol. Ecol.* **21**, 1834–1847. https://doi.org/10.1111/j.1365-294X.2012.05550.x

Compant, S., Duffy, B., Nowak, J., *et al*., 2005. Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. *Appl. Environ. Microbiol.* **71**, 4951–4959. https://doi.org/10.1128/AEM.71.9.4951-4959.2005

Cornelissen, J.H.C., Callaghan, T.V., Alatalo, J.M., *et al*., 2001. Global change and arctic ecosystems: is lichen decline a function of increases in vascular plant biomass? *J. Ecol.* **89**, 984–994. https://doi.org/10.1111/j.1365-2745.2001.00625.x

Courtin, J., Andreev, A.A., Raschke, E., *et al*., 2021. Vegetation changes in southeastern Siberia during the Late Pleistocene and the Holocene. *Front. Ecol. Evol*. **9**. https://doi.org/10.3389/fevo.2021.625096

Czimczik, C.I., Schmidt, M.W.I., Schulze, E.-D., 2005. Effects of increasing fire frequency on black carbon and organic matter in podzols of Siberian Scots pine forests. *Eur. J. Soil Sci*. **56**, 417–428. https://doi.org/10.1111/j.1365-2389.2004.00665.x

Deslippe, J.R., Hartmann, M., Simard, S.W., Mohn, W.W., 2012. Long-term warming alters the composition of Arctic soil microbial communities. *FEMS Microbiol. Ecol*. **82**, 303–315. https://doi.org/10.1111/j.1574-6941.2012.01350.x

Doležal, J., Yakubov, V., Hara, T., 2013. Plant diversity changes and succession along resource availability and disturbance gradients in Kamchatka. *Plant Ecol*. **214**, 477–488. https://doi.org/10.1007/s11258-013-0184-z

Drucker, D.G., Kind, C.-J., Stephan, E., 2011. Chronological and ecological information on Late Glacial and early Holocene reindeer from northwest Europe using radiocarbon (^{14}C) and stable isotope (^{13}C, ^{15}N) analysis of bone collagen: Case study in southwestern Germany. *Quat. Int.* **245**, 218–224. https://doi.org/10.1016/j.quaint.2011.05.007

El-Tarabily, K.A., Sivasithamparam, K., 2006. Potential of yeasts as biocontrol agents of soil-borne fungal plant pathogens and as plant growth promoters. *Mycoscience* **47**, 25–35. https://doi.org/10.1007/S10267-005-0268-2

Epp, L.S., Boessenkool, S., Bellemain, E.P., *et al*., 2012. New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Mol. Ecol*. **21**, 1821–1833. https://doi.org/10.1111/j.1365-294X.2012.05537.x

Epp, L., Gussarova, G., Boessenkool, S., *et al*., (2015): Lake sediment multi-taxon DNA from North Greenland records early post-glacial appearance of vascular plants and accurately tracks environmental changes. *Quat. Sci. Rev.* **117**, 152-163. . https://doi.org/10.1016/j.quascirev.2015.03.027

Epp, L.S., Kruse, S., Kath, N.J., *et al*., 2018. Temporal and spatial patterns of mitochondrial haplotype and species distributions in Siberian larches inferred from ancient environmental DNA and modeling. Sci. Rep. 8, 17436. https://doi.org/10.1038/s41598-018-35550-w

Fedorov, A.N., Iwahana, G., Konstantinov, P.Y., *et al*., 2017. Variability of permafrost and landscape conditions following clear cutting of larch forest in Central Yakutia. *Permafr. Periglac. Process*. **28**, 331–338. https://doi.org/10.1002/ppp.1897

Finlay, R.D., 2008. Ecological aspects of mycorrhizal symbiosis: with special emphasis on the functional diversity of interactions involving the extraradical mycelium. *J. Exp. Bot*. **59**, 1115–1126. https://doi.org/10.1093/jxb/ern059

Flannigan, M., Stocks, B., Turetsky, M., Wotton, M., 2009. Impacts of climate change on fire activity and fire management in the circumboreal forest. *Glob. Change Biol*. **15**, 549–560. https://doi.org/10.1111/j.1365-2486.2008.01660.x

Frąc, M., Hannula, S.E., Bełka, M., Jędryczka, M., 2018. Fungal Biodiversity and Their Role in Soil Health. *Front. Microbiol*. **9**. https://doi.org/10.3389/fmicb.2018.00707

Fraser, R.H., Lantz, T.C., Olthof, I., *et al.,* 2014. Warming-Induced Shrub Expansion and Lichen Decline in the Western Canadian Arctic. *Ecosystems* **17**, 1151–1168. https://doi.org/10.1007/s10021-014-9783-3

Frøslev, T.G., Kjøller, R., Bruun, H.H., *et al*., 2017. Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nat. Commun*. **8**, 1188. https://doi.org/10.1038/s41467-017-01312-x

Garbelotto, M., Gonthier, P., 2013. Biology, epidemiology, and control of *Heterobasidion* species worldwide. *Annu. Rev. Phytopathol*. **51**, 39–59. https://doi.org/10.1146/annurev-phyto-082712-102225

Garnica, S., Weiß, M., Oertel, B., Oberwinkler, F., 2005. A framework for a phylogenetic classification in the genus *Cortinarius* (Basidiomycota, Agaricales) derived from morphological and molecular data. *Can. J. Bot*. **83**, 21. https://doi.org/10.1139/b05-107

Geml, J., Morgado, L.N., Semenova, T.A., *et al*., 2015. Long-term warming alters richness and composition of taxonomic and functional groups of arctic fungi. *FEMS Microbiol. Ecol*. **91**. https://doi.org/10.1093/femsec/fiv095

Geml, J., Morgado, L.N., Semenova-Nelsen, T.A., 2021. Tundra type drives distinct trajectories of functional and taxonomic composition of Arctic fungal communities in response to climate change – results from long-term experimental summer warming and increased snow depth. *Front. Microbiol*. **12**, 628746. https://doi.org/10.3389/fmicb.2021.628746

Geml, J., Morgado, L.N., Semenova-Nelsen, T.A., Schilthuizen, M., 2017. Changes in richness and community composition of ectomycorrhizal fungi among altitudinal vegetation types on Mount Kinabalu in Borneo. *New Phytol*. **215**, 454–468. https://doi.org/10.1111/nph.14566

Goodwin, S., McPherson, J.D., McCombie, W.R., 2016. Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet*. **17**, 333–351. https://doi.org/10.1038/nrg.2016.49

Hernández-Fernández, M., Cordero-Bueso, G., Ruiz-Muñoz, M., Cantoral, J.M., 2021. Culturable yeasts as biofertilizers and biopesticides for a sustainable agriculture: a comprehensive review. Plants 10, 822. https://doi.org/10.3390/plants10050822

Hillebrand, H., Blasius, B., Borer, E.T., *et al*., 2018. Biodiversity change is uncoupled from species richness trends: consequences for conservation and monitoring. *J. Appl. Ecol*. **55**, 169–184. https://doi.org/10.1111/1365-2664.12959

Hosseini-Nasabnia, Z., Van Rees, K., Vujanovic, V., 2016. Preventing unwanted spread of invasive fungal species in willow (*Salix* spp.) plantations. *Can. J. Plant Pathol*. **38**, 325–337. https://doi.org/10.1080/07060661.2016.1228697

Huang, S., Stoof-Leichsenring, K.R., Liu, S., *et al*., 2020. Plant sedimentary ancient DNA from Far East Russia covering the last 28 ka reveals different assembly rules in cold and warm climates. *bioRxiv* 2020.12.11.406108. https://doi.org/10.1101/2020.12.11.406108

Ishida, T.A., Nara, K., Hogetsu, T., 2007. Host effects on ectomycorrhizal fungal communities: insight from eight host species in mixed conifer–broadleaf forests. *New Phytol*. **174**, 430–440. https://doi.org/10.1111/j.1469-8137.2007.02016.x

Iznova, T., Rukšėnienė, J., 2012. Ascomycete species new to Lithuania. *Bot. Lith*. **18**, 35–39. https://doi.org/10.2478/v10279-012-0005-7

Jarvis, P., Linder, S., 2000. Constraints to growth of boreal forests. *Nature* **405**, 904–905. https://doi.org/10.1038/35016154

Kachalkin, A.V., Yurkov, A.M., 2012. Yeast communities in *Sphagnum* phyllosphere along the temperature-moisture ecocline in the boreal forest-swamp ecosystem and description of *Candida sphagnicola* sp. nov. *Antonie Van Leeuwenhoek* **102**, 29–43. https://doi.org/10.1007/s10482-012-9710-6

Kandror, O., Bretschneider, N., Kreydin, E., *et al*., 2004. Yeast adapt to near-freezing temperatures by STRE/Msn2,4-dependent induction of trehalose synthesis and certain molecular chaperones. *Mol. Cell* **13**, 771–781. https://doi.org/10.1016/S1097-2765(04)00148-0

Kanz, C., Aldebert, P., Althorpe, N., *et al*., 2005. The EMBL nucleotide sequence database. *Nucleic Acids Res.* **33**, D29–D33. https://doi.org/10.1093/nar/gki098

Kerkhoff, A.J., Moriarty, P.E., Weiser, M.D., 2014. The latitudinal species richness gradient in New World woody angiosperms is consistent with the tropical conservatism hypothesis. *Proc. Natl. Acad. Sci*. **111**, 8125–8130. https://doi.org/10.1073/pnas.1308932111

Kharuk, V., Ranson, K.J., Dvinskaya, M., 2007. Evidence of evergreen conifer invasion into larch dominated forests during recent decades in central Siberia. *Eurasian J. For. Res*. **10**, 163-171.

Kharuk, V., Ranson, K.J., Im, S., Dvinskaya, M., 2009. Response of *Pinus sibirica* and *Larix sibirica* to climate change in southern Siberian alpine forest-tundra ecotone. *Scand. J. For. Res*. **24**, 130–139. https://doi.org/10.1080/02827580902845823

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. *Quat. Sci. Rev.* **147**, 422–434. https://doi.org/10.1016/j.quascirev.2015.08.015

Konstantinov, A.F., 2000. Environmental problems of Lake Bolshoe Toko. *Lakes of Cold Environments*, part V: Resource Study 85–93.

Kreveld, S. van, Sarnthein, M., Erlenkeuser, H., *et al*., 2000. Potential links between surging ice sheets, circulation changes, and the Dansgaard-Oeschger Cycles in the Irminger Sea, 60–18 Kyr. *Paleoceanography* **15**, 425–442. https://doi.org/10.1029/1999PA000464

Kreyling, J., Haei, M., Laudon, H., 2012. Absence of snow cover reduces understory plant cover and alters plant community composition in boreal forests. *Oecologia* **168**, 577–587. https://doi.org/10.1007/s00442-011-2092-z

Kurek, E., Korniłłowicz−kowalska, T., Słomka, A., Melke, J., 2007. Characteristics of soil filamentous fungi communities isolated from various micro-relief forms in the high Arctic tundra (Bellsund region, Spitsbergen). *Polar Res* **28**, 57–73.

Lebas, E., Krastel, S., Wagner, B., *et al*., 2019. Seismic stratigraphical record of lake Levinson-Lessing, Taymyr peninsula: evidence for ice-sheet dynamics and lake-level fluctuations since the Early Weichselian. *Boreas* **48**, 470–487. https://doi.org/10.1111/bor.12381

Leski, T., Rudawska, M., 2012. Ectomycorrhizal fungal community of naturally regenerated European larch (*Larix decidua*) seedlings. *Symbiosis* **56**, 45–53. https://doi.org/10.1007/s13199-012-0164-4

Liu, S., Stoof-Leichsenring, K. R., Kruse, S., *et al*., 2020. Holocene vegetation and plant diversity changes in the north-eastern Siberian treeline region from pollen and sedimentary ancient DNA. *Front. Ecol. Evol.* **8**, 18. https://doi.org/10.3389/fevo.2020.560243

Loughlin, N., Gosling, W., & Montoya, E. (2018). Identifying environmental drivers of fungal non-pollen palynomorphs in the montane forest of the eastern Andean flank, Ecuador. *Quaternary Research* **89**(1), 119-133. http.d://doi.org/10.1017/qua.2017.73

Lydolph, M.C., Jacobsen, J., Arctander, P., *et al*., 2005. Beringian paleoecology inferred from permafrost-preserved fungal DNA. *Appl. Environ. Microbiol*. **71**, 1012–1017. https://doi.org/10.1128/AEM.71.2.1012-1017.2005

MacDonald, G.M., Kremenetski, K.V., Beilman, D.W., 2008. Climate change and the northern Russian treeline zone. *Philos. Trans. R. Soc. B Biol. Sci*. **363**, 2285–2299. https://doi.org/10.1098/rstb.2007.2200

Marschner, H., Dell, B., 1994. Nutrient uptake in mycorrhizal symbiosis. *Plant Soil* **159**, 89–102. https://doi.org/10.1007/BF00000098

Matheny, P.B., Hobbs, A.M., Esteve-Raventós, F., 2020. Genera of Inocybaceae: New skin for the old ceremony. *Mycologia* **112**, 83–120. https://doi.org/10.1080/00275514.2019.1668906

McCalley, C.K., Woodcroft, B.J., Hodgkins, S.B., *et al*., 2014. Methane dynamics regulated by microbial community response to permafrost thaw. *Nature* **514**, 478–481. https://doi.org/10.1038/nature13798

McGuire, A.D., Anderson, L.G., Christensen, T.R., *et al.,* 2009. Sensitivity of the carbon cycle in the Arctic to climate change. *Ecological Monographs* **79**, 523-555. https://doi.org/10.1890/08-2025.1

McGuire, K.L., Allison, S.D., Fierer, N., Treseder, K.K., 2013. Ectomycorrhizal-dominated boreal and tropical forests have distinct fungal communities, but analogous spatial patterns across soil horizons. *PLOS ONE* **8**, e68278. https://doi.org/10.1371/journal.pone.0068278

Mercier, C., Boyer, F., Bonin, A., Coissac, E., 2013. SUMATRA and SUMACLUST: fast and exact comparison and clustering of sequences. [https://bioweb.pasteur.fr/docs/modules/suma_package/v1.0.00/sumatra_sumaclust_user_manual.pdf].

Merges, D., Bálint, M., Schmitt, I., *et al*., 2018. Spatial patterns of pathogenic and mutualistic fungi across the elevational range of a host plant. *J. Ecol.* **106**, 1545–1557. https://doi.org/10.1111/1365-2745.12942

Miyamoto, Y., Danilov, A.V., Bryanin, S.V., 2021. The dominance of *Suillus* species in ectomycorrhizal fungal communities on *Larix gmelinii* in a post-fire forest in the Russian Far East. *Mycorrhiza* **31**, 55–66. https://doi.org/10.1007/s00572-020-00995-3

Mohamed, D.J., Martiny, J.B., 2011. Patterns of fungal diversity and composition along a salinity gradient. *ISME J*. **5**, 379–388. https://doi.org/10.1038/ismej.2010.137

Mundra, S., Halvorsen, R., Kauserud, H., *et al*., 2016. Ectomycorrhizal and saprotrophic fungi respond differently to long-term experimentally increased snow depth in the High Arctic. *Microbiology Open* **5**, 856–869. https://doi.org/10.1002/mbo3.375

Myers-Smith, I., Elmendorf, S., Beck, P. et al, 2015. Climate sensitivity of shrub growth across the tundra biome. *Nature Clim Change*, **5,** 887–891. https://doi.org/10.1038/nclimate2697

Naranjo-Ortiz, M.A., Gabaldón, T., 2019. Fungal evolution: major ecological adaptations and evolutionary transitions. *Biol. Rev*. **94**, 1443–1476. https://doi.org/10.1111/brv.12510

Nash, T.H., 2002. Lichen flora of the greater Sonoran Desert Region: Most of the microlichens, balance of the macrolichens, and lichenicolous fungi. *Lichens Unlimited*, Arizona State University.

Nassar, A.H., El-Tarabily, K.A., Sivasithamparam, K., 2005. Promotion of plant growth by an auxin-producing isolate of the yeast *Williopsis saturnus* endophytic in maize (*Zea mays* L.) roots. *Biol. Fertil. Soils* **42**, 97–108. https://doi.org/10.1007/s00374-005-0008-y

Nelsen, M.P., Gargas, A., 2009. Symbiont flexibility in *Thamnolia vermicularis* (Pertusariales: Icmadophilaceae). *The Bryologist* **112**, 404–417.

Niemeyer, B., Epp, L.S., Stoof-Leichsenring, K.R., *et al*., 2017. A comparison of sedimentary DNA and pollen from lake sediments in recording vegetation composition at the Siberian treeline. *Mol. Ecol. Resour*. **17**, e46–e62. https://doi.org/10.1111/1755-0998.12689

Nilsson, R.H., Kristiansson, E., Ryberg, M., Larsson, K.-H., 2005. Approaching the taxonomic affiliation of unidentified sequences in public databases – an example from the mycorrhizal fungi. *BMC Bioinformatics* **6**, 178. https://doi.org/10.1186/1471-2105-6-178

Nilsson, R.H., Larsson, K.-H., Taylor, A.F.S., *et al*., 2019. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res*. **47**, D259–D264. https://doi.org/10.1093/nar/gky1022

Otrosina, W.J., Cobb, F.W., 1989. Biology, Ecology, and Epidemiology of Heterobasidion annosum. In: Otrosina, William J.; Scharpf, Robert F., technical coordinators. 1989. Proceedings of the symposium on research and management of annosus root disease (*Heterobasidion annosum*) in western North America; April 18-21, 1989; Monterey, CA. Gen. Tech. Rep. PSW-GTR-116. Berkeley, CA: Pacific Southwest Forest and Range Experiment Station, Forest Service, U.S. Department of Agriculture; p. 26-34.

Palamarchuk, M., Kirillov, D., 2019. Fungi (Agaricoid Basidiomycetes) of the Pechoro-Ilych reserve (Komi Republic, Russia). https://doi.org/10.15468/o9jk3m

Pande, V., Palni, U.T., Singh, S.P., 2004. Species diversity of ectomycorrhizal fungi associated with temperate forest of Western Himalaya: a preliminary assessment. *Curr. Sci*. **86**, 1619–1623.

Parducci, L., Bennett, K.D., Ficetola, G.F., *et al*., 2017. Ancient plant DNA in lake sediments. *New Phytol.* **214**, 924–942. https://doi.org/10.1111/nph.14470

Põlme, S., Bahram, M., Jacquemyn, H., *et al*., 2018. Host preference and network properties in biotrophic plant–fungal associations. *New Phytol*. **217**, 1230–1239. https://doi.org/10.1111/nph.14895

Polyakova, A., Chernov, I., 2001. Yeast diversity in hydromorphic soils with reference to a grass–*Sphagnum* wetland in western Siberia and a hummocky tundra region at Cape Barrow (Alaska). *Microbiology* **70**, 617–623. https://doi.org/10.1023/A:1012328710111

Porada, P., Ekici, A., Beer, C., 2016. Effects of bryophyte and lichen cover on permafrost soil temperature at large scale. *The Cryosphere* **10**, 2291–2315. https://doi.org/10.5194/tc-10-2291-2016

Quamar, Md. F., Stivrinsz, M. 2021. Modern pollen and non-pollen palynomorphs along an altitudinal transect in Jammu and Kashmir (Western Himalaya), India. *Palynology* **45** (4), 669–684. https://doi.org/10.1080/01916122.2021.1915402

Quince, C., Lanzén, A., Curtis, T.P., *et al*., 2009. Accurate determination of microbial diversity from 454 pyrosequencing data. *Nat. Methods* **6**, 639–641. https://doi.org/10.1038/nmeth.1361

Rasmussen, M., Li, Y., Lindgreen, S. *et al.* Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* **463,** 757–762 (2010). https://doi.org/10.1038/nature08835

R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reimer, P.J., Austin, W.E.N., Bard, E., *et al*., 2020. The IntCal20 northern hemisphere radiocarbon age calibration curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757. https://doi.org/10.1017/RDC.2020.41

Renker, C., Alphei, J., Buscot, F., 2003. Soil nematodes associated with the mammal pathogenic fungal genus *Malassezia* (Basidiomycota: Ustilaginomycetes) in central European forests. *Biol. Fertil. Soils* **37**, 70–72. https://doi.org/10.1007/s00374-002-0556-3

Romero-Olivares, A.L., Meléndrez-Carballo, G., Lago-Lestón, A., Treseder, K.K., 2019. Soil metatranscriptomes under long-term experimental warming and drying: fungi allocate resources to cell metabolic maintenance rather than decay. *Front. Microbiol.* **10**. https://doi.org/10.3389/fmicb.2019.01914

Romero-Olivares, A.L., Taylor, J.W., Treseder, K.K., 2015. *Neurospora discreta* as a model to assess adaptation of soil fungi to warming. *BMC Evol. Biol*. **15**, 198. https://doi.org/10.1186/s12862-015-0482-2

Russian Institute of Hydrometeorological Information: World Data Center, available at http://meteo.ru/english/climate/temp.php (last access: 18. Juni 2021), 2021.

Ryberg, M., Andreasen, M., Björk, R.G., 2011. Weak habitat specificity in ectomycorrhizal communities associated with *Salix herbacea* and *Salix polaris* in alpine tundra. *Mycorrhiza* **21**, 289–296. https://doi.org/10.1007/s00572-010-0335-1

Ryberg, M., Larsson, E., Molau, U., 2009. Ectomycorrhizal diversity on *Dryas octopetala* and *Salix reticulata* in an alpine cliff ecosystem. *Arct. Antarct. Alp. Res*. **41**, 506–514. https://doi.org/10.1657/1938-4246-41.4.506

Safronova, I., Yurkovsksya, T., 2019. The latitudinal distribution of vegetation cover in Siberia. *BIO Web Conf*. **16**, 00047. https://doi.org/10.1051/bioconf/20191600047

Santos, A., Sánchez, A., Marquina, D., 2004. Yeasts as biological agents to control *Botrytis cinerea*. *Microbiol. Res*. **159**, 331–338. https://doi.org/10.1016/j.micres.2004.07.001

Scheidt, S., Egli, R., Lenz, M., *et al*., 2021. Mineral magnetic characterization of high-latitude sediments from lake Levinson-Lessing, Siberia. *Geophys. Res. Lett*. **48**, e2021GL093026. https://doi.org/10.1029/2021GL093026

Schiro, G., Colangeli, P., Müller, M.E.H., 2019. A metabarcoding analysis of the mycobiome of wheat ears across a topographically heterogeneous field. *Front. Microbiol.* **10**. https://doi.org/10.3389/fmicb.2019.02095

Schulte, L., Bernhardt, N., Stoof-Leichsenring, K., *et al*., 2021. Hybridization capture of larch (*Larix* Mill.) chloroplast genomes from sedimentary ancient DNA reveals past changes of Siberian forest. *Mol. Ecol. Resour*. **21**, 801–815. https://doi.org/10.1111/1755-0998.13311

Schulze, E.-D., Mooney, H.A., 2012. Biodiversity and Ecosystem Function. *Springer* Science & Business Media.

Seeber, P.A., von Hippel, B., Kauserud, H., *et al*., 2022. Evaluation of lake sedimentary ancient DNA metabarcoding to assess fungal biodiversity in Arctic paleoecosystems. *Environmental DNA*, https://doi.org/10.1002/edn3.315

Seifert, K.A., 2009. Progress towards DNA barcoding of fungi. *Mol. Ecol. Resour*. **9**, 83–89. https://doi.org/10.1111/j.1755-0998.2009.02635.x

Sheard, J.W., 1977. Paleogeography, Chemistry and Taxonomy of the Lichenized Ascomycetes *Dimelaena* and *Thamnolia*. *The Bryologist* **80**, 100–118. https://doi.org/10.2307/3242516

Shevtsova, I., Heim, B., Kruse, S., *et al*., 2020. Strong shrub expansion in tundra-taiga, tree infilling in taiga and stable tundra in central Chukotka (north-eastern Siberia) between 2000 and 2017. *Environ. Res. Lett*. **15**, 085006. https://doi.org/10.1088/1748-9326/ab9059

Slaght, J.C., Milakovsky, B., Maksimova, D.A., *et al*., 2019. Anthropogenic influences on the distribution of a vulnerable coniferous forest specialist: habitat selection by the Siberian musk deer *Moschus moschiferus*. *Oryx* **53**, 174–180. https://doi.org/10.1017/S0030605316001617

Soininen, E.M., Gauthier, G., Bilodeau, F., *et al*., 2015. Highly overlapping winter diet in two sympatric lemming species revealed by DNA metabarcoding. *PLOS One* **10**, e0115335. https://doi.org/10.1371/journal.pone.0115335

Sønstebø, J.H., Gielly, L., Brysting, A.K., *et al*., 2010. Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate. *Mol. Ecol. Resour*. **10**, 1009–1018. https://doi.org/10.1111/j.1755-0998.2010.02855.x

Stoof-Leichsenring, K., Herzschuh, U., Pestrayakova, L., *et al*., 2015. Genetic data from algae sedimentary DNA reflect the influence of environment over geography. *Sci. Rep*. **5**. https://doi.org/10.1038/srep12924

Svendsen, J.I., Alexanderson, H., Astakhov, V.I., *et al*., 2004. Late Quaternary ice sheet history of northern Eurasia. *Quat. Sci. Rev*. **23**, 1229–1271. https://doi.org/10.1016/j.quascirev.2003.12.008

Swann, G.E.A., Mackay, A.W., Leng, M.J., Demory, F., 2005. Climatic change in central Asia during MIS 3/2: a case study using biological responses from lake Baikal. *Glob. Planet. Change* **46**, 235–253. https://doi.org/10.1016/j.gloplacha.2004.09.019

Taberlet, P., Coissac, E., Pompanon, F., *et al*., 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol. Ecol*. **21**, 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x

Taberlet, P., Coissac, E., Pompanon, F., *et al*., 2007. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Res*. **35**, e14. https://doi.org/10.1093/nar/gkl938

Talas, L., Stivrins, N., Veski, S., *et al*., 2021. Sedimentary ancient DNA (sedaDNA) reveals fungal Diversity and environmental drivers of community changes throughout the Holocene in the present boreal lake Lielais Svētiņu (Eastern Latvia). *Microorganisms* **9**, 719. https://doi.org/10.3390/microorganisms9040719

Tarnocai, C., Canadell, J. G., Schuur, E. A. G., *et al*., 2009. Soil organic carbon pools in the northern circumpolar permafrost region. Global Biogeochem. *Cycles* **23**, GB2023. https://doi.org/10.1029/2008GB003327.

Taylor, T., Osborn, J., 1996. The importance of fungi in shaping the paleoecosystem. https://doi.org/10.1016/0034-6667(95)00086-0

Tchebakova, N.M., Parfenova, E., Soja, A.J., 2009. The effects of climate, permafrost and fire on vegetation change in Siberia in a changing climate. *Environ. Res. Lett*. **4**, 045013. https://doi.org/10.1088/1748-9326/4/4/045013

Tedersoo, L., Mett, M., Ishida, T.A., Bahram, M., 2013. Phylogenetic relationships among host plants explain differences in fungal species richness and community composition in ectomycorrhizal symbiosis. *New Phytol*. **199**, 822–831. https://doi.org/10.1111/nph.12328

Thormann, M.N., 2006. Diversity and function of fungi in peatlands: a carbon cycling perspective. *Can. J. Soil Sci.* **86**, 281–293. https://doi.org/10.4141/S05-082

Timling, I., Dahlberg, A., Walker, D.A., *et al*., 2012. Distribution and drivers of ectomycorrhizal fungal communities across the North American Arctic. *Ecosphere* **3**, art111. https://doi.org/10.1890/ES12-00217.1

Treseder, K.K., Marusenko, Y., Romero-Olivares, A.L., Maltz, M.R., 2016. Experimental warming alters potential function of the fungal community in boreal forest. *Glob. Change Biol*. **22**, 3395–3404. https://doi.org/10.1111/gcb.13238

Treseder, K.K., Turner, K.M., Mack, M.C., 2007. Mycorrhizal responses to nitrogen fertilization in boreal ecosystems: potential consequences for soil carbon storage. *Glob. Change Biol*. **13**, 78–88. https://doi.org/10.1111/j.1365-2486.2006.01279.x

Urbanová, M., Šnajdr, J., Baldrian, P., 2015. Composition of fungal and bacterial communities in forest litter and soil is largely determined by dominant trees. *Soil Biol. Biochem*. **84**, 53–64. https://doi.org/10.1016/j.soilbio.2015.02.011

Van Geel, B., 2001. Non-Pollen Palynomorphs, in: Smol, J.P., Birks, H.J.B., Last, W.M., Bradley, R.S., Alverson, K. (Eds.), Tracking Environmental Change Using Lake Sediments, Vol. 3: Terrestrial, Algal, and Siliceous Indicators. *Springer* Netherlands, Dordrecht, 99–119. https://doi.org/10.1007/0-306-47668-1_6

Voříšková, J., Elberling, B., Priemé, A., 2019. Fast response of fungal and prokaryotic communities to climate change manipulation in two contrasting tundra soils. *Environ. Microbiome* **14**, 6. https://doi.org/10.1186/s40793-019-0344-4

Wallenstein, M.D., McMahon, S., Schimel, J., 2007. Bacterial and fungal community structure in Arctic tundra tussock and shrub soils: community structure in Arctic tundra tussock and shrub soils. *FEMS Microbiol. Ecol*. **59**, 428–435. https://doi.org/10.1111/j.1574-6941.2006.00260.x

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis, 2nd ed, Use R! *Springer* International Publishing. https://doi.org/10.1007/978-3-319-24277-4

Willerslev, E., Davison, J., Moora, M., *et al*., 2014. Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature* **506**, 47–51. https://doi.org/10.1038/nature12921

Yang, T., Adams, J.M., Shi, Y., *et al*., 2017. Soil fungal diversity in natural grasslands of the Tibetan Plateau: associations with plant diversity and productivity. *New Phytol*. **215**, 756–765. https://doi.org/10.1111/nph.14606

Yurkov, A.M., Kemler, M., Begerow, D., 2012. Assessment of yeast diversity in soils under different management regimes. *Fungal Ecol*. **5**, 24–35. https://doi.org/10.1016/j.funeco.2011.07.004

Zhang, K., Shi, Y., Jing, X., *et al*., 2016. Effects of short-term warming and altered precipitation on soil microbial communities in alpine grassland of the Tibetan plateau. *Front. Microbiol*. **7**, 1032. https://doi.org/10.3389/fmicb.2016.01032

Zhang, N., Yasunari, T., Ohta, T., 2011. Dynamics of the larch taiga–permafrost coupled system in Siberia under climate change. *Environ. Res. Lett*. **6**, 024003. https://doi.org/10.1088/1748-9326/6/2/024003

Zhurbenko, M.P., Yakovchenko, L.S., 2014. A new species, *Sagediopsis vasilyevae*, and other lichenicolous fungi from Zabaikal'skii Territory of Russia, southern Siberia. *Folia Cryptogam. Est*. **51**, 121–130. https://doi.org/10.12697/fce.2014.51.14

Zobel, M., Davison, J., Edwards, M.E., *et al*., 2018. Ancient environmental DNA reveals shifts in dominant mutualisms during the late Quaternary. *Nat. Commun*. **9**, 139. https://doi.org/10.1038/s41467-017-02421-3

4 Manuscript III

Postglacial bioweathering, soil nutrient cycling, and podzolization from palaeometagenomics of plants, fungi, and bacteria

Status

Submitted to *Nature Communications* (date: 06-13-2023)

Authors

Barbara von Hippel[1], Kathleen Stoof-Leichsenring[1], Martin Melles[2], Ulrike Herzschuh[1,3,4*]

Affiliations

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

2 Institute of Geology and Mineralogy, University of Cologne, Germany

3 Institute of Environmental Science and Geography, University of Potsdam, Germany

4 Institute of Biochemistry and Biology, University of Potsdam, Germany

*corresponding author (Ulrike.Herzschuh@awi.de)

Keywords

carbon capture, ecosystem dynamics, paleo metagenomics, sedimentary ancient DNA, soil development

4.1 Abstract

Past and recent warming-induced glacier retreat exposes bare rocks, facilitating the establishment of soils. The dynamic interplay between climate, vegetation cover, and soil formation, however, is poorly understood as time-series covering an appropriate time span are lacking. Here, we present post-glacial soil formation during the past 23,000 years as inferred from ancient DNA shotgun analyses of lake Lama sediments (northern-central Siberia) targeting plants, rhizosphere-associated fungi, and bacteria. In the Late Glacial, we reveal strong basaltic weathering with lichen-domination and high relative abundances of arsenic cyclers, shifting to mycorrhizae-domination in the Holocene. Additionally, the bacterial element cycling shifts from C to N dominance and a diversification of nitrogen pathways is detected. Further, we reconstruct podzolization starting with Holocene spruce-forest invasion showing soil acidification and, later, increased iron cycling. Our results show pedogenesis is an environmentally driven process, mainly by vegetation, although differences in early vs late Holocene larch forest soil communities suggest trajectory effects due to soil ageing. As well as basic knowledge on postglacial soil formation, our data provide a scientific knowledge base for the design of carbon-capture strategies using basalt weathering.

4.2 Introduction

Soils often feature as a static entity in terrestrial ecosystems, for example in dynamic global vegetation models, despite it being known that they develop and even dynamically respond to drivers[1–3]. This misconception originates, at least partly, from the lack of time-series portraying the major soil processes including weathering[4], element cycling, and podzolization[5] and their reflection in soil communities (mainly plants, fungi, and micro-organisms). For example, the initial soil establishment after deglaciation at the end of the last glacial and the subsequent soil development in response to climate-driven vegetation changes remains largely unexplored. However, understanding soil changes and their related drivers is necessary for decision-making to safeguard ecosystem services of soils including food production, forestry, and maintenance of ecosystem stability.

Pedogenesis is initialised by weathering of the parent material which is, amongst other processes, supported by plants, fungi, and bacteria[6–8]. Lichens, as characteristic early colonisers, enhance weathering by using their hyphae to penetrate mineral cleavage planes[9] as well as releasing organic acids[6,10,11]. Plant root exudates, for example low molecular weight organic acids deriving from respiratory CO_2[12], additionally increase weathering. In more developed soils, ectomycorrhizae further enhance weathering when supplying plants with ammonium, resulting in an efflux of H^+ and,

subsequently, soil acidification[13]. However, how basalt weathering changes on millennial time-scales in relation to compositional changes of plants, fungi, and bacteria remains largely unexplored .

Nutrient cycling by fungi and bacteria in the rhizosphere, particularly of carbon, nitrogen, phosphorus, and sulphur, determines plant productivity, diversity, and composition[14]. Most soil organic carbon originates from above- and below-ground plant litter degradation and transformation[15]. In addition to the plant-produced organic matter, soil organic carbon can be derived from atmospheric CO_2 being fixed by multiple photo- and chemoautotrophic microbes in the soil[16], while heterotrophic bacteria degrade these fixed carbon compounds, later using them as a metabolic substrate, and releasing smaller parts as metabolites or as CO_2 back into the atmosphere[17]. Besides soil bacteria, saprotrophic fungi are also important for a first degradation of complex carbon compounds such as lignin[18]. In the nitrogen cycle, N-fixing bacteria directly bind atmospheric nitrogen and convert it into a plant-available form[19]. The plant uptake of nitrogen from the soil is supported by mycorrhizal fungi[20–22]. Vegetation densification, such as forest establishment, results in an increased need for nutrient supply due to reduced turnover times of wood compared to soft tissue[23]. Whether nutrient cycling is also changing on long time-scales alongside soil development but with similar vegetation cover remains unknown. Also, whether a more complex nutrient demand in relation to vegetation densification results in a long-term diversification of the cycling pathways is currently unknown.

Podzols are the common soil type in boreal forests which are typically dominated by *Larix*, *Picea*, or *Pinus*[24]. These soils are characterised by low pH and show a high sensitivity towards further acidification due to low capacities for cation exchange and small amount of weatherable material[25]. During podzolization, organic acids induce the release of aluminium and iron ions from rocks which then form chelates with organic matter[24,26]. These complexes leach from the upper mineral horizons (bleaching) and become - at least partly - deposited in the subsoil leading to its characteristic reddish-brown colour[24,26]. So far, podzolization has mainly been described along spatial gradients, and such studies do not help our understanding of podzolization temporalities. However, *Larix-Rhododendron* succession, for example, has been found to accelerate podzolization[27], although the specific and unique impact of ecological processes and environmental drivers are poorly understood. When podzolization started in the boreal forest and whether vegetation compositional changes can reverse podzolization processes remain largely unknown.

Directly assessing soil dynamics would greatly improve existing knowledge on soil development. Through erosion, soil-derived matter, including substantial amounts of DNA, can be transported into a lake[28]. Consequently, the analysis of lake sedimentary ancient DNA (sedaDNA) has become a popular

palaeoecological method[29]. Hitherto, mostly metabarcoding approaches are applied to target single organism groups including plants[30–32], or, rarely, fungi[33,34]. Recently, metagenomic approaches emerged, enabling the study of complex ecosystems, for example through sequencing the whole DNA contained in a sample[35–37]. Such studies became possible because genome reference databases such as the widely used nucleotide (nt) database from NCBI (ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/nt.gz) have been markedly improved and extended recently. Despite the recent methodological improvements, palaeo-metagenomic studies targeting soil ecosystem development are hitherto entirely lacking.

Here we show how postglacial soils became established and further developed in response to climate-driven vegetation change during the last about 23,000 years, using sedaDNA records of plants, and rhizosphere-related bacteria and fungal taxa from lake Lama in north-central Siberia. We show that the vegetation as well as temperature variation has an impact on the establishment of the soil microbiome, while time itself is less important. We also trace the weathering progress of the basaltic bedrock in the lake catchment which shifted from a strong, lichen-dominated weathering during the Late Glacial to a generally weaker, mycorrhizae-dominated weathering in the Holocene. We also detected a turnover from carbon-dominated nutrient cycling during the Late Glacial to nitrogen-dominated cycling in the Holocene. Additionally, we reconstruct podzolization by showing increases in acidic-pH preferring taxa in all the assessed subsets as well as rising iron cycling with mid-Holocene spruce forest expansion.

4.3 Results and Discussion

4.3.1 Compositional changes of plants, fungi, and bacteria in ancient metagenomic datasets

Shotgun sequencing of 44 sediment samples, 5 extraction blanks, and 8 library blanks and subsequent read filtering yielded 1,240,658,912 reads. Of which, a total of 4,935,634 reads could be assigned to 5,967 unique terrestrial plant, fungal, and bacterial taxa in the nt database with an identity threshold of 0.8. A total of 257,710 reads belong to the Viridiplantae and are assigned to genus or species level; 3,988,224 reads belong to bacteria and are assigned to genus or species level; and 43,040 reads belong to fungi at all taxonomic levels. The read length distribution as well as the characteristic C-to-T substitution at the 5'-ends confirm the ancient origin of the reads (see supplement 5-7).

There are a median of 3,533 plant reads per sample, with 133 plant taxa identified. Among them, 25.2% of the reads are assigned to species level, while 74.8% are assigned to genus level. The overall compositional vegetation change from tundra-dominated Late Glacial to taiga-dominated Holocene reproduces the general trend known from a pollen[38] and a metabarcoding[34] record from the same lake,

and from pollen and aDNA data from other sites in northern Siberia (e.g. [31,39,40]). The Late Glacial is characterised by a typical glacial flora with high abundances of *Dryas* (avens) and Saxifragaceae as well as *Salix* (willow) in the river valleys (Figs 2, 4). After about 14 ka, *Betula* (birch) expands and, with the onset of the Holocene, *Alnus* (alder) and Pinaceae markedly increase in the record. For the early Holocene, the data show a massive expansion of *Larix* (larch), followed by *Picea* (spruce) during the mid-Holocene and a re-advance of *Larix* during the late Holocene (Figs 2, 5). With the invasion of *Picea*, the herbal community changes: Galegeae is absent from thereon and also *Dryas* (avens) also decreases in abundance, while Ericaceae, including *Pyrola rotundifolia* (round-leaved wintergreen), and Asteraceae increase in association with the *Picea* forests.

There are 318 unique fungal assignments, with a median of 225 reads per sample. Among the fungal reads, 24 % are assigned to species level, 31.1 % to genus level, and 32.44 % to family level. The remaining 12.46 % are at class level or above. We find a general trend from an Ascomycota-dominated Late Glacial with a relatively high abundance and diversity of lichens, to a Basidiomycota-dominated Holocene (many of them known as mycorrhiza-forming taxa), matching the spatial gradient observed in glacier forefields[41]. Modern studies on tundra and taiga soils from the Kola Peninsula show a *Penicillium* dominance in both biomes[42], which is contrary to our findings: Saprotrophs show a shift from *Penicillium* dominance during the Late Glacial towards *Mortierella* species in the late Holocene in our record (Figs 2, 4, 5). The overall high abundance of yeast taxa in the Late Glacial (*Malassezia* spp., *Komagataella* spp.) are in accordance with metabarcoding data[34]. Similarly to yeast, lichens (Peltigerales, *Peltigera* spp.) are more abundant in the Late Glacial than in the Holocene. Our data indicate that mycorrhizal taxa (Suillineae, Glomeraceae, *Rhizophagus*, *Laccaria*, *Hyaloscypha*, Tuberaceae) gained in importance with warming at the onset of the Holocene.

We recovered 1,251 bacterial assignments at genus and species level with a median of 24,285 reads per sample. We restricted the analyses to soil taxa (see Methods). Among them, we recovered 67.5 % reads at species level, and 32.5 % at genus level. Like with the vegetation and fungi, the major compositional shift for bacteria occurred at the Late Glacial-Holocene transition (Figs 2, 4, 5). We discovered *Brevundimonas* and *Hydrogenophaga* (both carbon-cycling genera) mainly in the Late Glacial. Additionally, arsenic cyclers from the genus *Herminiimonas*, which oxidise arsenite, are highly abundant throughout the Late Glacial[43]. In contrast, *Bradyrhizobium* (nitrogen fixation), *Ferrigenium* (iron oxidation), *Sideroxydans* (iron oxidation), and *Pseudolabrys* (ammonia oxidation) show high abundance in the Holocene samples.

4.3.2 Long-term soil development: a trajectory or environmentally driven processes?

Our time-series data on soil fungi and bacterial community changes trace, for the first time, long-term post-glacial soil development which has hitherto only been investigated along spatial gradients (e.g. [41,44,45]). Previous aDNA shotgun studies have focused on changes in above-ground terrestrial ecosystems (e.g. [30,35–37,46]), lakes[47] or oceans[48].

From variation partitioning using constrained ordination (Fig. 3), vegetation explains the highest unique amount of variance in the fungal compositional data, followed by temperature (see Methods), while time passed since deglaciation uniquely explains only a minor variation in the dataset. Similarly, vegetation and temperature uniquely explain a relatively high amount of variation in the bacterial dataset.

Generally, our findings on the importance of vegetation and temperature on soil development confirm spatial and experimental studies. For example, the invasion of *Betula nana* (dwarf birch) in arctic tundra has been identified as a main driver of soil microbial shifts after experimental warming[49]. Also, *Alnus* (alder) has been found to impact the establishment of bacterial communities after glacier retreat[50]. In tundra communities from the Taymyr Peninsula, vegetation cover also highly impacts the composition and biomass of fungi and bacteria[51]. Interestingly, we find that temperature has a greater impact on the bacterial community than on fungal composition, which is in contrast to experimental evidence from a pine forest[52].

Overall, our results indicate that post-glacial soil development on a millennial time-scale represents environmentally-driven processes rather than a pure trajectory, that is, time passed since glacial retreat explains only a small unique variance in the bacterial and fungal compositional changes. This agrees with the conclusion of Delgado-Baquerizo et al.[3] who compared multiple topsoils worldwide of varying ages, and showed that parent material, climate, vegetation, and topography have a much greater impact on soil development than soil age has. However, our results disagree with the finding that time since recent deglaciation is most important for soil microbiome establishment[53]. The majority of the variance in our fungal and bacterial data is not explained at all and most of the variance is explained by a combination of tested variables indicating that we may have missed major drivers and/or internal dynamics (e.g. external nutrient supply[54], variation in wetness[55]) and that soil-temporal-environmental relationships are complex.

4.3.3 Bioweathering supported by lichens and mycorrhiza

Basalt, forming the bedrock in the lake Lama catchment, is largely composed of feldspar silicates containing high amounts of potassium, which is a mobile element released through weathering[56,57].

Thus we interpret a high ratio of mobile K to immobile Ti (Fig. 6) in the sediment as a proxy for strong soil weathering[58–60].

We reveal enhanced weathering during the initial phase of soil formation shortly after deglaciation. According to the fungal record, this is at least partially related to lichens (and their release of organic acids) which are known to be early colonisers of basaltic rocks[9,61,62]. A second phase of high weathering occurred during the phase of maximum *Salix* abundance when glacial meltwater percolated the soils during the Bølling/Allerød warm period (Figs 2, 6).

Strong basalt weathering in the Late Glacial is also confirmed by the high abundance of arsenite oxidizers (mainly *Herminiimonas arsenicoxydans*, *H. arsenitoxidans*; Figs 2, 6) in our bacterial record. It is likely that the weathering of basalt led to the release of iron hydroxides[63] which have a high affinity to bind to arsenic[64]. The arsenic-cycling taxa oxidise the arsenic anion to arsenate[43,65], a process that is highly needed after the high arsenic release from initial rock weathering.

Interestingly, weathering declined with *Larix* forest expansion after the onset of the Holocene, alongside a decline in lichen and an increase in mycorrhiza relative abundance (Figs 2, 6). Previous short-term studies have demonstrated a decline in lichen abundance and diversity with warming[66,67]. This indicates that lichens in the catchment acted as the main primary rock weathering fungi for the initial rock breakdown[61], while mycorrhizae took over the role for finer mineral weathering by releasing inorganic nutrients from minerals after first soil establishment[68]. Warming is assumed to impact the diversity and composition of mycorrhizal communities rather than their relative abundance[69], an assumption supported by our data.

On the taxon level, we find an increase in mycorrhizal *Rhizophagus* as well as Glomerales with the onset of the Holocene (Fig. 2). Glomerales are arbuscular mycorrhizal fungi, living in symbiosis with around 80% of the vascular land plants[70]. Our data suggest a strong dependency of woody taxa on arbuscular mycorrhizae compared to tundra species (Figs 2, 6). Suillineae co-occurred with *Larix* invasion at the onset of the Holocene, disappeared during the *Picea* forest stage, and reappeared in the late Holocene with a second peak in *Larix*, as confirmed by co-occurrence analysis (Fig. 2, Supplement 3). This finding supports studies by Zhou and Hogetsu[71] and Praeg and Illmer[72], who highlight Suillineae as an important *Larix* mycorrhizal associate. Suillineae species are ectomycorrhizal fungi - usually a symbiosis of woody taxa and Asco- or Basidiomycetes[21] - which are known to enhance weathering by secreting oxalate[73,74].

We also note an increase in *Laccaria* species after 5 cal ka BP (Fig. 2). *Laccaria* are known to form mycorrhizae with Pinaceae, but also Salicaceae and Fabaceae[75]. In mycorrhizal associations with *Larix*, *Laccaria* is known to reduce the amount of phenolics, which defend plant roots against parasitic

fungi[76]. In the late Holocene, *Hyaloscypha* species became more abundant, coinciding with increasing *Vaccinium* abundance (blueberry; Ericaceae family), confirming known interactions from modern studies[77]. Overall, most of our detected mycorrhizal taxa are known to be non-specific to distinct plant species[21,70,78]. This suggests that mycorrhizal fungi in such an extremely cold and nutrient-poor habitat must be generalists, supporting a broad diversity of plants in their growth.

4.3.4 Turnover in carbon, nitrogen, and sulphur cycling

Our bacterial record reveals a dominance of carbon cyclers during the Late Glacial tundra phase while nitrogen cyclers become more abundant with woody-taxa densification from about 15 ka on (Fig. 6). This finding supports a 7-year monitoring study on the modern Siberian tundra-taiga ecotone that shows increasing nitrogen cycling with densification of the tree stands[79]. Additionally, the N content in tundra soils was generally increased when exposed to warming[80,81]. We find the major turnover from carbon to nitrogen cyclers occurred at the transition from light *Larix*-dominated forest to dark *Picea*-dominated coniferous taiga. Surprisingly, our study suggests that the forest composition has a larger effect on C/N cycling than the general invasion of forest.

As well as their relative share, the composition of the carbon-cycling community also shifted along with vegetation change (Figs 2, 6). While we recovered a high relative abundance of polyaromatic degraders in the Late Glacial (*Brevundimonas*, *Caulobacter*), polysaccharide degraders gained importance in the Holocene (e.g. *Paenibacillus* spp.). This aligns with modern spatial gradients where tundra soils in northern Siberia were found to contain polyaromatic compounds[82], while forest soils generally had a larger proportion of microbial polysaccharides[83].

A shift from bacterial-only carbon cycling to fungal-bacterial co-cycling is revealed, probably because the boreal litter is very difficult to decompose due to its high amount of phenol-rich substrates[84]. It is known that the increasing abundance of phenolic acids in soil has a stimulating effect on the abundance of saprotrophic fungi[85,86] as they can tolerate high concentrations of phenolic compounds and degrade them[87].

We also detect a diversification of the nitrogen pathways (Figs 2, 5, 6). During the Late Glacial, nitrogen fixation was mainly from the air (*Rhizobium* spp.). Wettening and warming in the Holocene[38] and subsequent establishment of dark taiga resulted in a diversification of nitrification processes, including ammonium oxidation (*Nitrosomas* spp.) as well as nitrite oxidation (*Nitrotoga*, *Pseudolabrys*). To date, current knowledge diverges: some studies showed that increasing moisture[88] as well as temperature, to a certain extent[89], stimulate nitrogen fixation in High Arctic ecosystems[90]. Experimental warming can also induce a decline in nitrogen fixation in arctic tundra sites[88]. Previous studies[91–93] have shown

that, depending on the tree species, between 20-40% (*Picea* plantation) and 70% (*Picea-Abies* forest) of the nitrogen fixation from the air is retained in the canopy. Our study suggests that the establishment of (dark) evergreen taiga results in higher foliage retention, meaning that nitrogen is directly captured in the crowns without microbial biomass being involved[94]. Subsequently, the fixation of nitrogen in soil from the air is hindered[95,96] and the nitrogen cycling pathways diverge.

Sulphur cyclers show a peak dominated by *Thiobacillus* during the initial plant establishment when sulphur cycling was vital for amino acids (methionine, cysteine) and such protein synthesis (Figs 2, 6). In contrast, we find that low S cycling occurs during phases of increased weathering. During these phases, S adsorbed onto Fe and Al hydroxides likely became a good source of plant-available sulphur as well as a hindering S leaching [97] such that the need for additional bacterial cycling was low. A warmer climate and the establishment of forest soils resulted in stronger sulphur cycling, as indicated by high abundances of sulphur oxidizers, including *Sulfuriferula plumbiphila*, mostly replacing *Thiobacillus*. Warming has previously been demonstrated to result in a higher relative abundance of sulphur cycling genes in tundra soils[98] and to induce high amino acid turnover (mineralization and subsequently re-uptake) in Alaskan taiga soils[99]. Our data underline the demand for diversified sulphur sources with slower plant turnover in boreal forests.

4.3.5 Tracing podzolization

We reveal an increase in taxa preferring acidic soil alongside soil development (Fig. 6). The peak of acidophilus plants is observed at 2.5 cal ka BP when they make up around 50% of the relative abundance (mainly spruce, larch, alder, birch, pine, blueberry - *Picea, Larix, Alnus, Betula, Pinus*, *Vaccinium*) of taxa with known pH preferences. This agrees with modern data comparing multiple forest sites, which show that *Picea* forests have the lowest pH[100]. Slowly decomposing litter[101] as well as poor buffer capacity[102] in evergreen *Picea* forests leads to recalcitrance (e.g. high C/N ratios and lignin concentrations[101,103]), while *Larix* litter has comparatively high base cation fluxes[104]. Additionally, less nutrients in the soils in *Picea* forest inhibit organic matter breakdown, leading to the formation of organic acids and subsequently to acidification of the soil[24,26].

The bacterial community also shows a strong peak in acidophilus taxa in the late Holocene (Fig. 6). The impact of soil acidification on the composition of bacterial communities is known to be driven by ecological filtering (i.e. better adapted taxa invade in the area)[105]. The acidic-preferring bacteria in our data show a shift from mainly *Delftia* species during the Late Glacial towards *Sideroxydans* dominance in the Holocene.

Fungi also show an increasing acid-preference in the late Holocene, though the signal is not as pronounced as for the other organismic groups, being underlined by many acid-tolerant fungi such as *Trichoderma* or *Hyaloscypha,* which occur throughout the whole record (Figs 2, 6). Rousk and Baath[106] find that a lowering of the pH generally leads to increased fungal growth while bacterial growth is decreased. Our data takes this further by showing that a lowering of the pH leads to more acid-tolerant bacteria, while fungal communities are less affected by pH changes. Alongside the acidification of the soil, we see evidence for podzolization with increasing iron cycling around 7-6 cal ka BP in our record (Figs 2, 6), coincident with the onset of the temperature and moisture maximum in the region[38], suggesting that podzolization is, to some extent, also warmth- and rain-induced.

The strong peak in iron-cycling bacteria (mainly *Ferrigenium kumadai* and *Sideroxydans*) in the late Holocene, alongside the re-expansion of *Larix* into *Picea* forests, indicates an iron deficiency induced by leaching (Figs 2, 6). We further detect an increased relative abundance of *Vaccinium* in the late Holocene, suggesting compositional differences between the early *Larix* forest at the onset of the Holocene and the late Holocene *Larix* forest that resulted from podzolization during the preceding *Picea* phase. A comparison of modern *Larix* and *Picea* communities reveals a higher Fe concentration in the soil with *Picea* than in that with *Larix*[107]. We assume that the iron cycling is not impacted by the general presence of *Larix*, but the change in soil composition induced by the preceding *Picea* is evidence of a trajectory of soil development in the area.

4.4 Implications and conclusions

By analysing sedaDNA shotgun metagenomics from sediments from lake Lama (northern-central Siberia) we reveal a pronounced vegetation shift from tundra towards taiga at the Late Glacial-Holocene transition. Alongside, we find a lichen decline but an increase in mycorrhizae as well as a shift from carbon-dominated nutrient cycling towards nitrogen-dominance.

With our study we have shown that lake sedaDNA is not only a valuable tool for analysing compositional changes of plants, fungi, and bacteria but also allows the reconstruction of soil development. We show that the establishment of fungal and bacterial soil communities is, to a great extent, influenced by the vegetation cover, followed by temperature variation, while the time since initial soil development only plays a minor role. This suggests there could be significant turnover in the soil microbiome under future global warming alongside shifting treelines. As the relationship between plants and their associated microorganisms is rather tight, understanding drivers of soil microbiome communities is an asset when developing advanced fertilisers adapted to global warming scenarios.

We found evidence of rapid initial weathering of basalt after glacier retreat with herb- or shrub-dominated tundra, which declined with the warming-induced taiga invasion. Understanding past weathering enables the application of its mechanisms to address ongoing global challenges: weathering of basalt is a known carbon sink for atmospheric CO_2 [108,109]. Powdered basalt grains can be applied to soils to enhance weathering and thus lock-up large amounts of carbon dioxide, removing it from the global carbon cycle[110]. We show that the most basalt weathering occurred during *Salix* dominance in the river valleys, which is of particular interest when assessing general soil development under ongoing global change and making use of basaltic carbon capturing potentials.

We could show a shift from carbon-dominated nutrient cycling in the Late Glacial towards intensified nitrogen cycling in the Holocene. An intensified need for diverse nitrogen cycling in taiga vegetation in comparison to tundra is noted, highlighting the differences between tundra and taiga turnover times. Our data additionally reveal a diversification of sulphur sources in boreal forests for plant establishment and their amino acid synthesis. Altogether, the data provide strong evidence for a relationship between nutrient pools and cycling in relation to plant life cycles.

Finally, our data also trace the establishment of podzol in the study area with increasing iron cycling in the Holocene as well as soil acidification. We highlight that the early and the late larch forests show differences in their underlying herbaceous taxa, inferred from changing soil and therefore growing conditions. The podzolization process was initiated with the establishment of dark evergreen taiga. The re-invasion of larch forest into the area in the late Holocene indicates that soil establishment is a trajectory as podzolization is irreversible despite changing vegetation cover. This might be an important result for forecasting future plant establishment or even be applicable to foster soil development in agricultural settings where multiple plant types need to be grown on the same ground.

Our study provides basic knowledge for understanding and forecasting treeline advance under future global warming. It forms the base for the development of potential afforestation strategies and as such highlights the potential of large-scale carbon-capture enhancement through boreal forest establishment alongside basalt grain weathering.

4.5 Material and methods

4.5.1 Geographical setting and study site

Lake Lama (69.32°N, 90.12°E; 53 m a.s.l.) is located on the Taymyr Peninsula, northern-central Siberia, at the western rim of the basaltic Putorana Plateau (Fig. 1). The current vegetation in the area is comprised of dense taiga with *Picea*, *Larix*, and *Betula*, as well as shrubs such as *Alnus fructicosa*, *Salix,* and *Juniperus communis*, and dwarf shrubs[38]. Modern mean temperatures are 13.8°C for July and -28.8°C

for January (Volochanka weather station; distance to the lake: 247 km[111]). In 1997, an 18.85 m long sediment core (PG1341) was retrieved from the lake at a depth of 66 m, dating back to the last about 23 ka. Prior to processing, the sediment has been stored in the dark and at 4°C. We refined the age-depth model of von Hippel et al.[34] (Supplement 1).

4.5.2 X-ray fluorescence scanning of the sediment core

X-ray fluorescence (XRF) scanning was conducted at the University of Cologne, Germany, on one core half using an Itrax core scanner (Cox Analytical Systems, Sweden) equipped with a Cr-tube and a silicon-drift detector (SDD) in combination with a multi-channel analyser. Analyses were performed at 30 kV and 55 mA, at a resolution of 2 mm and an integration time of 6 seconds. Results are semi-quantitative estimates of relative concentrations of the detected elements[112], derived from the detected peak area intensities and given in total counts per second (cps). The K and Ti count data were normalised to the K/Ti element ratio to account for variations in organic content and other elements[113,114].

4.5.3 Core sub-sampling

The sub-core segments were sampled for sedaDNA in the climate chamber of the Helmholtz Centre Potsdam - German Research Centre for Geosciences (GFZ). During the sampling process, protective clothing as well as face masks and hair nets were worn. Before sampling, the surfaces of the cores were scraped twice with clean knives. The samples were taken using four knives and were then placed in sterile 8 ml Sarstedt tubes and frozen to - 20°C until further processing. We included a total of 44 samples in the study, with an interval of approximately 500 years between samples. A more detailed description of the procedure for preparing and cleaning the chamber as well as the subsampling is provided by von Hippel et al.[34].

4.5.4 DNA extraction

The extraction of the sedaDNA was conducted in the dedicated ancient DNA laboratories at AWI Potsdam, using the DNeasy PowerMax Soil DNA Isolation Kit (Qiagen), following the manufacturer's instruction. An additional incubation step overnight in a rotation incubator at 56°C was added and the elution of the DNA was performed as described in von Hippel et al.[34]. After the extraction, the DNA was concentrated using the GeneJET PCR purification Kit (Thermo Fisher Scientific, Germany) by which 1 mL of the DNA extract was reduced to a volume of 50 µL. The concentrated extracts were measured with a Qubit dsDNA BR assay kit using a Qubit 4.0 Fluorometer (Thermo Fisher Scientific, Germany),

diluted to a final concentration of 3 ng/µL and stored in aliquots of 15 µL to avoid extensive freeze-thaw cycles.

4.5.5 Single stranded DNA library build

The DNA libraries were built following the single stranded DNA library preparation protocol of Gansauge et al.[115] with the ligation of the second adapter in a rotating incubator as described by Schulte et al.[116], using 30 ng of DNA as input. Furthermore, the libraries were quantified with qPCR[117]. Further details on the protocol are described in Schulte et al.[116]. For the setup of the index PCR, we used 1x AccuPrime Pfx reaction mix, 2.5 U/µL AccuPrime Pfx Polymerase, 4 µL of P7_X indexing primer (10 µM) and P5_X indexing primer (10 µM), and 57 µL of deionized water. 24 µL of the final DNA library were added to the reaction. The PCR was conducted according to the following protocol: 2 min at 95 °C, 20 s at 95°C, 30 s at 60°C, 1 min at 68°C and final elongation for 5 min at 68°C. The appropriate number of amplification cycles (steps 2-4) for the index PCR was calculated from the qPCR results and varied between 11 and 13 cycles for samples and blank controls.

The PCR products were purified with MinElute (Qiagen, Switzerland) according to the manufacturer's instructions and eluted in 30 µL elution buffer. The DNA library concentration was determined using a Qubit 4.0 Fluorometer dsDNA BR assay kit (Thermo Fisher Scientific, Germany). For the quality control and to measure the fragment length composition, we loaded the libraries on a TapeStation (Agilent, United States). Mean fragment length and concentration of indexed libraries was used to calculate the molarity of each library and equimolar library pools were prepared. In total, we compiled 3 library pools.

The library pools APMG-37 (10 samples, 4 library blanks, 1 extraction blank) and APMG-38 (10 samples, 3 library blanks, 2 extraction blanks) were sent to Fasteris SA, Switzerland, and were run on a NovaSeq device (2x100 bp). A table with the sample composition of the sequencing runs as well as their metadata is provided in supplement table 1. The second (23 samples, 6 library blanks, 3 extraction blanks) and third (15 samples, whereof 14 were sample replicates to increase read counts for poorly sequenced samples) library pools were sequenced on an NextSeq 2000 platform (2x100 bp) at AWI Bremerhaven, Germany.

4.5.6 Bioinformatic pipeline for the analysis of the sequencing results

The analysis of the raw sequencing data included a quality check using *fastQC* (version 0.11.[118]) and a deduplication step (removing identical reads) with *clumpify* (BBmap version 38.87,

https://sourceforge.net/projects/bbmap/). The paired-end forward and reverse reads were merged with *fastp* (version 0.20.1[119]) applying a low complexity filter in order to remove reads of low complexity from the dataset. Taxonomic classification was done with *Kraken2*[120] against the nt database by NCBI (ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/nt.gz; download: 10/2022, with default k-mer size 35) with a confidence threshold of 0.8. We also tested the taxonomic classification of bacteria against the refseq database by NCBI[121], yielding comparable patterns (Supplement 4). The Kraken report files were converted into a txt-file using the command awk as an input in R.

4.5.7 Data analysis

The analysis of the processed DNA data was done in R, version 4.0.3 [123]. As a first step we combined the converted kraken file, with metadata (depth and age of the sediment samples) and a lineage file, which adds the full taxonomic lineage of the identified taxa via TaxID (Supplement table 1). The raw reads of all three sequencing runs were finally merged in R.

Three taxonomic data subsets (plants, fungi, bacteria) were created. For the plant dataset, we extracted all reads assigned to the clade "Viridiplantae" and kept those reads which were at least assigned to genus level. We cleaned the plant subset from aquatic and non-Siberian taxa (list of taxa: Supplement table 2) and kept those taxa which occurred in at least three samples. The fungus subset is defined by all reads of the kingdom "Fungi". Due to generally poorly sequenced fungal genomes and therefore their absence in databases, we decided to work on taxa assigned to at least phylum level for terrestrial fungi (list of taxa: Supplement table 3) and kept all taxa occurring in at least three samples.

The bacterium subset contains reads assigned to the domain of "Bacteria". Among bacteria, we only kept reads which were assigned to at least genus level and excluded taxa, which are characteristic of aquatic habitats (Supplement table 4). Further, we filtered for taxa that occurred in at least three samples and with a minimum of 20 reads in all samples.

4.5.8 Analysis of the ancient patterns

To assess the ancient origin of the analysed reads, we selected key taxa for ancient damage pattern analysis with MapDamage (v. 2.0.8, [122]), including *Larix sibirica* (plants), *Hyaloscypha bicolor* (fungi) and *Herminiimonas arsenitoxidans* (bacteria). Prior to the analyses, we grouped the samples into two subgroups (Holocene and Pleistocene samples) and merged their raw sequencing data using the cat command. The Holocene group includes samples from 0–10.4 cal ka BP (18 samples), and the Pleistocene group from 11.5–23 cal ka BP (27 samples) respectively. With the merged files, we

repeated the bioinformatic pipeline as described above. From the newly generated kraken output, we extracted taxon-specific reads on species level and mapped them against their reference genomes (*Larix sibirica*, accession number MT797187; *Hyaloscypha bicolor*, accession number GCF_002865645.1; *Herminiimonas arsenitoxidans*, accession number GCF_900130075.1) using MapDamage (v. 2.0.8 [122]) with the options 'rescale' and 'single-stranded'. Fragment length, incorporation plots, distribution of the C to T changes, and predicted nucleotide changes for the last 25 nucleotides of the analysed reads are presented in the supplement 5-7.

4.5.9 Statistical analysis of the dataset

All statistical analyses were carried out with the software R, version 4.0.3 [123]. For the analysis, we resampled 100 times each taxonomic subset to the basecount of the sample mean value to balance uneven read counts between the samples. The resampling of the fungi had a mean count of 272, the plant data of 3,533, and the bacteria of 24,285. We followed the github script of Kruse (2019, https://github.com/StefanKruse/R_Rarefaction). All statistical analyses were performed on the resampled datasets.

For analysis of the long-term soil development and its driving forces, we assessed the time (i.e. the age of the sediment) and the impact of multiple environmental variables (vegetation, temperature). We defined "soil development" as the bacterial communities on one side and fungal soil communities on the other side (whole subsets). Constrained ordination analyses were run to statistically relate the environmental variables to the variation in the composition of either soil community. To yield the variable "Vegetation", we performed a redundancy analysis (RDA) on the double-square rooted vegetation subset and determined the significant PC axes using PCAsignificance(), provided the first two PC axes were significant. We used the scores of the PC axes 1 and 2 and merged them as a dataframe. We used a reconstruction of the temperature variation in the Northern Hemisphere as further input for the environmental variables to yield the variable "Temperature". The reconstruction of the temperature variation followed the script of Kruse et al. (https://github.com/StefanKruse/R_PastElevationChange). In brief, it is based on the mean temperature reconstructions by Shakun et al.[125] and Marcott et al.[126]. For the variable "Time", we used the respective ages of the samples as the input.

The K/Ti element ratio data derived from the XRF scanning was used as a proxy for weathering[58–60]. The analysis of the XRF data was done by smoothing the scanning data with the function predict (package: stats[123]). All data were plotted with ggplot2 (package: tidyverse[127]).

We analysed co-occurrence patterns between mycorrhizal fungi and respective tree taxa. To do so, we analysed the Spearman correlation between the fungal dataset and the plant dataset using the function cor (package stats[123]). The correlation matrix was converted in a dataframe and only positively correlated mycorrhizal taxa with a correlation value of at least 0.2 were selected. With the final taxa selection, we plotted the reduced data using the function corrplot (package corrplot[124]).

For assessing the pH preferences of all data subsets, we assigned the taxa to five categories of preference, namely acidic, slightly acidic, neutral, slightly alkaline, and alkaline (supplement tables 2-4). We merged the percentages of the slightly alkaline and alkaline preferring taxa to yield the overall alkaline preference for plotting. The displayed acidic-preferring taxa are only those being assigned to strong acidic preference.

Acknowledgements

We are grateful to Janine Klimke for the help with pooling the libraries, and to Volker Wennrich and Nicole Mantke for performing the XRF measurements of the sediment halves. We are also grateful to Amedea Perfumo for the fruitful discussions about the interpretation of the bacterial patterns. We are thankful for the help of Lars Harms with the bioinformatic analysis of the ancient patterns. Also, we thank Cathy Jenkins for the English proofreading of the manuscript.

4.6 References

1. Wardle, D. A., Walker, L. R. & Bardgett, R. D. Ecosystem properties and forest decline in contrasting long-term chronosequences. *Science* **305**, 509–513 (2004).
2. Davidson, E. A. & Janssens, I. A. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature* **440**, 165–173 (2006).
3. Delgado-Baquerizo, M. *et al.* The influence of soil age on ecosystem structure and function across biomes. *Nat. Commun.* **11**, 4721 (2020).
4. Jackson, T. A. Weathering, secondary mineral genesis, and soil formation caused by lichens and mosses growing on granitic gneiss in a boreal forest environment. *Geoderma* **251–252**, 78–91 (2015).
5. Ewing, H. A. The influence of substrate on vegetation history and ecosystem development. *Ecology* **83**, 2766–2781 (2002).
6. Kelly, E. F., Chadwick, O. A. & Hilinski, T. E. The effect of plants on mineral weathering. *Biogeochemistry* **42**, 21–53 (1998).
7. Finlay, R. *et al.* The role of fungi in biogenic weathering in boreal forest soils. *Fungal Biol. Rev.* **23**, 101–106 (2009).

8. Berthelin, J. *et al.* Bioreduction of ferric species and biogenesis of green rusts in soils. *Comptes Rendus Geosci.* **338**, 447–455 (2006).

9. Chen, J., Blume, H.-P. & Beyer, L. Weathering of rocks induced by lichen colonization — a review. *CATENA* **39**, 121–146 (2000).

10. Shaler, N. S. *The Origin and Nature of Soils*. (U.S. Government Printing Office, 1892).

11. Courty, P.-E. *et al.* The role of ectomycorrhizal communities in forest ecosystem processes: New perspectives and emerging concepts. *Soil Biol. Biochem.* **42**, 679–698 (2010).

12. Lambers, H., Atkin, O. K. & Millenaar, F. F. Respiratory patterns in roots in relation to their functioning. chap. 32 *Plant Roots* (CRC Press, 2002).

13. Hoffland, E. *et al.* The role of fungi in weathering. *Front. Ecol. Environ.* **2**, 258–264 (2004).

14. Van Der Heijden, M. G. A., Bardgett, R. D. & Van Straalen, N. M. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol. Lett.* **11**, 296–310 (2008).

15. Cotrufo, M. F., Wallenstein, M. D., Boot, C. M., Denef, K. & Paul, E. The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: do labile plant inputs form stable soil organic matter? *Glob. Change Biol.* **19**, 988–995 (2013).

16. Trumbore, S. Carbon respired by terrestrial ecosystems – recent progress and challenges. *Glob. Change Biol.* **12**, 141–153 (2006).

17. Liang, C. & Balser, T. C. Microbial production of recalcitrant organic matter in global soils: implications for productivity and climate policy. *Nat. Rev. Microbiol.* **9**, 75; author reply 75 (2011).

18. Lindahl, B. D. *et al.* Spatial separation of litter decomposition and mycorrhizal nitrogen uptake in a boreal forest. *New Phytol.* **173**, 611–620 (2007).

19. Dixon, R. & Kahn, D. Genetic regulation of biological nitrogen fixation. *Nat. Rev. Microbiol.* **2**, 621–631 (2004).

20. George, E., Marschner, H. & Jakobsen, I. Role of arbuscular mycorrhizal fungi in uptake of phosphorus and nitrogen from soil. *Crit. Rev. Biotechnol.* **15**, 257–270 (1995).

21. Smith, S. E. & Read, D. J. *Mycorrhizal Symbiosis*. (Academic Press, 2010).

22. Miransari, M. Arbuscular mycorrhizal fungi and nitrogen uptake. *Arch. Microbiol.* **193**, 77–81 (2011).

23. Pugh, T. A. M. *et al.* Understanding the uncertainty in global forest carbon turnover. *Biogeosciences* **17**, 3961–3989 (2020).

24. Lundström, U. S., van Breemen, N. & Bain, D. The podzolization process. A review. *Geoderma* **94**, 91–107 (2000).

25. Wiklander, L. & Andersson, A. The replacing efficiency of hydrogen ion in relation to base saturation and pH. *Geoderma* **7**, 159–165 (1972).

26. Sauer, D. *et al.* Podzol: Soil of the Year 2007. A review on its genesis, occurrence, and functions. *J. Plant Nutr. Soil Sci.* **170**, 581–597 (2007).

27. D'Amico, M. E., Freppaz, M., Filippa, G. & Zanini, E. Vegetation influence on soil formation rate in a proglacial chronosequence (Lys Glacier, NW Italian Alps). *CATENA* **113**, 122–137 (2014).

28. Giguet-Covex, C. *et al.* New insights on lake sediment DNA from the catchment: importance of taphonomic and analytical issues on the record quality. *Sci. Rep.* **9**, 14676 (2019).

29. Capo, E. *et al.* Lake sedimentary DNA research on past terrestrial and aquatic biodiversity: overview and recommendations. *Quaternary* **4**, 6 (2021).

30. Parducci, L. *et al.* Shotgun environmental DNA, pollen, and macrofossil analysis of lateglacial lake sediments from southern Sweden. *Front. Ecol. Evol.* **7**, (2019).

31. Huang, S. *et al.* Plant sedimentary ancient DNA from Far East Russia covering the last 28,000 years reveals different assembly rules in cold and warm climates. *Front. Ecol. Evol.* **9**, (2021).

32. Rijal, D. P. *et al.* Sedimentary ancient DNA shows terrestrial plant richness continuously increased over the Holocene in northern Fennoscandia. *Sci. Adv.* **7**, eabf9557 (2021).

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

34. von Hippel, B. *et al.* Long-term fungus–plant covariation from multi-site sedimentary ancient DNA metabarcoding. *Quat. Sci. Rev.* **295**, 107758 (2022).

35. Courtin, J. *et al.* Pleistocene glacial and interglacial ecosystems inferred from ancient DNA analyses of permafrost sediments from Batagay megaslump, East Siberia. *Environ. DNA* **4**, (2022).

36. Kjær, K. H. *et al.* A 2-million-year-old ecosystem in Greenland uncovered by environmental DNA. *Nature* **612**, 283–291 (2022).

37. Wang, Y. *et al.* Late Quaternary dynamics of Arctic biota from ancient environmental genomics. *Nature* **600**, 86–92 (2021).

38. Andreev, A. A. *et al.* Vegetation and climate changes around the Lama Lake, Taymyr Peninsula, Russia during the Late Pleistocene and Holocene. *Quat. Int.* **122**, 69–84 (2004).

39. Zimmermann, H. H. *et al.* The history of tree and shrub taxa on Bol'shoy Lyakhovsky Island (New Siberian Archipelago) since the last interglacial uncovered by sedimentary ancient DNA and pollen data. *Genes* **8**, 273 (2017).

40. Clarke, C. L. *et al.* A 24,000-year ancient DNA and pollen record from the Polar Urals reveals temporal dynamics of arctic and boreal plant communities. *Quat. Sci. Rev.* **247**, 106564 (2020).

41. Zumsteg, A. *et al.* Bacterial, archaeal and fungal succession in the forefield of a receding glacier. *Microb. Ecol.* **63**, 552–564 (2012).

42. Korneikova, M. V. Comparative analysis of the number and structure of the complexes of microscopic fungi in tundra and taiga soils in the north of the Kola Peninsula. *Eurasian Soil Sci.* **51**, 89–95 (2018).

43. Muller, D. *et al. Herminiimonas arsenicoxydans* sp. nov., a metalloresistant bacterium. *Int. J. Syst. Evol. Microbiol.* **56**, 1765–1769.

44. Maestre, F. T. & Cortina, J. Spatial patterns of surface soil properties and vegetation in a Mediterranean semi-arid steppe. *Plant Soil* **241**, 279–291 (2002).

45. Kabala, C. *et al.* Soil development and spatial differentiation in a glacial river valley under cold and extremely arid climate of East Pamir Mountains. *Sci. Total Environ.* **758**, 144308 (2021).

46. Pedersen, M. W. *et al.* Postglacial viability and colonization in North America's ice-free corridor. *Nature* **537**, 45–49 (2016).

47. Garner, R. E., Gregory-Eaves, I. & Walsh, D. A. Sediment metagenomes as time capsules of lake microbiomes. *mSphere* **5**, e00512-20 (2020).

48. Zimmermann, H. H. *et al.* Marine ecosystem shifts with deglacial sea-ice loss inferred from ancient DNA shotgun sequencing. *Nat. Commun.* **14**, 1650 (2023).

49. Deslippe, J. R., Hartmann, M., Simard, S. W. & Mohn, W. W. Long-term warming alters the composition of Arctic soil microbial communities. *FEMS Microbiol. Ecol.* **82**, 303–315 (2012).

50. Krishna, M. *et al.* Successional trajectory of bacterial communities in soil are shaped by plant-driven changes during secondary succession. *Sci. Rep.* **10**, 9864 (2020).

51. Schmidt, N. & Bölter, M. Fungal and bacterial biomass in tundra soils along an arctic transect from Taimyr Peninsula, central Siberia. *Polar Biol.* **25**, 871–877 (2002).

52. Berg, M. P., Kniese, J. P. & Verhoef, H. A. Dynamics and stratification of bacteria and fungi in the organic layers of a Scots pine forest soil. *Biol. Fertil. Soils* **26**, 313–322 (1998).

53. Franzetti, A. *et al.* Early ecological succession patterns of bacterial, fungal and plant communities along a chronosequence in a recently deglaciated area of the Italian Alps. *FEMS Microbiol. Ecol.* **96**, fiaa165 (2020).

54. Jiang, Y. *et al.* Divergent assemblage patterns and driving forces for bacterial and fungal communities along a glacier forefield chronosequence. *Soil Biol. Biochem.* **118**, 207–216 (2018).

55. Voříšková, J., Elberling, B. & Priemé, A. Fast response of fungal and prokaryotic communities to climate change manipulation in two contrasting tundra soils. *Environ. Microbiome* **14**, 6 (2019).

56. Liu, S. K., Han, C., Liu, J. M. & Li, H. Hydrothermal decomposition of potassium feldspar under alkaline conditions. *RSC Adv.* **5**, 93301–93309 (2015).

57. Manning, D. A. C., Baptista, J., Sanchez Limon, M. & Brandt, K. Testing the ability of plants to access potassium from framework silicate minerals. *Sci. Total Environ.* **574**, 476–481 (2017).

58. Clift, P. D. *et al.* Correlation of Himalayan exhumation rates and Asian monsoon intensity. *Nat. Geosci.* **1**, 875–880 (2008).

59. Gebregiorgis, D. *et al.* What can we learn from X-ray fluorescence core scanning data? A paleomonsoon case study. *Geochem. Geophys. Geosystems* **21**, (2020).

60. Vyse, S. A. *et al.* Geochemical and sedimentological responses of arctic glacial Lake Ilirney, Chukotka (far east Russia) to palaeoenvironmental change since ~51.8 ka BP. *Quat. Sci. Rev.* **247**, 106607 (2020).

61. Zambell, C. B., Adams, J. M., Gorring, M. L. & Schwartzman, D. W. Effect of lichen colonization on chemical weathering of hornblende granite as estimated by aqueous elemental flux. *Chem. Geol.* **291**, 166–174 (2012).

62. Varadachari, C., Barman, A. K. & Ghosh, K. Weathering of silicate minerals by organic acids II. Nature of residual products. *Geoderma* **61**, 251–268 (1994).

63. Vodyanitskii, Yu. N. Iron hydroxides in soils: A review of publications. *Eurasian Soil Sci.* **43**, 1244–1254 (2010).

64. Pedersen, H. D., Postma, D. & Jakobsen, R. Release of arsenic associated with the reduction and transformation of iron oxides. *Geochim. Cosmochim. Acta* **70**, 4116–4129 (2006).

65. Koh, H.-W. *et al.* Physiological and genomic insights into the lifestyle of arsenite-oxidizing *Herminiimonas arsenitoxidans*. *Sci. Rep.* **7**, 15007 (2017).

66. Fraser, R. H., Lantz, T. C., Olthof, I., Kokelj, S. V. & Sims, R. A. Warming-induced shrub expansion and lichen decline in the western Canadian Arctic. *Ecosystems* **17**, 1151–1168 (2014).

67. Lang, S. I. *et al.* Arctic warming on two continents has consistent negative effects on lichen diversity and mixed effects on bryophyte diversity. *Glob. Change Biol.* **18**, 1096–1107 (2012).

68. Landeweert, R., Hoffland, E., Finlay, R. D., Kuyper, T. W. & van Breemen, N. Linking plants to rocks: ectomycorrhizal fungi mobilize nutrients from minerals. *Trends Ecol. Evol.* **16**, 248–254 (2001).

69. Law, S. R. *et al.* Metatranscriptomics captures dynamic shifts in mycorrhizal coordination in boreal forests. *Proc. Natl. Acad. Sci.* **119**, e2118852119 (2022).

70. Brundrett, M. C. & Tedersoo, L. Evolutionary history of mycorrhizal symbioses and global host plant diversity. *New Phytol.* **220**, 1108–1115 (2018).

71. Zhou, Z. & Hogetsu, T. Subterranean community structure of ectomycorrhizal fungi under *Suillus grevillei* sporocarps in a *Larix kaempferi* forest. *New Phytol.* **154**, 529–539 (2002).

72. Praeg, N. & Illmer, P. Microbial community composition in the rhizosphere of Larix decidua under different light regimes with additional focus on methane cycling microorganisms. *Sci. Rep.* **10**, 22324 (2020).

73. Balogh-Brunstad, Z. *et al.* Biotite weathering and nutrient uptake by ectomycorrhizal fungus, *Suillus tomentosus*, in liquid-culture experiments. *Geochim. Cosmochim. Acta* **72**, 2601–2618 (2008).

74. Bonneville, S. *et al.* Tree-mycorrhiza symbiosis accelerate mineral weathering: Evidences from nanometer-scale elemental fluxes at the hypha–mineral interface. *Geochim. Cosmochim. Acta* **75**, 6988–7005 (2011).

75. Mueller, G. M. Systematics of *Laccaria* (Agaricales) in the continental United States and Canada, with discussions on extralimital taxa and descriptions of extant types. *Fieldiana Bot.* **30**, (1992).

76. Münzenberger, B., Kottke, I. & Oberwinkler, F. Reduction of phenolics in mycorrhizas of *Larix decidua* Mill. *Tree Physiol.* **15**, 191–196 (1995).

77. Fehrer, J., Réblová, M., Bambasová, V. & Vohník, M. The root-symbiotic *Rhizoscyphus ericae* aggregate and Hyaloscypha (Leotiomycetes) are congeneric: Phylogenetic and experimental evidence. *Stud. Mycol.* **92**, 195–225 (2019).

78. Vohník, M., Figura, T. & Réblová, M. *Hyaloscypha gabretae* and *Hyaloscypha gryndleri* spp. nov. (Hyaloscyphaceae, Helotiales), two new mycobionts colonizing conifer, ericaceous and orchid roots. *Mycorrhiza* **32**, 105–122 (2022).

79. Hewitt, R. E. *et al.* Increasing tree density accelerates stand-level nitrogen cycling at the taiga–tundra ecotone in northeastern Siberia. *Ecosphere* **13**, e4175 (2022).

80. Aerts, R., Cornelissen, J. H. C. & Dorrepaal, E. Plant performance in a warmer world: general responses of plants from cold, northern biomes and the importance of winter and spring events, chap. 5 *Plants and Climate Change* (eds. Rozema, J., Aerts, R. & Cornelissen, H.) 65–78 (Springer Netherlands, 2006). doi:10.1007/978-1-4020-4443-4_5.

81. Gao, W., Sun, W. & Xu, X. Permafrost response to temperature rise in carbon and nutrient cycling: Effects from habitat-specific conditions and factors of warming. *Ecol. Evol.* **11**, 16021–16033 (2021).

82. Gabov, D. N. & Beznosikov, V. A. Polycyclic aromatic hydrocarbons in tundra soils of the Komi Republic. *Eurasian Soil Sci.* **47**, 18–25 (2014).

83. Murayama, S. & Sugiura, Y. Origin of Soil Polysaccharides, and Ectomycorrhizal Fungal Sclerotia as Sources of Forest Soil Polysaccharides, chap. 6 in *Sclerotia Grains in Soils: A New Perspective from Pedosclerotiology* (ed. Watanabe, M.) 91–117 (Springer, 2021). doi:10.1007/978-981-33-4252-1_6.

84. Nilsson, M.-C., Wardle, D. A. & DeLuca, T. H. Belowground and aboveground consequences of interactions between live plant species mixtures and dead organic substrate mixtures. *Oikos* **117**, 439–449 (2008).

85. Wang, Y. *et al.* Environmental behaviors of phenolic acids dominated their rhizodeposition in boreal poplar plantation forest soils. *J. Soils Sediments* **16**, 1858–1870 (2016).

86. Clocchiatti, A., Hannula, S. E., van den Berg, M., Hundscheid, M. P. J. & de Boer, W. Evaluation of phenolic root exudates as stimulants of saptrophic fungi in the rhizosphere. *Front. Microbiol.* **12**, 644046 (2021).

87. Brant, J. B., Sulzman, E. W. & Myrold, D. D. Microbial community utilization of added carbon substrates in response to long-term carbon input manipulation. *Soil Biol. Biochem.* **38**, 2219–2232 (2006).

88. Rousk, K., Sorensen, P. L. & Michelsen, A. What drives biological nitrogen fixation in high arctic tundra: Moisture or temperature? *Ecosphere* **9**, e02117 (2018).

89. Gundale, M. J., Nilsson, M., Bansal, S. & Jäderlund, A. The interactive effects of temperature and light on biological nitrogen fixation in boreal forests. *New Phytol.* **194**, 453–463 (2012).

90. Stewart, K. J., Grogan, P., Coxson, D. S. & Siciliano, S. D. Topography as a key factor driving atmospheric nitrogen exchanges in arctic terrestrial ecosystems. *Soil Biol. Biochem.* **70**, 96–112 (2014).

91. Cape, J. N., Dunster, A., Crossley, A., Sheppard, L. J. & Harvey, F. J. Throughfall chemistry in a Sitka spruce plantation in response to six different simulated polluted mist treatments. *Water. Air. Soil Pollut.* **130**, 619–624 (2001).

92. Chiwa, M., Crossley, A., Sheppard, L. J., Sakugawa, H. & Cape, J. N. Throughfall chemistry and canopy interactions in a Sitka spruce plantation sprayed with six different simulated polluted mist treatments. *Environ. Pollut.* **127**, 57–64 (2004).

93. Bryan Dail, D. *et al.* Distribution of nitrogen-15 tracers applied to the canopy of a mature spruce-hemlock stand, Howland, Maine, USA. *Oecologia* **160**, 589–599 (2009).

94. Sievering, H., Tomaszewski, T. & Torizzo, J. Canopy uptake of atmospheric N deposition at a conifer forest: part I -canopy N budget, photosynthetic efficiency and net ecosystem exchange. *Tellus B Chem. Phys. Meteorol.* **59**, 483–492 (2007).

95. Reed, S. C., Cleveland, C. C. & Townsend, A. R. Tree species control rates of free-living nitrogen fixation in a tropical rain forest. *Ecology* **89**, 2924–2934 (2008).

96. Sparks, J. P. Ecological ramifications of the direct foliar uptake of nitrogen. *Oecologia* **159**, 1–13 (2009).

97. McLaren, R. G., Cameron, K. C., McLaren, R. G. & Cameron, K. C. *Soil Science: Sustainable Production and Environmental Protection*. (Oxford University Press, 1996).

98. Feng, J. *et al.* Warming-induced permafrost thaw exacerbates tundra soil carbon decomposition mediated by microbial community. *Microbiome* **8**, 3 (2020).

99. Jones, D. L. & Kielland, K. Soil amino acid turnover dominates the nitrogen flux in permafrost-dominated taiga forest soils. *Soil Biol. Biochem.* **34**, 209–219 (2002).

100. Priha, O. & Smolander, A. Nitrogen transformations in soil under *Pinus sylvestris*, *Picea abies* and *Betula pendula* at two forest sites. *Soil Biol. Biochem.* **31**, 965–977 (1999).

101. Berger, T. W. & Berger, P. Greater accumulation of litter in spruce (*Picea abies*) compared to beech (*Fagus sylvatica*) stands is not a consequence of the inherent recalcitrance of needles. *Plant Soil* **358**, 349–369 (2012).

102. Desie, E. *et al.* Litter share and clay content determine soil restoration effects of rich litter tree species in forests on acidified sandy soils. *For. Ecol. Manag.* **474**, 118377 (2020).

103. Ellenberg, H. *et al*. Ökosystemforschung. Ergebnisse des Sollingprojekts 1966 - 1986. Ulmer Verlag (1986).

104. Kim, C., An, H.-C., Cho, H.-S. & Choo, G.-C. Base cation fluxes and release by needle litter in three adjacent coniferous plantations. *For. Sci. Technol.* **9**, 225–228 (2013).

105. Zhang, X., Liu, W., Zhang, G., Jiang, L. & Han, X. Mechanisms of soil acidification reducing bacterial diversity. *Soil Biol. Biochem.* **81**, 275–281 (2015).

106. Rousk, J. & Bååth, E. Growth of saprotrophic fungi and bacteria in soil. *FEMS Microbiol. Ecol.* **78**, 17–30 (2011).

107. Bares, R. H. & Wali, M. K. Chemical relations and litter production of *Picea mariana* and *Larix laricina* stands on an alkaline peatland in Northern Minnesota. *Vegetatio* **40**, 79–94 (1979).

108. Gislason, S. *et al.* Rapid solubility and mineral storage of CO_2 in basalt. *Energy Procedia* **63**, 4561–4574 (2014).

109. Snæbjörnsdóttir, S. Ó. *et al.* Carbon dioxide storage through mineral carbonation. *Nat. Rev. Earth Environ.* **1**, 90–102 (2020).

110. Goll, D. S. *et al.* Potential CO_2 removal from enhanced weathering by ecosystem responses to powdered rock. *Nat. Geosci.* **14**, 545–549 (2021).

111. Russian Institute of Hydrometeorological Information: World Data Center, available at http://meteo.ru/english/climate/temp.php (last access: 18. Juni 2021), 2021.

112. Croudace, I., Rindby, A. & Rothwell, R. ITRAX: Description and evaluation of a new multi-function X-ray core scanner. *Geol. Soc. Lond. Spec. Publ.* **267**, 51–63 (2006).

113. Löwemark, L. *et al.* Normalizing XRF-scanner data: A cautionary note on the interpretation of high-resolution records from organic-rich lakes. *J. Asian Earth Sci.* **40**, 1250–1256 (2011).

114. Shala, S. *et al.* Palaeoenvironmental record of glacial lake evolution during the early Holocene at Sokli, NE Finland. *Boreas* **43**, 362–376 (2014).

115. Gansauge, M.-T. *et al.* Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. *Nucleic Acids Res.* **45**, e79 (2017).

116. Schulte, L. *et al.* Hybridization capture of larch (*Larix* Mill.) chloroplast genomes from sedimentary ancient DNA reveals past changes of Siberian forest. *Mol. Ecol. Resour.* **21**, 801–815 (2021).

117. Gansauge, M.-T. & Meyer, M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat. Protoc.* **8**, 737–748 (2013).

118. Andrews, S. A Quality Control Tool for High Throughput Sequence Data [Online]. (2010).

119. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).

120. Malik, I., Pawlik, Ł., Ślęzak, A. & Wistuba, M. A study of the wood anatomy of *Picea abies* roots and their role in biomechanical weathering of rock cracks. *CATENA* **173**, 264–275 (2019).

121. Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **33**, D501–D504 (2005).

122. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).

123. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2021).

124. Wei, T. & Simko, V. R package 'corrplot': Visualization of a Correlation Matrix. (2021).

125. Shakun, J. D. *et al.* Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation. *Nature* **484**, 49–54 (2012).

126. Marcott, S. A., Shakun, J. D., Clark, P. U. & Mix, A. C. A reconstruction of regional and global temperature for the past 11,300 years. *Science* **339**, 1198–1201 (2013).

127. Wickham, H. *et al.* Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).

Declarations

Funding

This research was funded by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (Grant Agreement No. 772852, ERC Consolidator Grant "Glacial Legacy") and the Initiative and Networking Fund of the Helmholtz Association.

Conflict of interest/competing interest

We hereby confirm that this paper has not been published elsewhere nor submitted to another journal for consideration. All authors have approved the manuscript and agree with its submission.

Availability of data and material

The data will be publicly available after the acceptance of the manuscript.

Dryad: sequencing data of all sequencing runs as well as XRF data

Pangaea: upload of revised age-depth model

DNA data: ENA

Author contribution

BvH: study design, subsampling of Lama sediment, DNA extraction, built library, data analysis, writing first draft of the manuscript; KSL: supervision laboratory part and bioinformatics; MM: collection of the sediment record, XRF analysis; UH: study design, supervision of data analysis & writing of first manuscript version. All authors commented on earlier versions of the manuscript.

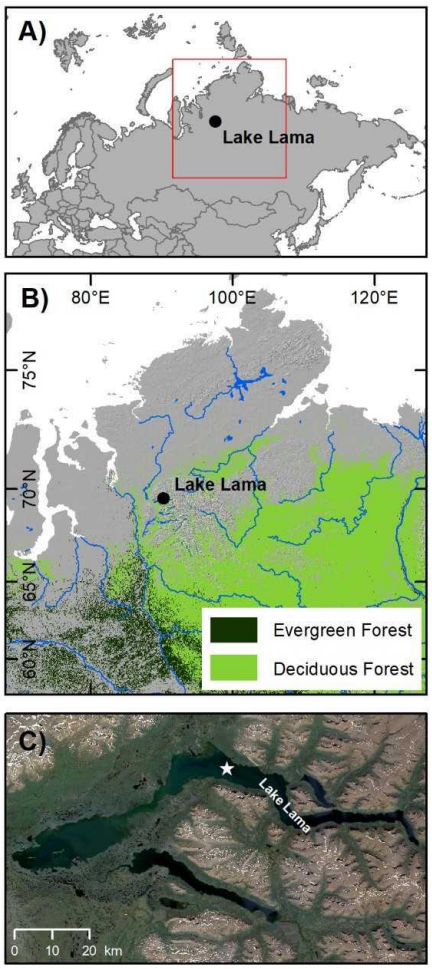


Fig. 1: Map of central Russia showing the location of the study site (A+B). (C): satellite image of the lake and its surroundings. The coring location in the lake is marked with a star. For the "distribution of deciduous and evergreen forests", data from the ESA CCI Land Cover time-series v2.0.7 (1992–2015)- data set were used (https://www.esa-landcover-cci.org/). The land cover classes "70" ("Tree cover, needleleaved, evergreen") and "80" ("Tree cover, needleleaved, deciduous") were extracted for the illustration of the figure.

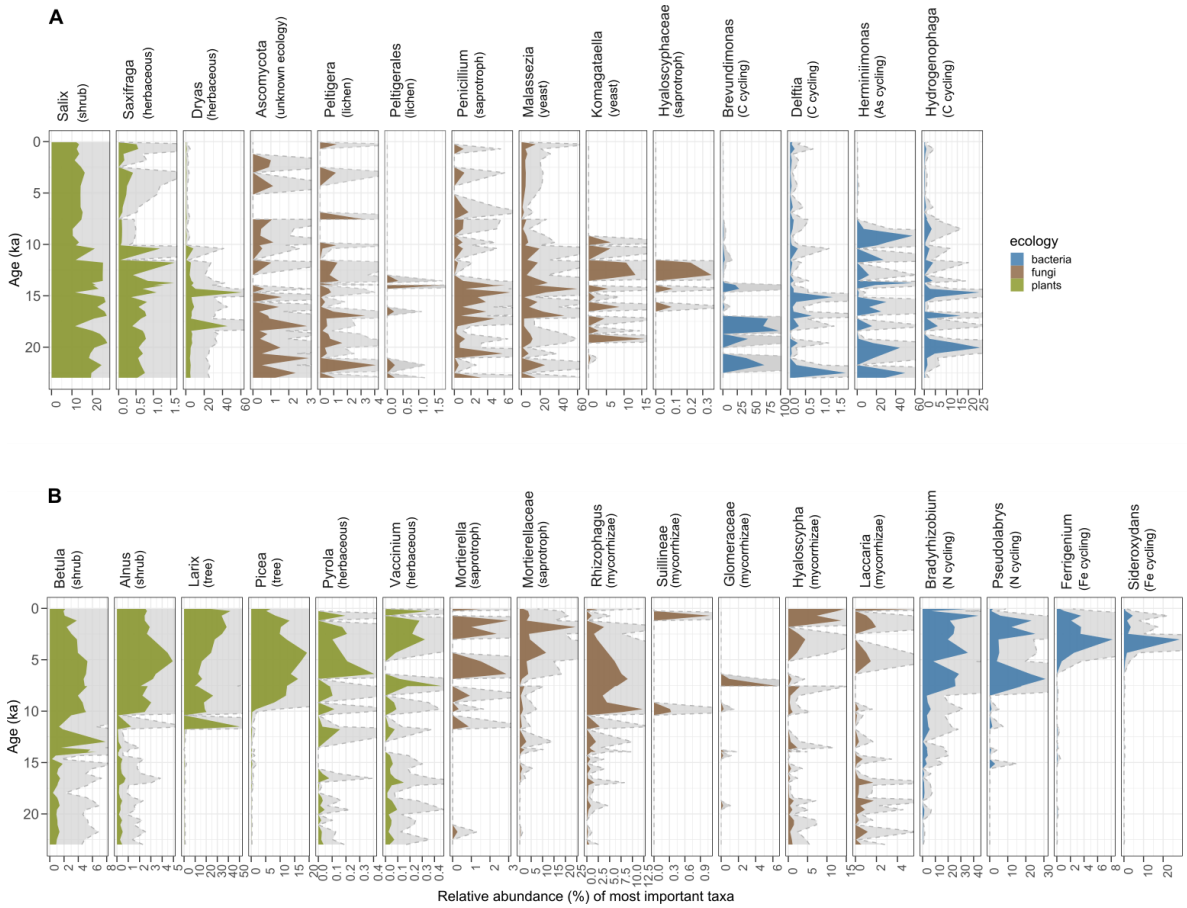


Fig. 2: Relative abundance of the most prominent plants, fungal, and bacterial taxa recovered from the sediment of lake Lama. The abundance is relative to the taxa recovered in the respective subset (plants, fungi, bacteria). In brackets is the respective vegetation type, fungal ecology, or bacterial element cycle. A: Most prominent taxa in the Late Glacial, B: Most prominent taxa in the Holocene.

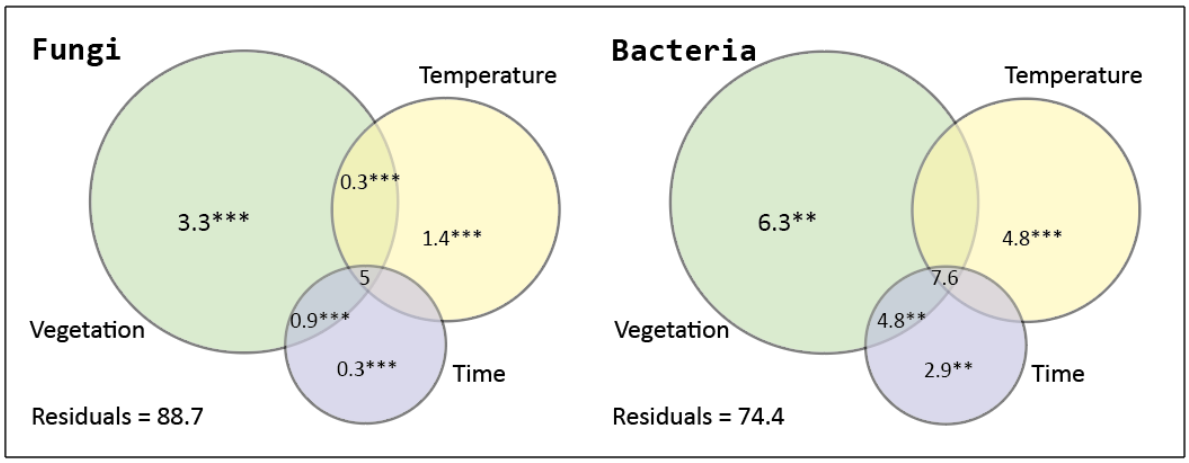


Fig. 3: Variation partitioning of fungi and soil bacteria. The numbers are the percentage of explained variation for the respective driver (for overlapping areas the combined variation). We show that vegetation as a single variable has the greatest impact on the establishment of either community, followed by the temperature. The unexplained variation is given as the percentage of the residuals. The asterisks indicate the statistical significance of each single result (** = $p \leq 0.01$, *** = $p \leq 0.001$). Values <0 are not shown.

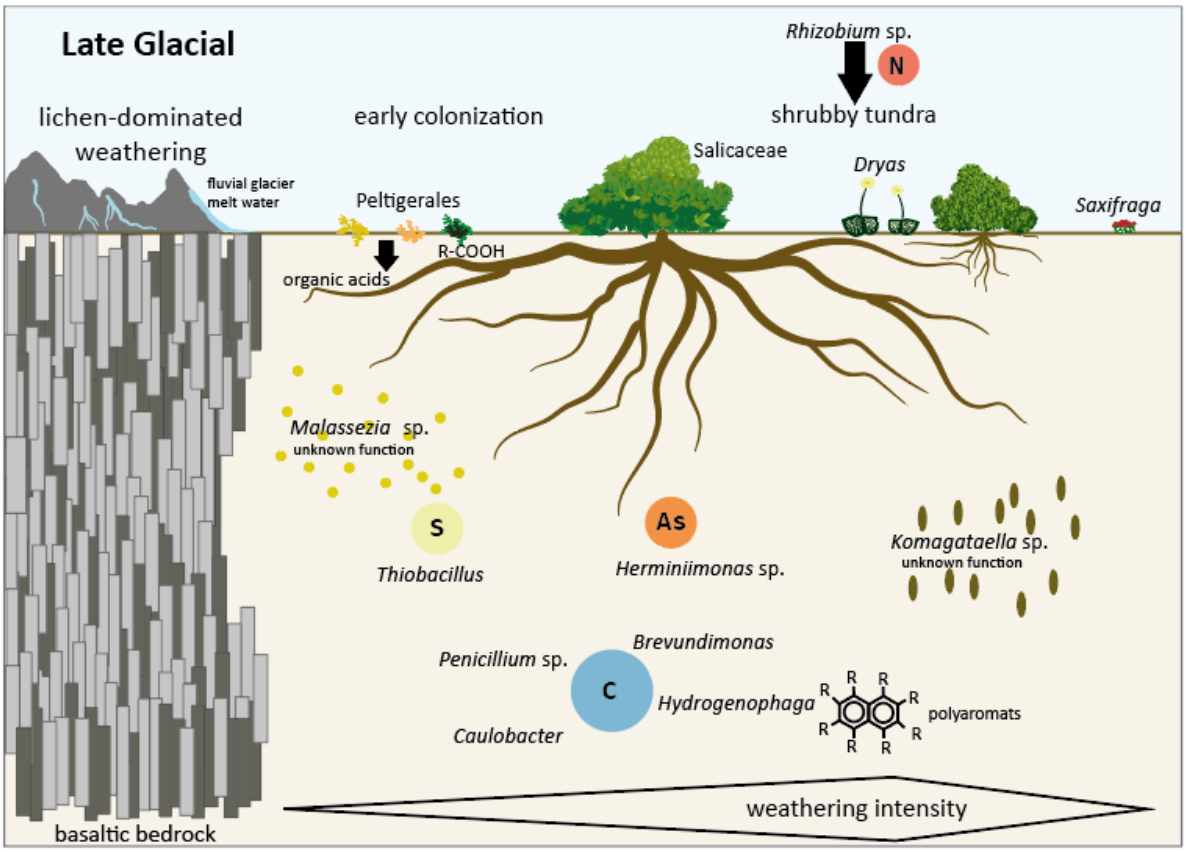


Fig. 4: Ecosystem and soil dynamic changes over time. Upper part: development and changes in the ecosystem over time during the Late Glacial, when glacial melt-water flow and lichen cover lead to strong weathering of the basaltic bedrock. After early colonisation, the vegetation subsequently developed towards shrubby tundra. Lower part: Carbon (C) cycling in the soil was high. Sulphur (S) cycling was high during early colonisation, while arsenic (AsI cycling was prominent until the onset of the Holocene. The size of the circles represents the importance of the respective element cycling process. All cyclers besides *Penicillium* (fungus) are representing bacterial taxa. Yeast taxa were highly abundant throughout the Late Glacial.

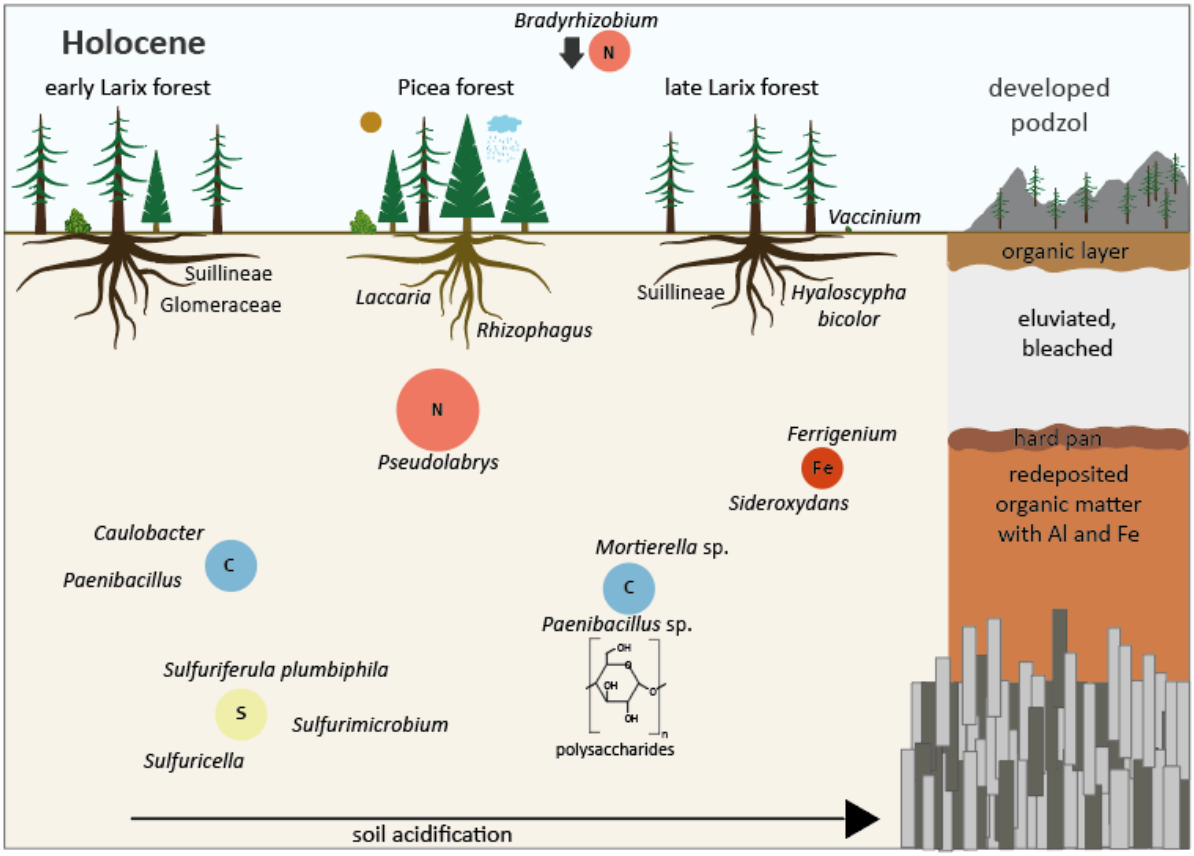


Fig. 5: Soil establishment during the Holocene, resulting in the formation of podzol. With the onset of the Holocene, Pinaceae invaded the area and an early *Larix* (larch) forest was established. During the thermal optimum in the Holocene, this early *Larix* forest was replaced by *Picea* (spruce), leading to an increase in iron (Fe) cycling in the soil and the start of podzol development. The changes in the soil properties enabled the re-invasion of *Larix* with *Vaccinium* taxa as a herbaceous soil cover. The nutrient cycling in the Holocene is dominated by nitrogen (N) cyclers in the soil. Additionally, carbon (C) cycling changed towards polysaccharide cycling and sulphur cycling diversified. In the late Holocene, iron cycling increased. All cyclers besides *Mortierella* (fungus) are representing bacterial taxa.

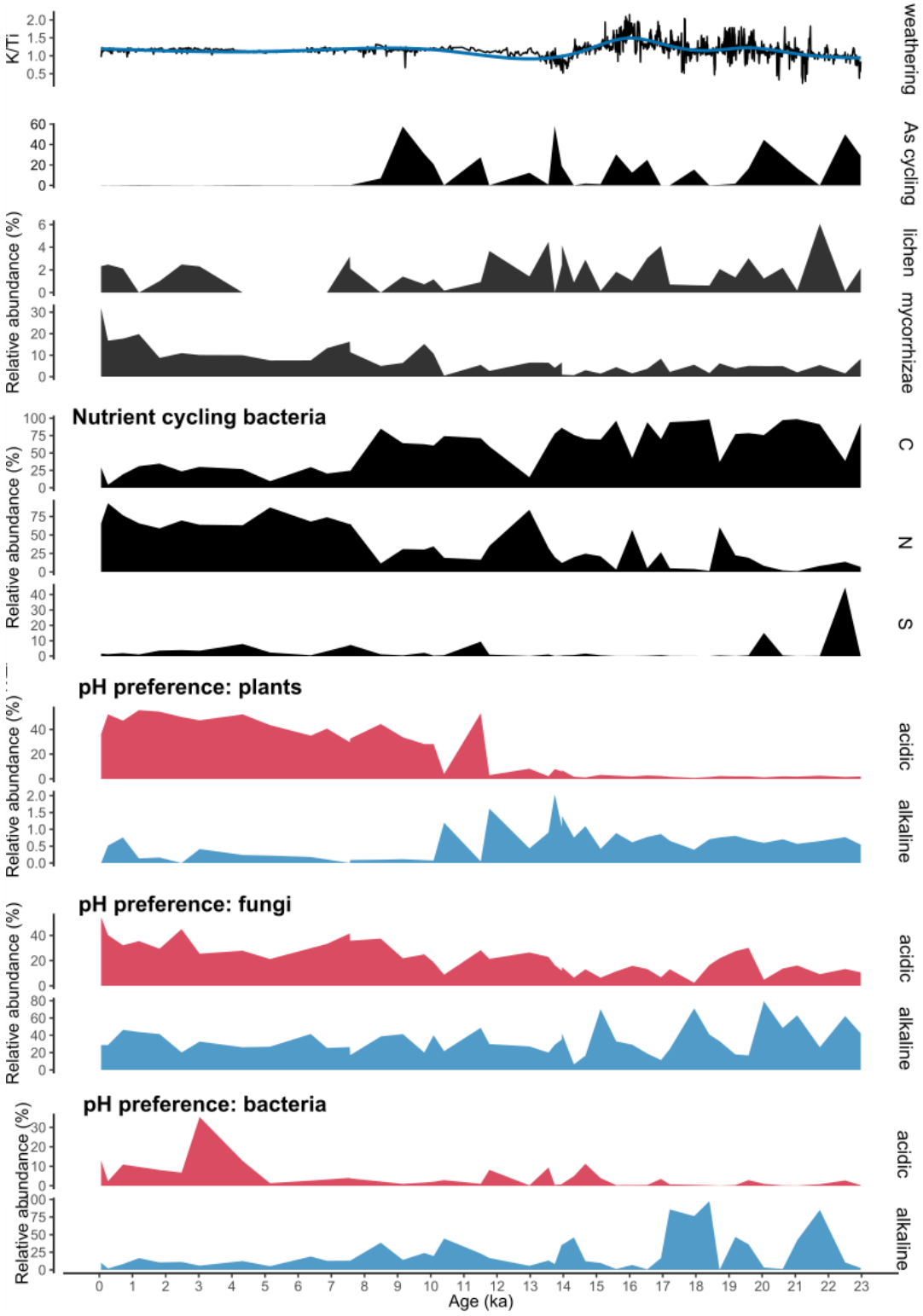


Fig. 6: Reconstruction of soil development from lake Lama sediment. General weathering is shown by the K/Ti ratio of the XRF data. Arsenic cycling bacteria are prominent in the samples from the Late Glacial up until 7.5 ka BP. A transition in relative abundance from lichen towards mycorrhizae is apparent throughout the core. We display a turn-over in nutrient cycling bacteria from carbon (C) dominated cycling in the Late Glacial to nitrogen (N) dominated cycling in the Holocene and show the sulphur (S) cycling alongside. The pH preferences (acidic/alkaline) of the taxa of the respective subsets show an acidification with increasing soil age and ongoing development.

5 Discussion and synthesis

The main goal of this thesis was to shed light onto long-term community dynamics of plants and soil microorganisms using lake sedimentary ancient DNA. In this work, I assessed the long-term establishment of rhizosphere communities in relation to changing vegetation cover and identified distinct community trends for the (Late) Glacial and Holocene. Here, I discuss the included manuscripts forming the base of this thesis in this context.

5.1 Long-term rhizosphere establishment in tundra and taiga areas

During the Late Glacial, northern-central Siberia, central Yakutia, and the Tibetan Plateau were covered by shrubby tundra, while with the onset of the Holocene and local warm periods, mainly larch-dominated coniferous forests with spruce and herbs invaded the areas (Manuscripts II, III, IV). I highlight in my thesis the long-term dependency of boreal forest establishment on mycorrhizal associations. I also show turnover from initial vegetation with lichen cover after glacier-retreat towards boreal forests with strong mycorrhizae-domination (Manuscripts II, III). Additionally, I show that changes in the element cycling reflect the needs of higher trees for diversified nitrogen cycling in nutrient poor habitats (Manuscript III). Whereas for fungus-plant covariation, the major ecological trends were comparable amongst sites, specific local signals were detected on taxon level, reflecting the dependency of microbiome establishment on soil properties (Manuscripts II, IV).

5.1.1 SedaDNA as a proxy for soil microbiome

Until recently, the use of lake sedaDNA for the reconstruction of past ecosystem dynamics has mainly focused on plant communities, diatoms or mammals (e.g. Pansu et al., 2015; Stivrins et al., 2018; Parducci et al., 2019), while the analysis of soil communities is a relatively new approach (Talas et al., 2021). This is partially because databases for fungi are still far from being comprehensive, but also because short metabarcoding primer pairs for assessing the broad fungal community have been lacking. To close this gap, we first reassessed existing primer pairs and established a new, relatively short primer combination targeting the ITS1 region (Manuscript I). Then, I applied this primer pair on a variety of lake sediment cores for target specific metabarcoding (Manuscript II), demonstrating that fungal ecological groups show distinct trends in their abundance under long-term climatic change. For comparison, I assessed fungal terrestrial communities as well using shotgun sequencing (Manuscripts III, IV), showing overall similar results with the two approaches, though also revealing bias in the metabarcoding. With the shotgun sequenced data, also bacterial communities were included in the analysis (Manuscripts III, IV), demonstrating the possibility to reconstruct the diverse soil microbiome.

5.1.1.1 Fungal DNA metabarcoding

We underlined the possibility to trace fungal community dynamics using ITS metabarcoding on lake seaDNA with establishing and applying a primer pair targeting the ITS1 region. In fungal metabarcoding, the ITS region spanning ITS1 and ITS2 is mainly being targeted. Though, due to varying lengths of the ITS between different species and a subsequent bias in PCR amplification towards shorter sequences, an accurate community reconstruction is limited (Baldrian, 2019). Commonly, ITS2 is used for these studies as it is more consistent than ITS1 (Walters et al., 2015), while I amplified ITS1 as it yielded the most promising primer pair for palaeo-reconstructions with short sequence lengths as well as high species recovery (Manuscripts I, II). A comparative study using ITS1 vs. ITS2, targeting lichen diversity (Fernández-Mendoza et al., 2017; Banchi et al., 2018), showed differing compositional patterns between the barcodes. This underlines the bias in the chosen barcode, emphasising the need to target ITS1 as well as ITS2 when assessing a broad ecological diversity of fungi. As an alternative to the ITS region, the 16S region which is commonly used for bacterial metabarcoding (Stackebrandt and Goebel, 1994), can also be applied for fungal metabarcoding, though it has less hypervariable regions in fungi (Schoch et al., 2012). The abundance of 16S in the bacterial as well as fungal ribosome renders it a great asset to the shotgun sequencing data of Manuscript III. While not only strengthening the bacterial shotgun data, it could potentially verify the fungal community composition as derived form ITS metabarcoding in Manuscript II. Furthermore, the protein-coding RPB1 subunit of RNA Polymerase II has been proposed as a fungal barcode and has in combination with the ITS primer been suggested to yield high species recovery rates (McLaughlin et al., 2009; Schoch et al., 2012).This would yield it in combination with our established primer (Manuscript I) an additional great target region for ancient DNA metabarcoding. Nonetheless, each barcode shows slight differences in amplified species (Schoch et al., 2012). Due to time and money constraints during the work on this thesis, only the primer pair published in Manuscript I could be applied on a large scale. Though in future, a metabarcoding approach comparing adapted primers for ITS1 as well as ITS2 or RPB1 for fungal communities deriving from lake sedaDNA would be a great asset.

5.1.1.2 Targeting soil communities with shotgun sequencing

Besides the recovery of fungi and their community compositional trends, I could also show for the first time the possibility to reconstruct terrestrial bacterial communities from lake sedaDNA and assign them to their respective roles in the rhizosphere (Manuscripts III, IV). As plant-growth promotion in the rhizosphere is not solely deriving by fungi, assessing the bacterial community alongside the fungal changes is of great importance for understanding ecological dependencies. So far, rhizobial bacteria have been assessed by amplifying modern rhizosphere samples of tree individuals (Praeg and Illmer,

2020), or shotgun sequencing of rhizosphere soil (Babalola et al., 2020). For such studies on modern plant individuals, rhizosphere analyses are rather straight-forward, as a comparison to bulk soil as a control is possible, enabling the assessment of direct biotic dependencies. Those direct relations between specific plant species and their rhizobiome are more complicated to assess from lake sediment as a correlation can only be tested using statistics, while no real associations are observed (Manuscript IV). Addressing bacterial changes throughout rhizospheres alongside, for example, major vegetational changes is of advantage as it helps identifying important drivers for general plant establishment or treeline migration in nutrient poor habitats. Nonetheless, the assessment of a range of modern rhizospheres in the target area would be beneficial for the comparison to the reconstructed sedaDNA rhizosphere dependencies. This is necessary as the rhizobiome is highly dependent on the surrounding environment and influenced by factors such as bedrock (Vieira et al., 2020) or altitude (Li et al., 2014; Praeg et al., 2019). In the case of this thesis, this could entail sequencing the rhizosphere of several dominant tree, shrub, herb, and grass individuals in the vicinity of the lakes, comparing it to sequenced bulk samples, and relating it then to the data of the manuscripts II, III, and IV respectively.

5.1.1.3 Comparison between metabarcoding and shotgun sequencing for the soil microbiome

I applied metabarcoding and shotgun sequencing to recover ancient fungal communities in a variety of lake sediment cores (Manuscripts I, II, III, IV), showing that major trends in fungi composition are comparable between metabarcoding and shotgun sequencing (Manuscripts II, III, IV). Nonetheless, I discovered slight differences between the methods when comparing trends in fungal ecology: I showed an increasing relative abundance for mycorrhizal fungi with warming and ongoing soil development using shotgun sequencing (Manuscript III), while the fungal metabarcoding data of the same lake exhibited no clear trends (Manuscripts I, II). I observed differences between the two approaches also for plant-parasitic fungi, where for metabarcoding a high covariation with woody plant taxa was obtained (Manuscript II). In contrast, in the shotgun data, not many correlations between (woody) plants and parasitic fungi were detected (Manuscript IV). Arbuscular mycorrhizal fungi form part of the phylum of Glomeromycota which shows a high intraspecific variation in the ITS region (Lloyd-Macgilp et al., 1996) and therefore also in barcoding (Schoch et al., 2012). Thus, this indicates a bias in the metabarcoding amplification towards species with an overall shorter ITS, explaining the differences between the shotgun and metabarcoding results, while indicating that we might missed some species in the barcoding approach. Also, varying copy numbers of the ITS per fungal genome were described, ranging up to > 100 per genome, leading to further bias in the metabarcoding data (Raidl et al., 2005; Debode et al., 2009; Baldrian, 2019). Additionally, the reconstruction of the lichen community differed prominently between the approaches: while the shotgun sequencing

showed a broad abundance of lichens in the Pleistocene with a strong decline until the onset of the Holocene (Manuscript III), the metabarcoding did not reveal pronounced trends in lichen abundance (Manuscript II). Therefore, the combination of metabarcoding multiple target regions for one organism group, such as ITS, RPB1, as well as 16S for fungi and bacteria, and additional shotgun sequencing will give the best overview of the soil microbiome. Overall, my data suggest that the application of barcoding multiple target sites, coupled with whole genome sequencing methods is an appropriate choice when reconstructing a broad fungal community. When only applying metabarcoding, the primer selection should be re-evaluated using specific barcodes for selected fungal groups, for example when focusing on arbuscular mycorrhizae reconstruction, as stated by Lee et al. (2008).

For validation of the shotgun data for the bacterial communities (Manuscripts III, IV), a metabarcoding approach such as for fungi (Manuscripts I, II) will be an asset. As an initial step, the barcoding primers for bacteria need to be re-evaluated as well to adapt them for the amplification of ancient DNA. Commonly used primer pairs for bacterial metabarcoding on aquatic sediment target regions of >300 nucleotides (Stoeck et al., 2010; Caporaso et al., 2012; Leray et al., 2013; Fonseca et al., 2022), while an established primer pair for targeting the bacterial rhizosphere amplifies a region of 580 nt (Baker et al., 2003; Lasa et al., 2019). This renders those primers not suitable for ancient DNA due to too long amplicon sizes. A re-evaluation of the primer pairs for the amplification of bacterial 16S, comparable to the reassessment of fungal primers as in Manuscript I, and their adaptation to the amplification of shorter fragments would be an asset. The primer pair used by Lasa et al. (2019) includes three variable gene regions. For example, using a different reverse primer or varying the forward and reverse primer combinations for the amplification of a shorter RNA region with only one variable gene region could be promising. A profound characterisation of ancient rhizosphere bacterial communities will enable further understanding of the impact of long-term global warming on terrestrial nutrient cycles.

5.1.2 Fungi-vegetation interaction changes over time

Plant-fungus interactions are complex and thus so far only understood on spatial gradients within a country (Silva et al., 2014) or global comparisons (Bahram et al., 2013), for single individuals (Praeg and Illmer, 2020), or on time-scales covering a couple of years (Nuñez et al., 2009; Liu et al., 2019; Weber et al., 2019). As climate change is a slow, but steady process, especially when assessing the Last Glacial-Holocene transition, long-term monitoring of sedaDNA datasets are a great tool for the understanding long-term dependencies.

In my thesis, I demonstrate the possibility to reconstruct complex fungal-plant interactions from lake sediment and highlight the long-term dependency of Pinaceae establishment on mycorrhizal fungi

including Suillineae, Inocybaceae, and *Hyaloscypha* species (Manuscripts II, III). So far, such dependency was shown only on a short temporal gradient of several years (Nuñez et al., 2009). It is known that some fungal species show a high host fidelity towards specific plant taxa, such as *Suillus* species being associates for Pinaceae (Kretzer et al., 1996; Nguyen et al., 2016). However, it has remained unknown if mycorrhizal associations were dynamic during ecosystem establishment and if they did adapt to changing environmental conditions. In my thesis, I provide evidence that such mycorrhizal associations are dynamic throughout time and dependent on the stage of soil development. Shotgun sequencing revealed that the mycorrhizal community composition of an early invaded larch forest is different to the mycorrhizal communities of the reinvaded late larch forest multiple thousand years later (Manuscript III). For the invaded larch forest at the site, the herbaceous understorey vegetation of the forest changed and Ericaceae such as *Vaccinium* species became more abundant in the late larch forest. In British woodlands, an influence of arbuscular mycorrhiza on herbaceous understorey vegetation was demonstrated (Guy et al., 2022). The presented data show that mycorrhizal plant-fungus interactions are complex, while indicating also a vice-versa impact of the herbaceous taxa on fungi. Besides, I also demonstrate that mycorrhizal interactions are differing in the same plant genera across geographical locations. This indicates a strong adaptation of mycorrhizae to the local environment and as such to soil properties independent from the soil development stage (Manuscript IV). These include environmental parameters such as the bedrock, which varied in the assessed datasets from basalt over sandstone to granite, showing *Suillus* as the known Pinaceae associate only as a strong correlated mycorrhizal fungus for *Larix* in the basaltic location. In Finnish boreal forests, a major impact of geographical location on the establishment of fungal communities has been demonstrated (Qu et al., 2021), underlining that the large spatial gradient covered by my data and therefore the unique environmental factors such as elevation, precipitation, or local soil properties are shaping the rhizosphere to a great extent. My data show varying mycorrhizal associations depending on the age of the forest and the understorey vegetation as well as the impact of different geographical locations (Manuscripts III, IV). As an example, the data present increasing relative abundance of Suillineae and *Hyaloscypha* with the re-invaded larch forest with *Vaccinium* understorey in the late Holocene (Manuscript III). This underlines the complexity of mycorrhizal associations, proving that they are far from being fully understood and thus parameters shaping these associations need to be further assessed.

I demonstrate in my thesis that the relative abundance of saprotrophic fungal communities is shifting with the presence or absence of woody taxa and that their community composition is alternating depending on the respective abundant plant taxa (Manuscripts II, III, IV). Mainly, this includes a shift from *Penicillium* domination with tundra vegetation to increasing *Mortierella* abundance in boreal forests, indicating a high selective pressure of plants towards saprotrophic fungi. A positive feedback

of herbaceous understorey vegetation on saprotrophic fungal biomass as well as on the fungi:bacteria ratio has already been demonstrated for deciduous forests (Stefanowicz et al., 2022). In boreal forests, protease activity was strongly correlated only with soil fungal communities but not with bacterial communities (Sakurai et al., 2007; Vuong et al., 2020). As the litter of boreal taxa compared to herbaceous taxa or shrubs is differing in its lignin content (Rahman et al., 2013; Peng et al., 2022), the enzymes required for breakdown of the sugar or polymer molecules are also shifting, explaining the different compositional patterns of saprotrophs. A recent study showed differences in cell wall structures in herbaceous and woody plants due to differing hemicellulose, cellulose, and lignin interaction networks and subsequent host-specific parasitic fungal plant-cell-wall degrading enzymes (Dou et al., 2021), being in good agreement with my data (Manuscripts II, III, IV). This concludes the occurrence of host-specific saprotrophic fungi, potentially evolved due to differences in the cell wall structures. The relative increases in saprotrophic fungi occurrence with woody taxa might have also emerged from increased abundance in mycorrhizal fungi, as litter decomposition at most sites is decreased with the presence of ectomycorrhizal fungi, leading to enhanced carbon storage in forest soils, termed the "Gadgil effect" (Fernandez and Kennedy, 2016). These interspecies competitions have yet not been fully understood using modern data sets. Therefore, it remains unclear whether also arbuscular mycorrhizae are enhancing the effect.

For plant-parasitic fungi, I obtained contradictory results with metabarcoding (lakes: CH12, Lama, Bolshoe Toko, Kyutyunda, Levinson Lessing) and shotgun sequencing (lakes: Lama, Bolshoe Toko, Ximencuo) (Manuscripts II, IV). Metabarcoding indicates that parasitic fungal communities are strongly co-occurring with woody plant taxa and following their dynamics (Manuscript II), indicating that not only a changing climate is causing their ecology trend. In contrast, shotgun sequencing of three lake sediments and the assessment of correlations between specific plant taxa and their rhizobiome did not reveal high correlations for parasitic fungi (Manuscript IV). So far, increasing plant parasitic fungi richness with warming (Geml et al., 2015) as well as a great host-specificity between plants and fungal parasites (Põlme et al., 2018) were shown. Nonetheless, the generally warming-related changes in plant-parasite interplay are not fully understood (Burdon and Zhan, 2020). The contradicting results of my manuscripts indicate that the increase in parasitic fungi alongside woody taxa is potentially rather warming-induced, than being dependent on the specific boreal plant taxa. This assumption is supported by Sepp et al. (2021) who assessed the impact of woody patches in grassland areas on fungal community composition, showing that mainly symbiotic fungal communities are changing between the different vegetation cover. Nonetheless, as demonstrated by my results, more research on ecosystem wide plant-parasitic interaction under different environmental conditions is needed to delimit main causes of compositional shifts.

On an ecology level, I also highlight the habitat loss of yeast, including *Komagataella* and *Malassezia* species, due to natural long-term warming. I revealed a decline in their relative abundance with warming which subsequently impacts the whole ecosystem (Manuscripts II, III). This strengthens a study showing declining yeast abundance with experimental warming on a relatively small timeframe of 8 years (Treseder et al., 2016), while proving this scenario also for natural warming. The ecological roles of yeast in soil are vast: while some species possess saprotrophic properties (Mašínová et al., 2018), others are plant parasites (Hernández-Fernández et al., 2021), produce biotoxins (Santos et al., 2004) or act as plant-growth promoters (El-Tarabily and Sivasithamparam, 2006). This indicates that yeast may play a role in the establishment and survival of plants. The decline in yeast abundance with warming is potentially also relatable to the relative increase and taxon shift of saprotrophs, pointing out that less carbon decaying yeast are needed with ongoing warming. Also, the abundance of some yeast was found to impact the establishment of mycorrhizal fungi by reducing their overall colonization, resulting in less plant-defence related genes (Mestre et al., 2022). This indicates that there is a strong interplay between yeast and further fungal taxa, suggesting that the yeast decline with warming (Manuscripts II, III) also favours the subsequent establishment of mycorrhizal taxa. By underlining the complexity of yeast functional ecology and their tremendous impact on terrestrial ecosystems, I show the need for further research on their distinct functionalities and biotic dependencies.

A large remaining problem is that the modern rhizosphere of many plants is not fully understood due to its complexity. Assessing long-term changes and subsequently relating covariations would be easier if the modern analogue is better understood. Therefore, it will be of great advantage assessing selected species in the catchment areas and sequencing their rhizobiomes. Thus, a relation to changing rhizosphere related taxa in the past will be more straight-forward and more promising.

5.1.3 Soil development on a temporal gradient

In my thesis, I was able to reconstruct post-glacial development of podzol from the sequencing data of lake Lama in northern central Siberia. Alongside, I showed that soil development is to a large extent driven by the vegetation cover, while time passed by since initial deglaciation is of neglectable importance.

Assessing the trends in bacterial element cycling alongside the dynamics of plant community changes, I grouped the taxa based on their preferred soil pH. The results directly pointed to a strong trend towards acidification in the Holocene, which I proved also for the fungi (Manuscript III). I found further indication for ongoing pedogenesis with strong basaltic weathering dominated by lichen after

deglaciation (Manuscript III) and a high abundance of arsenic cyclers, mainly *Herminiimonas* species. In the Holocene, fungal weathering changed to mycorrhizae-domination (Manuscript III) and the soil stabilised. So far, lake sedaDNA has been used mainly to assess terrestrial or marine ecosystem dynamics. In Manuscript III, I used sedimentary ancient DNA for the first time to reconstruct large-scale geological processes in a catchment area, highlighting new potentials for sedaDNA research.

I could show shifts from polyaromatic to polysaccharide carbon cycling for lake Lama with an additional increase in saprotrophic fungi alongside the invasion of boreal forests, indicating a strong connection between bacterial and fungal carbon cycling (Manuscripts II, III). A large contribution to soil organic carbon derives from plants via rhizodeposition through the roots or through litter decomposition (Gougoulias et al., 2014). Rhizodeposition mainly consists of simple molecules such as sugars, amino acids, or secondary metabolites which can be rapidly respired and degraded (Bais et al., 2006). In contrast, more complex compounds including cellulose or the polyaromatic lignin need prior depolymerisation by extracellular enzymes before microbial degradation (Baldrian and Valášková, 2008; Wallenstein and Weintraub, 2008). In line with this, my data indicate that saprotrophic fungi are more specialised in the degradation of the rhizodeposits of taiga than bacterial carbon cyclers. This implies a niche separation between fungal and bacterial carbon cycling dependent on the ecosystem type and plant cover with the need of highly specific enzymes for the cleavage of the respective plant litter.

For the bacterial community of basaltic bedrock, I was able to reconstruct a turnover in nutrient cycling with changing vegetation cover from carbon-dominated cycling in shrubby tundra towards nitrogen-dominated cycling in boreal forests. The major turnover occurred with the invasion of dark taiga, including a diversification of nitrogen pathways from *Rhizobium* dominated nitrogen fixation to nitrogen cycling by *Pseudolabrys* and *Bradyrhizobium* (Manuscript III). For sites with granitic or sandstone bedrock, nitrogen cycling was continuously at a high level (Manuscript IV). For the basaltic bedrock, that might be an indication for higher foliage retention in dark coniferous forests compared to deciduous coniferous forests or tundra and the reduced turnover time with trees (Manuscript III). However, the comparison to the other sites does not fully support this interpretation being a sole root-cause (Manuscript IV). The general increase in nitrogen cycling for the basaltic site is also relatable to the changing soil structure with the *Picea* invasion (Manuscript III). Surprisingly, these major changes in nutrient cycling throughout the core were not detected alongside pedogenesis or in relation to vegetation changes for the other sites with granitic and sandstone bedrock (Manuscript IV). A major reason here could be that the sites are not only geologically different, but also vary in their altitudes, ranging from 50 to 4000 m above sea level. Additionally, my thesis indicates while demonstrating the complexity of nutrient cycles in soil, that pedogenesis and the resulting impact on the nutrient cycles

is highly bedrock specific. This specificity is being relatable to the varying content in silicate in the rock types, different pore sizes in the bedrock and the emerging soil, and therefore varying aeration and water drainage (Voroney, 2007).

I could show that podzolization is reconstructable from a basaltic bedrock using lake sedaDNA: I detected an acidification of the soil with ongoing soil age starting from the onset of the Holocene by assessing the pH preferences of plants, fungi, and bacteria, alongside the establishment of boreal forest in the area (Manuscript III). Additionally, I reconstructed increasing iron cycling in the late Holocene after spruce invasion (Manuscript III). The release of iron and aluminium ions from rocks which is induced by organic acids, resulting in complexation with organic matter, and the subsequent leaching of the complexes from the upper soil horizons to the subsoil is characteristic for podzol establishment (Lundström et al., 2000; Sauer et al., 2007). The duration of podzolization is so far unknown and varying time spans have been indicated: Various studies specified the development of ferric podzol and the differentiation of horizons to take several hundred years (e.g. Singleton and Lavkulich, 1987; Lundström et al., 2000). In contrast, Barrett and Schaetzl (1992) indicate large time spans and estimate that the development of a spodic horizon near lake Michigan took between 4 – 10 kyr. In accordance with this study, my data underline the extremely long period of time from initial weathering to podzol development. I showed a drastically increased and ongoing iron cycling over several thousand years in the late Holocene, potentially marking the beginning of podzolization and indicating a time-span of multiple thousand years for the development. Nonetheless, the timing of the respective processes remains unknown, while the discusses contrasting time-scales indicate a site-specific duration of pedogenesis which is highly dependent on local biotic and abiotic factors.

I demonstrate the possibility to trace weathering of bedrock alongside soil development mainly using bacterial and fungal compositional changes (Manuscript III). Nonetheless, the data indicate that these trends are highly catchment as well as bedrock specific when compared to shotgun data of other sites (Manuscripts III, IV). When assessing soil processes, next to the community composition, it would be highly interesting to also measure metabolic rates and link them to the microorganisms (Baldrian, 2019). As the metabolome is a rather instable value (Gil et al., 2015), proteomics might be a more promising alternative for palaeo studies. The metabolome and the proteome are highly intertwined such that proteins are driving the cellular metabolisms through enzyme activity, while metabolites as signalling molecules influence protein expression (Kresnowati et al., 2006; Bradley et al., 2009; Buescher et al., 2012). So far, ancient proteomics were mainly done on tissue isolates (Welker et al., 2015, 2016; Chen et al., 2019), while data on lake sediments or generally tracing temporal community changes are lacking. Further application on lake sediment also for the analysis of rhizosphere processes seems promising: an increasing fungal protein richness in forests compared to shrublands has recently

been demonstrated (Fernandes et al., 2022), suggesting that a profound comparison to the ecosystem trends derived from the metabarcoding (Manuscripts I, II) and the shotgun data (Manuscripts III, IV) is possible.

5.2 Conclusion and future perspectives

The thesis objective was to analyse long-term plant-microbe interactions which drive ecosystem and soil establishment in nutrient poor permafrost landscapes. This could be fulfilled while showing the relevance of cross discipline research in such highly complex systems.

Little is known about long-term plant dependencies towards soil microorganisms and on the dynamic establishment of rhizobial communities. Most studies assessing rhizobial dynamics focus on up to only several years due to complex sampling and experimental set-up, though initial rhizosphere as well as plant establishment itself are processes which can themselves take several years. However, understanding the rhizosphere leads to a better understanding of treeline dynamics. With the base of this thesis, an assessment of future treeline dynamics is facilitated by unfolding microbial associations needed for tree establishment.

To validate the way of reconstructing pedogenesis, it will be necessary to apply this on further sediment cores. If applicable on different bedrock and catchment types, the method can be widely used, helping to understand global soil development and turnover. Soil development forms the base for the survival of modern society in terms of agriculture and livestock. Intensified use of resources and overproduction in dry areas do not allow soils to recover and deserts arise. A further desertification leads to an increase of exposed bare soils and higher absorption of solar radiation, while hindered plant growth reduces the capacity to store excess carbon. As such, a better knowledge of long-term pedogenesis will facilitate dealing with future declining resources.

In a warming world, it is inevitable to counteract excess carbon dioxide emission with capturing and storage to meet the goals of the Paris agreement, limiting anthropogenic warming to 1.5 - 2 °C. Carbon capturing by application of basalt grains on soil and its subsequent weathering is an emerging strategy, resulting in mineral carbonation. Our data provide a great basis for future improvement of the technique as we highlight natural, molecular mechanisms alongside the rock weathering and its impact on the development of ecosystems. Further intensified research on the molecular processes behind basalt weathering will be of great advatage when globally extending its application on large scales.

6 References

Ahkami, A.H., Allen White, R., Handakumbura, P.P., Jansson, C., 2017. Rhizosphere engineering: enhancing sustainable plant ecosystem productivity. *Rhizosphere*, New understanding of rhizosphere processes enabled by advances in molecular and spatially resolved techniques **3**, 233–243. https://doi.org/10.1016/j.rhisph.2017.04.012

Alsos, I.G., Lammers, Y., Yoccoz, N.G., *et al*., 2018. Plant DNA metabarcoding of lake sediments: How does it represent the contemporary vegetation. *PLOS ONE* **13**, e0195403. https://doi.org/10.1371/journal.pone.0195403

Alsos, I.G., Sjögren, P., Edwards, M.E., *et al*., 2016. Sedimentary ancient DNA from lake Skartjørna, Svalbard: Assessing the resilience of arctic flora to Holocene climate change. *The Holocene* **26**, 627–642. https://doi.org/10.1177/0959683615612563

Andersen, K., Bird, K.L., Rasmussen, M., *et a*l., 2012. Meta-barcoding of 'dirt' DNA from soil reflects vertebrate biodiversity. *Mol. Ecol.* **21**, 1966–1979. https://doi.org/10.1111/j.1365-294X.2011.05261.x

Andree, M., Oeschger, H., Siegenthaler, U., *et al*., 1986. 14C dating of plant macrofossils in lake sediment. *Radiocarbon* **28**, 411–416. https://doi.org/10.1017/S0033822200007529

Babalola, O.O., Alawiye, T.T., Lopez, C.R., Ayangbenro, A.S., 2020. Shotgun metagenomic sequencing data of sunflower rhizosphere microbial community in South Africa. *Data Brief* **31**, 105831. https://doi.org/10.1016/j.dib.2020.105831

Bahram, M., Kõljalg, U., Courty, P.-E., *et al*., 2013. The distance decay of similarity in communities of ectomycorrhizal fungi in different ecosystems and scales. *J. Ecol*. **101**, 1335–1344. https://doi.org/10.1111/1365-2745.12120

Bais, H.P., Weir, T.L., Perry, L.G., Gilroy, S., Vivanco, J.M., 2006. The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu. Rev. Plant Biol.* **57**, 233–266. https://doi.org/10.1146/annurev.arplant.57.032905.105159

Baker, G.C., Smith, J.J., Cowan, D.A., 2003. Review and re-analysis of domain-specific 16S primers. *J. Microbiol. Methods* **55**, 541–555. https://doi.org/10.1016/j.mimet.2003.08.009

Baldrian, P., 2019. The known and the unknown in soil microbial ecology. *FEMS Microbiol. Ecol.* **95**, fiz005. https://doi.org/10.1093/femsec/fiz005

Baldrian, P., Valášková, V., 2008. Degradation of cellulose by basidiomycetous fungi. *FEMS Microbiol. Rev.* **32**, 501–521. https://doi.org/10.1111/j.1574-6976.2008.00106.x

Banchi, E., Stankovic, D., Fernández-Mendoza, F., *et al.*, 2018. ITS2 metabarcoding analysis complements lichen mycobiome diversity data. *Mycol. Prog.* **17**, 1049–1066. https://doi.org/10.1007/s11557-018-1415-4

Barrett, L.R., Schaetzl, R.J., 1992. An examination of podzolization near lake Michigan using chronofunctions. *Can. J. Soil Sci.* **72**, 527–541. https://doi.org/10.4141/cjss92-044

Beeck, M.O.D., Lievens, B., Busschaert, P., *et al.*, 2014. Comparison and validation of some ITS primer pairs useful for fungal metabarcoding studies. *PLOS ONE* **9**, e97629. https://doi.org/10.1371/journal.pone.0097629

Bellemain, E., Davey, M.L., Kauserud, H., *et al.*, 2013. Fungal palaeodiversity revealed using high-throughput metabarcoding of ancient DNA from arctic permafrost. *Environ. Microbiol*. **15**, 1176–1189. https://doi.org/10.1111/1462-2920.12020

Bigelow, N.H., Brubaker, L.B., Edwards, M.E., *et al.* 2003. Climate change and Arctic ecosystems: 1. Vegetation changes north of 55°N between the last glacial maximum, mid-Holocene, and present. *Journal of Geophysical Research*: Atmospheres **108,** 2156-2202. https://doi.org/10.1029/2002JD002558 (accessed 5.9.23).

Bradley, P.H., Brauer, M.J., Rabinowitz, J.D., Troyanskaya, O.G., 2009. Coordinated concentration changes of transcripts and metabolites in *Saccharomyces cerevisiae*. *PLOS Comput. Biol.* **5**, e1000270. https://doi.org/10.1371/journal.pcbi.1000270

Bradshaw, C.J.A., Warkentin, I.G., 2015. Global estimates of boreal forest carbon stocks and flux. *Glob. Planet. Change* **128**, 24–30. https://doi.org/10.1016/j.gloplacha.2015.02.004

Brandt, J.P., Flannigan, M.D., Maynard, D.G., *et al.*, 2013. An introduction to Canada's boreal zone: ecosystem processes, health, sustainability, and environmental issues. *Environ. Rev.* **21**, 207–226. https://doi.org/10.1139/er-2013-0040

Briggs, A.W., Stenzel, U., Johnson, P.L.F.,*et al*., 2007. Patterns of damage in genomic DNA sequences from a Neandertal. P*roc. Natl. Acad. Sci.* **104**, 14616–14621. https://doi.org/10.1073/pnas.0704665104

Brown, J., Romanovsky, V.E., 2008. Report from the International Permafrost Association: state of permafrost in the first decade of the 21st century. *Permafr. Periglac. Process*. **19**, 255–260. https://doi.org/10.1002/ppp.618

Brundrett, M.C., 2009. Mycorrhizal associations and other means of nutrition of vascular plants: understanding the global diversity of host plants by resolving conflicting information and developing reliable means of diagnosis. *Plant Soil* **320**, 37–77. https://doi.org/10.1007/s11104-008-9877-9

Buescher, J.M., Liebermeister, W., Jules, M., *et al.* 2012. Global Network Reorganization During Dynamic Adaptations of Bacillus subtilis Metabolism. *Science* **335**, 1099–1103. https://doi.org/10.1126/science.1206871

Burdon, J.J., Zhan, J., 2020. Climate change and disease in plant communities. *PLOS Biol.* **18**, e3000949. https://doi.org/10.1371/journal.pbio.3000949

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J.A., Smith, G., Knight, R., 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6**, 1621–1624. https://doi.org/10.1038/ismej.2012.8

Chapin, F.S., Sturm, M., Serreze, M.C., *et al.*, 2005. Role of land-surface changes in Arctic summer warming. *Science* **310**, 657–660. https://doi.org/10.1126/science.1117368

Chen, F., Welker, F., Shen, C.-C., *et al.*, 2019. A late Middle Pleistocene Denisovan mandible from the Tibetan Plateau. *Nature* **569**, 409–412. https://doi.org/10.1038/s41586-019-1139-x

Courtin, J., Perfumo, A., Andreev, A.A., *et al*., 2022. Pleistocene glacial and interglacial ecosystems inferred from ancient DNA analyses of permafrost sediments from Batagay megaslump, East Siberia. *Environ. DNA* **4**, 1265–1283. https://doi.org/10.1002/edn3.336

Debode, J., Van Hemelrijck, W., Baeyen, S., *et al*., 2009. Quantitative detection and monitoring of *Colletotrichum acutatum* in strawberry leaves using real-time PCR. *Plant Pathol*. **58**, 504–514. https://doi.org/10.1111/j.1365-3059.2008.01987.x

Delgado-Baquerizo, M., Reich, P.B., Bardgett, R.D., *et al*., 2020. The influence of soil age on ecosystem structure and function across biomes. *Nat. Commun.* **11**, 4721. https://doi.org/10.1038/s41467-020-18451-3

DeMarco, J., Mack, M.C., Bret-Harte, M.S., *et al.*, 2014. Long-term experimental warming and nutrient additions increase productivity in tall deciduous shrub tundra. *Ecosphere* **5**, art72. https://doi.org/10.1890/ES13-00281.1

Dou, Y., Yang, Y., Mund, N.K., *et al*., 2021. Comparative analysis of herbaceous and woody cell wall digestibility by pathogenic fungi. *Molecules* **26**, 7220. https://doi.org/10.3390/molecules26237220

Edwards, M.E., Brubaker, L.B., Lozhkin, A.V., Anderson, P.M., 2005. Structurally novel biomes: A response to past warming in Beringia. *Ecology* **86**, 1696–1703. https://doi.org/10.1890/03-0787

Elmendorf, S.C., Henry, G.H.R., Hollister, R.D., *et al*., 2012. Plot-scale evidence of tundra vegetation change and links to recent summer warming. *Nat. Clim. Change* **2**, 453–457. https://doi.org/10.1038/nclimate1465

El-Tarabily, K.A., Sivasithamparam, K., 2006. Potential of yeasts as biocontrol agents of soil-borne fungal plant pathogens and as plant growth promoters. *Mycoscience* **47**, 25–35. https://doi.org/10.1007/S10267-005-0268-2

Epp, L.S., Boessenkool, S., Bellemain, E.P., *et al*., 2012. New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Mol. Ecol.* **21**, 1821–1833. https://doi.org/10.1111/j.1365-294X.2012.05537.x

Epp, L.S., Gussarova, G., Boessenkool, S., *et al.*, 2015. Lake sediment multi-taxon DNA from North Greenland records early post-glacial appearance of vascular plants and accurately tracks environmental changes. *Quat. Sci. Rev.* **117**, 152–163. https://doi.org/10.1016/j.quascirev.2015.03.027

Feibelman, T., Bayman, P., Cibula, W.G., 1994. Length variation in the internal transcribed spacer of ribosomal DNA in chanterelles. *Mycol. Res.* **98**, 614–618. https://doi.org/10.1016/S0953-7562(09)80407-3

Fernandes, M.L.P., Bastida, F., Jehmlich, N., *et al*., 2022. Functional soil mycobiome across ecosystems. *J. Proteomics* **252**, 104428. https://doi.org/10.1016/j.jprot.2021.104428

Fernandez, C.W., Kennedy, P.G., 2016. Revisiting the 'Gadgil effect': do interguild fungal interactions control carbon cycling in forest soils? *New Phytol.* **209**, 1382–1394. https://doi.org/10.1111/nph.13648

Fernández-Mendoza, F., Fleischhacker, A., Kopun, T., *et al*., 2017. ITS1 metabarcoding highlights low specificity of lichen mycobiomes at a local scale. *Mol. Ecol.* **26**, 4811–4830. https://doi.org/10.1111/mec.14244

Flannigan, M., Stocks, B., Turetsky, M., Wotton, M., 2009. Impacts of climate change on fire activity and fire management in the circumboreal forest. *Glob. Change Biol.* **15**, 549–560. https://doi.org/10.1111/j.1365-2486.2008.01660.x

Fonseca, V.G., Kirse, A., Giebner, H., *et al*., 2022. Metabarcoding the Antarctic Peninsula biodiversity using a multi-gene approach. *ISME Commun*. **2**, 1–11. https://doi.org/10.1038/s43705-022-00118-3

Forbes, B.C., Fauria, M.M., Zetterberg, P., 2010. Russian Arctic warming and 'greening' are closely tracked by tundra shrub willows. *Glob. Change Biol*. **16**, 1542–1554. https://doi.org/10.1111/j.1365-2486.2009.02047.x

Franzetti, A., Pittino, F., Gandolfi, I., *et al*., 2020. Early ecological succession patterns of bacterial, fungal and plant communities along a chronosequence in a recently deglaciated area of the Italian Alps. *FEMS Microbiol. Ecol.* **96**, fiaa165. https://doi.org/10.1093/femsec/fiaa165

Geml, J., Morgado, L.N., Semenova, T.A., *et al*., 2015. Long-term warming alters richness and composition of taxonomic and functional groups of arctic fungi. *FEMS Microbiol. Ecol*. **91**, fiv095. https://doi.org/10.1093/femsec/fiv095

Giguet-Covex, C., Ficetola, G.F., Walsh, K., *et al*., 2019. New insights on lake sediment DNA from the catchment: importance of taphonomic and analytical issues on the record quality. *Sci. Rep.* **9**, 14676. https://doi.org/10.1038/s41598-019-50339-1

Gil, A., Siegel, D., Permentier, H., *et al*., 2015. Stability of energy metabolites—An often overlooked issue in metabolomics studies: A review. *ELECTROPHORESIS* **36**, 2156–2169. https://doi.org/10.1002/elps.201500031

Ginolhac, A., Rasmussen, M., Gilbert, M., *et al.*, 2011. mapDamage: Testing for damage patterns in ancient DNA sequences. *Bioinforma. Oxf. Engl.* **27**, 2153–5. https://doi.org/10.1093/bioinformatics/btr347

Gougoulias, C., Clark, J.M., Shaw, L.J., 2014. The role of soil microbes in the global carbon cycle: tracking the below-ground microbial processing of plant-derived carbon for manipulating carbon dynamics in agricultural systems. *J. Sci. Food Agric.* **94**, 2362–2371. https://doi.org/10.1002/jsfa.6577

Grace, J., Berninger, F., Nagy, L., 2002. Impacts of climate change on the tree line. *Ann. Bot.* **90**, 537–544. https://doi.org/10.1093/aob/mcf222

Guy, P., Sibly, R., Smart, S.M., *et al*., 2022. Mycorrhizal type of woody plants influences understory species richness in British broadleaved woodlands. *New Phytol.* **235**, 2046–2053. https://doi.org/10.1111/nph.18274

Hack, H.R.G.K., 2020. Weathering, erosion, and susceptibility to weathering, in: Kanji, M., He, M., Ribeiro e Sousa, L. (Eds.), Soft Rock Mechanics and Engineering. *Springer International Publishing*, pp. 291–333. https://doi.org/10.1007/978-3-030-29477-9_11

Hart, S.A., Chen, H.Y.H., 2006. Understory vegetation dynamics of North American boreal forests. *Crit. Rev. Plant Sci.* **25**, 381–397. https://doi.org/10.1080/07352680600819286

Hernández-Fernández, M., Cordero-Bueso, G., Ruiz-Muñoz, M., Cantoral, J.M., 2021. Culturable yeasts as biofertilizers and biopesticides for a sustainable agriculture: a comprehensive review. *Plants* **10**, 822. https://doi.org/10.3390/plants10050822

Herzschuh, U., Tarasov, P., Wünnemann, B., Hartmann, K., 2004. Holocene vegetation and climate of the Alashan Plateau, NW China, reconstructed from pollen data. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **211**, 1–17. https://doi.org/10.1016/j.palaeo.2004.04.001

Hiltner L, 1904. Über neuere Erfahrungen und Probleme auf dem Gebiete der Bodenbakteriologie unter besonderer Berücksichtigung der Gründüngung und Brache. *Arb. Deutschen Landwirtschaftlichen Ges*. **98**, 59–78.

Hoffland, E., Giesler, R., Jongmans, T., Breemen, N. van, 2002. Increasing feldspar tunneling by fungi across a North Sweden podzol chronosequence. *Ecosystems* **5**, 11–22. https://doi.org/10.1007/s10021-001-0052-x

Hoffland, E., Kuyper, T.W., Wallander, H., *et al.*, 2004. The role of fungi in weathering. *Front. Ecol. Environ.* **2**, 258–264. https://doi.org/10.2307/3868266

Holmgren, M., Lin, C.-Y., Murillo, J.E., *et al*., 2015. Positive shrub–tree interactions facilitate woody encroachment in boreal peatlands. *J. Ecol.* **103**, 58–66. https://doi.org/10.1111/1365-2745.12331

Iwańska, O., Latoch, P., Suchora, M., *et al*., 2022. Lake microbiome and trophy fluctuations of the ancient hemp rettery. *Sci. Rep.* **12**, 8846. https://doi.org/10.1038/s41598-022-12761-w

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., Orlando, L., 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684. https://doi.org/10.1093/bioinformatics/btt193

Kasischke, E.S., 2000. Boreal ecosystems in the global carbon cycle, in: Kasischke, E.S., Stocks, B.J. (Eds.), Fire, Climate Change, and Carbon Cycling in the Boreal Forest, Ecological Studies. *Springer*, *New York*, pp. 19–30. https://doi.org/10.1007/978-0-387-21629-4_2

Kelly, E.F., Chadwick, O.A., Hilinski, T.E., 1998. The effect of plants on mineral weathering. *Biogeochemistry* **42**, 21–53. https://doi.org/10.1023/A:1005919306687

Kisand, V., Talas, L., Kisand, A., *et al*., 2018. From microbial eukaryotes to metazoan vertebrates: Wide spectrum paleo-diversity in sedimentary ancient DNA over the last 14,500 years. *Geobiology* **16**, 628–639. https://doi.org/10.1111/gbi.12307

Kjær, K.H., Winther Pedersen, M., De Sanctis, B., *et al*., 2022. A 2-million-year-old ecosystem in Greenland uncovered by environmental DNA. *Nature* **612**, 283–291. https://doi.org/10.1038/s41586-022-05453-y

Kresnowati, M.T. a P., van Winden, W.A., Almering, M.J.H., *et al.*, 2006. When transcriptome meets metabolome: fast cellular responses of yeast to sudden relief of glucose limitation. *Mol. Syst. Biol*. **2**, 49. https://doi.org/10.1038/msb4100083

Kretzer, A., Li, Y., Szaro, T., Bruns, T.D., 1996. Internal transcribed spacer sequences from 38 recognized species of Suillus sensu lato: Phylogenetic and taxonomic implications. *Mycologia* **88**, 776–785. https://doi.org/10.1080/00275514.1996.12026715

Kruse, S., Herzschuh, U., 2022. Regional opportunities for tundra conservation in the next 1000 years. *eLife* **11**, e75163. https://doi.org/10.7554/eLife.75163

Lasa, A.V., Fernández-González, A.J., Villadas, P.J., Toro, N., Fernández-López, M., 2019. Metabarcoding reveals that rhizospheric microbiota of *Quercus pyrenaica* is composed by a relatively small number of bacterial taxa highly abundant. *Sci. Rep*. **9**, 1695. https://doi.org/10.1038/s41598-018-38123-z

Lee, J., Lee, S., Young, J.P.W., 2008. Improved PCR primers for the detection and identification of arbuscular mycorrhizal fungi. *FEMS Microbiol. Ecol*. **65**, 339–349.

https://doi.org/10.1111/j.1574-6941.2008.00531.x
Leray, M., Yang, J.Y., Meyer, C.P., *et al*., 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Front. Zool.* **10**, 34. https://doi.org/10.1186/1742-9994-10-34
Li, X., Gai, J., Cai, X., *et al.*, 2014. Molecular diversity of arbuscular mycorrhizal fungi associated with two co-occurring perennial plant species on a Tibetan altitudinal gradient. *Mycorrhiza* **24**, 95–107. https://doi.org/10.1007/s00572-013-0518-7
Liu, H., Liang, C., Ai, Z., *et al*., 2019. Plant-mycorrhizae association affects plant diversity, biomass, and soil nutrients along temporal gradients of natural restoration after farmland abandonment in the Loess Plateau, China. *Land Degrad. Dev.* **30**, 1677–1690. https://doi.org/10.1002/ldr.3372
Liu, S., Stoof-Leichsenring, K.R., Kruse, S., Pestryakova, L.A., Herzschuh, U., 2020. Holocene vegetation and plant diversity changes in the north-eastern Siberian treeline region from pollen and sedimentary ancient DNA. *Front. Ecol. Evol*. **8**. https://doi.org/10.3389/fevo.2020.560243
Lloyd-Macgilp, S.A., Chambers, S.M., Dodd, J.C., *et al.*, 1996. Diversity of the ribosomal internal transcribed spacers within and among isolates of *Glomus mosseae* and related mycorrhizal fungi. *New Phytol*. **133**, 103–111. https://doi.org/10.1111/j.1469-8137.1996.tb04346.x
Loughlin, N.J.D., Gosling, W.D., Montoya, E., 2018. Identifying environmental drivers of fungal non-pollen palynomorphs in the montane forest of the eastern Andean flank, Ecuador. *Quat. Res.* **89**, 119–133. https://doi.org/10.1017/qua.2017.73
Lundström, U. S, van Breemen, N., Bain, D., 2000. The podzolization process. A review. *Geoderma* **94**, 91–107. https://doi.org/10.1016/S0016-7061(99)00036-1
Lundström, U. S., van Breemen, N., Bain, D.C., *et al*., 2000. Advances in understanding the podzolization process resulting from a multidisciplinary study of three coniferous forest soils in the Nordic Countries. *Geoderma* **94**, 335–353. https://doi.org/10.1016/S0016-7061(99)00077-4
Lynch, J.M., Whipps, J.M., 1990. Substrate flow in the rhizosphere. *Plant Soil* **129**, 1–10. https://doi.org/10.1007/BF00011685
MacDonald, G.M., Beukens, R.P., Kieser, W.E., 1991. Radiocarbon dating of limnic sediments: A comparative analysis and discussion. *Ecology* **72**, 1150–1155. https://doi.org/10.2307/1940612
Mašínová, T., Yurkov, A., Baldrian, P., 2018. Forest soil yeasts: Decomposition potential and the utilization of carbon sources. *Fungal Ecol.* **34**, 10–19. https://doi.org/10.1016/j.funeco.2018.03.005
Mcguire, A.D., Hayes, D.J., Kicklighter, D.W., *et al.*, 2010. An analysis of the carbon balance of the Arctic Basin from 1997 to 2006. *Tellus B Chem. Phys. Meteorol.* **62**, 455–474. https://doi.org/10.1111/j.1600-0889.2010.00497.x
McLaughlin, D.J., Hibbett, D.S., Lutzoni, F., Spatafora, J.W., Vilgalys, R., 2009. The search for the fungal tree of life. *Trends Microbiol.* **17**, 488–497. https://doi.org/10.1016/j.tim.2009.08.001
Melles, M., Svendsen, J.I., Fedorov, G., Brigham-Grette, J., Wagner, B., 2022. Quaternary environmental and climatic history of the northern high latitudes – recent contributions and perspectives from lake sediment records. *J. Quat. Sci.* **37**, 721–728. https://doi.org/10.1002/jqs.3456
Mestre, M.C., Tamayo Navarrete, M.I., García Garrido, J.M., 2022. Exploring the yeast-mycorrhiza-plant interaction: *Saccharomyces eubayanus* negative effects on arbuscular mycorrhizal formation in tomato plants. *Plant Soil* **479**, 529–542. https://doi.org/10.1007/s11104-022-05538-7
Meyer, M., Kircher, M., Gansauge, M.-T., *et al.*, 2012. A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222–226. https://doi.org/10.1126/science.1224344
Myers-Smith, I.H., Forbes, B.C., Wilmking, M., *et al*., 2011. Shrub expansion in tundra ecosystems: dynamics, impacts and research priorities. *Environ. Res. Lett.* **6**, 045509. https://doi.org/10.1088/1748-9326/6/4/045509

Nguyen, N.H., Vellinga, E.C., Bruns, T.D., Kennedy, P.G., 2016. Phylogenetic assessment of global *Suillus* ITS sequences supports morphologically defined species and reveals synonymous and undescribed taxa. *Mycologia* **108**, 1216–1228. https://doi.org/10.3852/16-106

Nuñez, M.A., Horton, T.R., Simberloff, D., 2009. Lack of belowground mutualisms hinders Pinaceae invasions. *Ecology* **90**, 2352–2359. https://doi.org/10.1890/08-2139.1

Pan, Y., Birdsey, R.A., Fang, J., *et al*., 2011. A large and persistent carbon sink in the world's forests. *Science* **333**, 988–993. https://doi.org/10.1126/science.1201609

Pandey, P., Irulappan, V., Bagavathiannan, M.V., Senthil-Kumar, M., 2017. Impact of combined abiotic and biotic stresses on plant growth and avenues for crop improvement by exploiting physio-morphological traits. *Front. Plant Sci.* **8**. https://doi.org/10.3389/fpls.2017.00537

Pansu, J., Giguet-Covex, C., Ficetola, G.F., *et al*., 2015. Reconstructing long-term human impacts on plant communities: an ecological approach based on lake sediment DNA. *Mol. Ecol.* **24**, 1485–1498. https://doi.org/10.1111/mec.13136

Parducci, L., Alsos, I.G., Unneberg, P., *et al*., 2019. Shotgun environmental DNA, pollen, and macrofossil analysis of lateglacial lake sediments from southern Sweden. *Front. Ecol. Evol.* **7**. https://doi.org/10.3389/fevo.2019.00189

Parducci, L., Bennett, K.D., Ficetola, G.F., *et al.*, 2017. Ancient plant DNA in lake sediments. *New Phytol.* **214**, 924–942. https://doi.org/10.1111/nph.14470

Parducci, L., Matetovici, I., Fontana, S.L., *et al*., 2013. Molecular- and pollen-based vegetation analysis in lake sediments from central Scandinavia. *Mol. Ecol.* **22**, 3511–3524. https://doi.org/10.1111/mec.12298

Parducci, L., Väliranta, M., Salonen, J.S., *et al.*, 2015. Proxy comparison in ancient peat sediments: pollen, macrofossil and plant DNA. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20130382. https://doi.org/10.1098/rstb.2013.0382

Parniske, M., 2008. Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nat. Rev. Microbiol.* **6**, 763–775. https://doi.org/10.1038/nrmicro1987

Pearson, R.G., Phillips, S.J., Loranty, M.M., *et al.*, 2013. Shifts in Arctic vegetation and associated feedbacks under climate change. *Nat. Clim. Change* **3**, 673–677. https://doi.org/10.1038/nclimate1858

Pedersen, M.W., Ginolhac, A., Orlando, L., *et al.*, 2013. A comparative study of ancient environmental DNA to pollen and macrofossils from lake sediments reveals taxonomic overlap and additional plant taxa. *Quat. Sci. Rev*. **75**, 161–168. https://doi.org/10.1016/j.quascirev.2013.06.006

Pedersen, M.W., Ruter, A., Schweger, C., *et al*., 2016. Postglacial viability and colonization in North America's ice-free corridor. *Nature* **537**, 45–49. https://doi.org/10.1038/nature19085

Peng, Y., Yuan, J., Heděnec, P., *et al*., 2022. Mycorrhizal association and life form dominantly control plant litter lignocellulose concentration at the global scale. *Front. Plant Sci.* **13**, 926941. https://doi.org/10.3389/fpls.2022.926941

Pietramellara, G., Ascher, J., Borgogni, F., *et al*., 2009. Extracellular DNA in soil and sediment: fate and ecological relevance. *Biol. Fertil. Soils* **45**, 219–235. https://doi.org/10.1007/s00374-008-0345-8

Põlme, S., Bahram, M., Jacquemyn, H., *et al.*, 2018. Host preference and network properties in biotrophic plant–fungal associations. *New Phytol.* **217**, 1230–1239. https://doi.org/10.1111/nph.14895

Potapov, P., Yaroshenko, A., Turubanova, S., *et al.*, 2008. Mapping the world's intact forest landscapes by remote sensing. *Ecol. Soc.* **13**. https://doi.org/10.5751/ES-02670-130251

Praeg, N., Illmer, P., 2020. Microbial community composition in the rhizosphere of *Larix decidua* under different light regimes with additional focus on methane cycling microorganisms. *Sci. Rep*. **10**, 22324. https://doi.org/10.1038/s41598-020-79143-y

Praeg, N., Pauli, H., Illmer, P., 2019. Microbial diversity in bulk and rhizosphere soil of *Ranunculus glacialis* along a high-alpine altitudinal gradient. *Front. Microbiol.* **10**. https://doi.org/10.3389/fmicb.2019.01429

Qian, H., Klinka, K., Kayahara, G.J., 1998. Longitudinal patterns of plant diversity in the North American boreal forest. *Plant Ecol.* **138**, 161–178. https://doi.org/10.1023/A:1009756318848

Qu, Z.-L., Santalahti, M., Köster, K., *et al*., 2021. Soil fungal community structure in boreal pine forests: from southern to subarctic areas of Finland. *Front. Microbiol.* **12**, 653896. https://doi.org/10.3389/fmicb.2021.653896

Quamar, Md.F., Stivrins, N., 2021. Modern pollen and non-pollen palynomorphs along an altitudinal transect in Jammu and Kashmir (Western Himalaya), India. *Palynology* **45**, 669–684. https://doi.org/10.1080/01916122.2021.1915402

Rahman, M.M., Tsukamoto, J., Motiur Rahman, Md., *et al*., 2013. Lignin and its effects on litter decomposition in forest ecosystem. *Chem. Ecol*. **29**. https://doi.org/10.1080/02757540.2013.790380

Raidl, S., Bonfigli, R., Agerer, R., 2005. Calibration of quantitative real-time taqman PCR by correlation with hyphal biomass and ITS copies in mycelia of *Piloderma croceum*. *Plant Biol.* **7**, 713–717. https://doi.org/10.1055/s-2005-873003

Read, D.J., Perez-Moreno, J., 2003. Mycorrhizas and nutrient cycling in ecosystems – a journey towards relevance? *New Phytol*. **157**, 475–492. https://doi.org/10.1046/j.1469-8137.2003.00704.x

Sakurai, M., Suzuki, K., Onodera, M., *et al*., 2007. Analysis of bacterial communities in soil by PCR–DGGE targeting protease genes. *Soil Biol. Biochem*. **39**, 2777–2784. https://doi.org/10.1016/j.soilbio.2007.05.026

Santos, A., Sánchez, A., Marquina, D., 2004. Yeasts as biological agents to control *Botrytis cinerea*. *Microbiol. Res*. **159**, 331–338. https://doi.org/10.1016/j.micres.2004.07.001

Sasaki, K., Tsunekawa, M., Ohtsuka, T., Konno, H., 1998. The role of sulfur-oxidizing bacteria *Thiobacillus thiooxidans* in pyrite weathering. *Colloids Surf. Physicochem. Eng. Asp*. **133**, 269–278. https://doi.org/10.1016/S0927-7757(97)00200-8

Sauer, D., Sponagel, H., Sommer, M., *et al.*, 2007. Podzol: soil of the year 2007. A review on its genesis, occurrence, and functions. *J. Plant Nutr. Soil Sci*. **170**, 581–597. https://doi.org/10.1002/jpln.200700135

Schlesinger, W.H., Bernhardt, E.S., 2013. Biogeochemistry: An Analysis of Global Change. *Academic Press*.

Schoch, C.L., Robbertse, B., Robert, V., *et a*l., 2014. Finding needles in haystacks: linking scientific names, reference specimens and molecular data for fungi. *Database* **2014**, bau061. https://doi.org/10.1093/database/bau061

Schoch, C.L., Seifert, K.A., Huhndorf, S., K., *et al*., 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for fungi. *Proc. Natl. Acad. Sci.* **109**, 6241–6246. https://doi.org/10.1073/pnas.1117018109

Sepp, S.-K., Davison, J., Moora, M., *et al.*, 2021. Woody encroachment in grassland elicits complex changes in the functional structure of above- and belowground biota. *Ecosphere* **12**, e03512. https://doi.org/10.1002/ecs2.3512

Sharma, S.K., Singh, U.B., Sahu, P.K., *et al*. (Eds.), 2020. Rhizosphere Microbes: Soil and Plant Functions, Microorganisms for Sustainability. *Springer*, Singapore. https://doi.org/10.1007/978-981-15-9154-9

Shevtsova, I., Heim, B., Kruse, S., *et al*., 2020. Strong shrub expansion in tundra-taiga, tree infilling in taiga and stable tundra in central Chukotka (north-eastern Siberia) between 2000 and 2017. *Environ. Res. Lett*. **15**, 085006. https://doi.org/10.1088/1748-9326/ab9059

da Silva, I.R., de Mello, C.M.A., Ferreira Neto, R.A., *et al*., 2014. Diversity of arbuscular mycorrhizal fungi along an environmental gradient in the Brazilian semiarid. *Appl. Soil Ecol.* **84**, 166–175. https://doi.org/10.1016/j.apsoil.2014.07.008

Singleton, G.A., Lavkulich, L.M., 1987. A soil chronosequence on beach sands, Vancouver island, British Columbia. *Can. J. Soil Sci.* **67**, 795–810. https://doi.org/10.4141/cjss87-077

Smith, S.E., Read, D.J., 2010. Mycorrhizal Symbiosis. *Academic Press*.

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Avery, K., M.Tignor, Miller, H., 2007. Climate Change 2007: The Physical Science Basis. Working group I - Contribution to the fourth

assessment report of the IPCC. Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.). *Cambridge University Press*, Cambridge, United Kingdom and New York, NY, USA

Sønstebø, J.H., Gielly, L., Brysting, A.K., *et al*., 2010. Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate. *Mol. Ecol. Resour*. **10**, 1009–1018. https://doi.org/10.1111/j.1755-0998.2010.02855.x

Stackebrandt, E., Goebel, B.M., 1994. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Evol. Microbiol.* **44**, 846–849. https://doi.org/10.1099/00207713-44-4-846

Stefanowicz, A.M., Kapusta, P., Stanek, M., *et al*., 2022. Herbaceous plant species support soil microbial performance in deciduous temperate forests. *Sci. Total Environ.* **810**, 151313. https://doi.org/10.1016/j.scitotenv.2021.151313

Stivrins, N., Soininen, J., Tõnno, I., *et al*., 2018. Towards understanding the abundance of non-pollen palynomorphs: A comparison of fossil algae, algal pigments and sedaDNA from temperate lake sediments. *Rev. Palaeobot. Palynol.* 249, 9–15. https://doi.org/10.1016/j.revpalbo.2017.11.001

Stoeck, T., Bass, D., Nebel, M., *et al.*, 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol*. **19**, 21–31. https://doi.org/10.1111/j.1365-294X.2009.04480.x

Stonestrom, D.A., White, A.F., C. Akstin, K., 1998. Determining rates of chemical weathering in soils—solute transport versus profile evolution. *J. Hydrol.* **209**, 331–345. https://doi.org/10.1016/S0022-1694(98)00158-9

Taberlet, P., Coissac, E., Pompanon, F., *et al*., 2007. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Res.* **35**, e14. https://doi.org/10.1093/nar/gkl938

Talas, L., Stivrins, N., Veski, S., *et al*., 2021. Sedimentary ancient DNA (sedaDNA) reveals fungal diversity and environmental drivers of community changes throughout the Holocene in the present boreal lake Lielais Svētiņu (Eastern Latvia). *Microorganisms* **9**, 719. https://doi.org/10.3390/microorganisms9040719

Tautenhahn, S., Lichstein, J.W., Jung, M., *et al.*, 2016. Dispersal limitation drives successional pathways in central Siberian forests under current and intensified fire regimes. *Glob. Change Biol.* **22**, 2178–2197. https://doi.org/10.1111/gcb.13181

Tedersoo, L., Anslan, S., Bahram, M., *et al.*, 2015. Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi. *MycoKeys* **10**, 1–43. https://doi.org/10.3897/mycokeys.10.4852

Treseder, K.K., Marusenko, Y., Romero-Olivares, A.L., Maltz, M.R., 2016. Experimental warming alters potential function of the fungal community in boreal forest. *Glob. Change Biol*. **22**, 3395–3404. https://doi.org/10.1111/gcb.13238

Uroz, S., Calvaruso, C., Turpault, M.-P., Frey-Klett, P., 2009. Mineral weathering by bacteria: ecology, actors and mechanisms. *Trends Microbiol.* **17**, 378–387. https://doi.org/10.1016/j.tim.2009.05.004

Vacheron, J., Desbrosses, G., Bouffaud, M.-L., *et al*., 2013. Plant growth-promoting rhizobacteria and root system functioning. *Front. Plant Sci*. **4**. https://doi.org/10.3389/fpls.2013.00356

Van Cleve, K., Viereck, L.A., 1981. Forest succession in relation to nutrient cycling in the boreal forest of Alaska, in: West, D.C., Shugart, H.H., Botkin, D.B. (Eds.), Forest Succession: Concepts and Application, Springer Advanced Texts in Life Sciences. *Springer*, New York, NY, pp. 185–211. https://doi.org/10.1007/978-1-4612-5950-3_13

van der Heijden, M.G.A., Martin, F.M., Selosse, M.-A., Sanders, I.R., 2015. Mycorrhizal ecology and evolution: the past, the present, and the future. *New Phytol*. **205**, 1406–1423. https://doi.org/10.1111/nph.13288

Van Geel, B., 2001. Non-pollen palynomorphs, in: Smol, J.P., Birks, H.J.B., Last, W.M., Bradley, R.S., Alverson, K. (Eds.), Tracking Environmental Change Using Lake Sediments: Terrestrial, algal,

and siliceous indicators, developments in paleoenvironmental research. *Springer Netherlands*, Dordrecht, pp. 99–119. https://doi.org/10.1007/0-306-47668-1_6

Vieira, S., Sikorski, J., Dietz, S., *et al*., 2020. Drivers of the composition of active rhizosphere bacterial communities in temperate grasslands. *ISME J.* **14**, 463–475. https://doi.org/10.1038/s41396-019-0543-4

Viers, J., Oliva, P., Dandurand, J.-L., *et al*., 2014. Chemical weathering rates, CO_2 consumption, and control parameters deduced from the chemical composition of rivers, in: Holland, H.D., Turekian, K.K. (Eds.), Treatise on Geochemistry (2nd Edition). *Elsevier*, Oxford, pp. 175–194. https://doi.org/10.1016/B978-0-08-095975-7.00506-4

Voroney, R.P., 2007. The soil habitat, in: Soil Microbiology, Ecology and Biochemistry. *Elsevier*, pp. 25–49. https://doi.org/10.1016/B978-0-08-047514-1.50006-8

Vuong, T.M.D., Zeng, J.Y., Man, X.L., 2020. Soil fungal and bacterial communities in southern boreal forests of the Greater Khingan Mountains and their relationship with soil properties. *Sci. Rep.* **10**, 22025. https://doi.org/10.1038/s41598-020-79206-0

Walker, D.A., Raynolds, M.K., Daniëls, F.J.A., *et al*., 2005. The circumpolar Arctic vegetation map. *J. Veg. Sci.* **16**, 267–282. https://doi.org/10.1111/j.1654-1103.2005.tb02365.x

Walker, M.D., Wahren, C.H., Hollister, R.D., *et al*., 2006. Plant community responses to experimental warming across the tundra biome. *Proc. Natl. Acad. Sci.* **103**, 1342–1346. https://doi.org/10.1073/pnas.0503198103

Wallenstein, M.D., Weintraub, M.N., 2008. Emerging tools for measuring and modeling the *in situ* activity of soil extracellular enzymes. *Soil Biol. Biochem.*, Special Section: Enzymes in the Environment **40**, 2098–2106. https://doi.org/10.1016/j.soilbio.2008.01.024

Walters, W., Hyde, E.R., Berg-Lyons, D., *et al*., 2015. Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems* **1**, e00009-15. https://doi.org/10.1128/mSystems.00009-15

Wang, P., Limpens, J., Nauta, A., *et al*., 2018. Depth-based differentiation in nitrogen uptake between graminoids and shrubs in an Arctic tundra plant community. *J. Veg. Sci*. **29**, 34–41. https://doi.org/10.1111/jvs.12593

Wang, Y., Pedersen, M.W., Alsos, I.G., *et al*., 2021. Late Quaternary dynamics of Arctic biota from ancient environmental genomics. *Nature* **600**, 86–92. https://doi.org/10.1038/s41586-021-04016-x

Weber, S.E., Diez, J.M., Andrews, L.V., *et al*., 2019. Responses of arbuscular mycorrhizal fungi to multiple coinciding global change drivers. *Fungal Ecol.*, Ecology of Mycorrhizas in the Anthropocene **40**, 62–71. https://doi.org/10.1016/j.funeco.2018.11.008

Welch, S.A., Taunton, A.E., Banfield, J.F., 2002. Effect of microorganisms and microbial metabolites on apatite dissolution. *Geomicrobiol. J.* **19**, 343–367. https://doi.org/10.1080/01490450290098414

Welker, F., Collins, M.J., Thomas, J.A., *et al*., 2015. Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. *Nature* **522**, 81–84. https://doi.org/10.1038/nature14249

Welker, F., Hajdinjak, M., Talamo, S., *et al.*, 2016. Palaeoproteomic evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne. *Proc. Natl. Acad. Sci.* **113**, 11162–11167. https://doi.org/10.1073/pnas.1605834113

Whipps, J.M., Lynch, J.M., 1985. Energy losses by the plant in rhizodeposition. *Ann. Proc. Phytochem. Soc. Eur.* **26**, 59–71.

White, Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA Genes for phylogenetics. pp. 315–322. In: PCR - Protocols and Applications - A Laboratory Manual. *Academic Press*, Cambridge

Wiklander, L., Andersson, A., 1972. The replacing efficiency of hydrogen ion in relation to base saturation and pH. *Geoderma* **7**, 159–165. https://doi.org/10.1016/0016-7061(72)90002-X

Yang, J., Kloepper, J.W., Ryu, C.-M., 2009. Rhizosphere bacteria help plants tolerate abiotic stress. *Trends Plant Sci.* **14**, 1–4. https://doi.org/10.1016/j.tplants.2008.10.004

Zhang, W., Miller, P.A., Smith, B., *et al*., R., 2013. Tundra shrubification and tree-line advance amplify arctic climate warming: results from an individual-based dynamic vegetation model. *Environ. Res. Lett.* **8**, 034023. https://doi.org/10.1088/1748-9326/8/3/034023

Zimov, S.A., Schuur, E.A.G., Chapin, F.S., 2006. Permafrost and the global carbon budget. *Science* **312**, 1612–1613. https://doi.org/10.1126/science.1128908

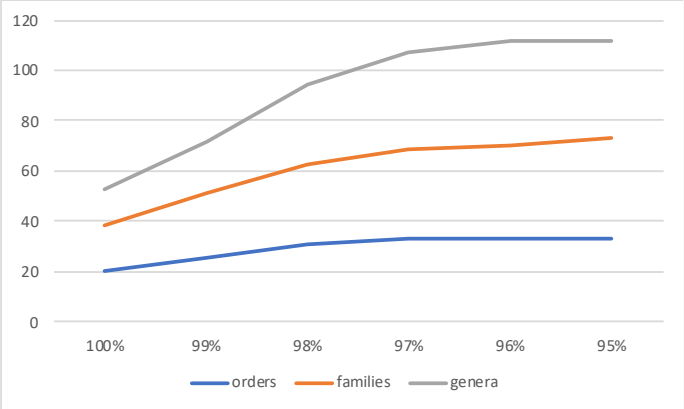
7 Appendix

Due to their overall sizes, all supplementary tables to the manuscripts are provided separately on a CD.

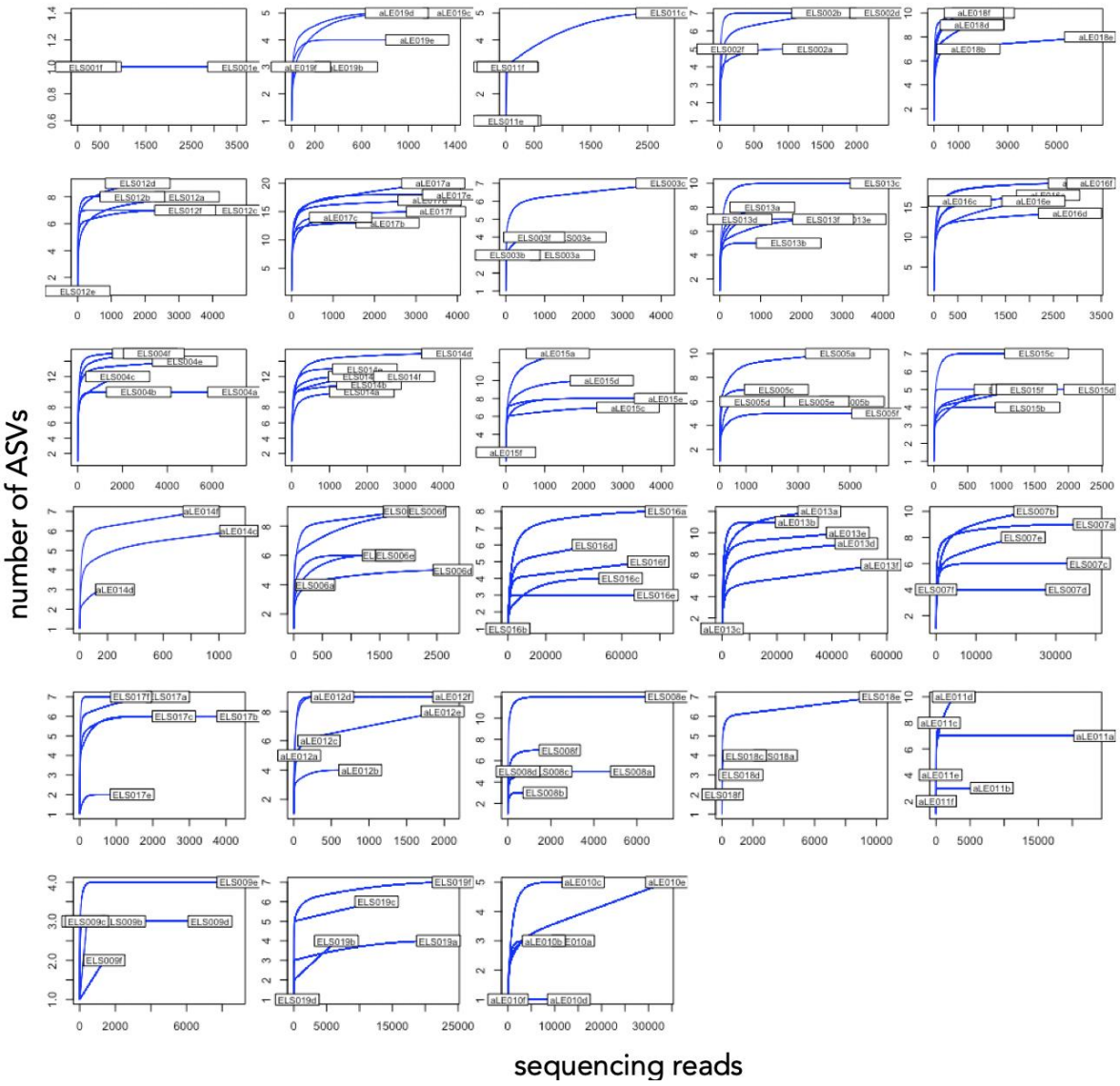
7.1 Appendix to manuscript I

Supplemental figures for

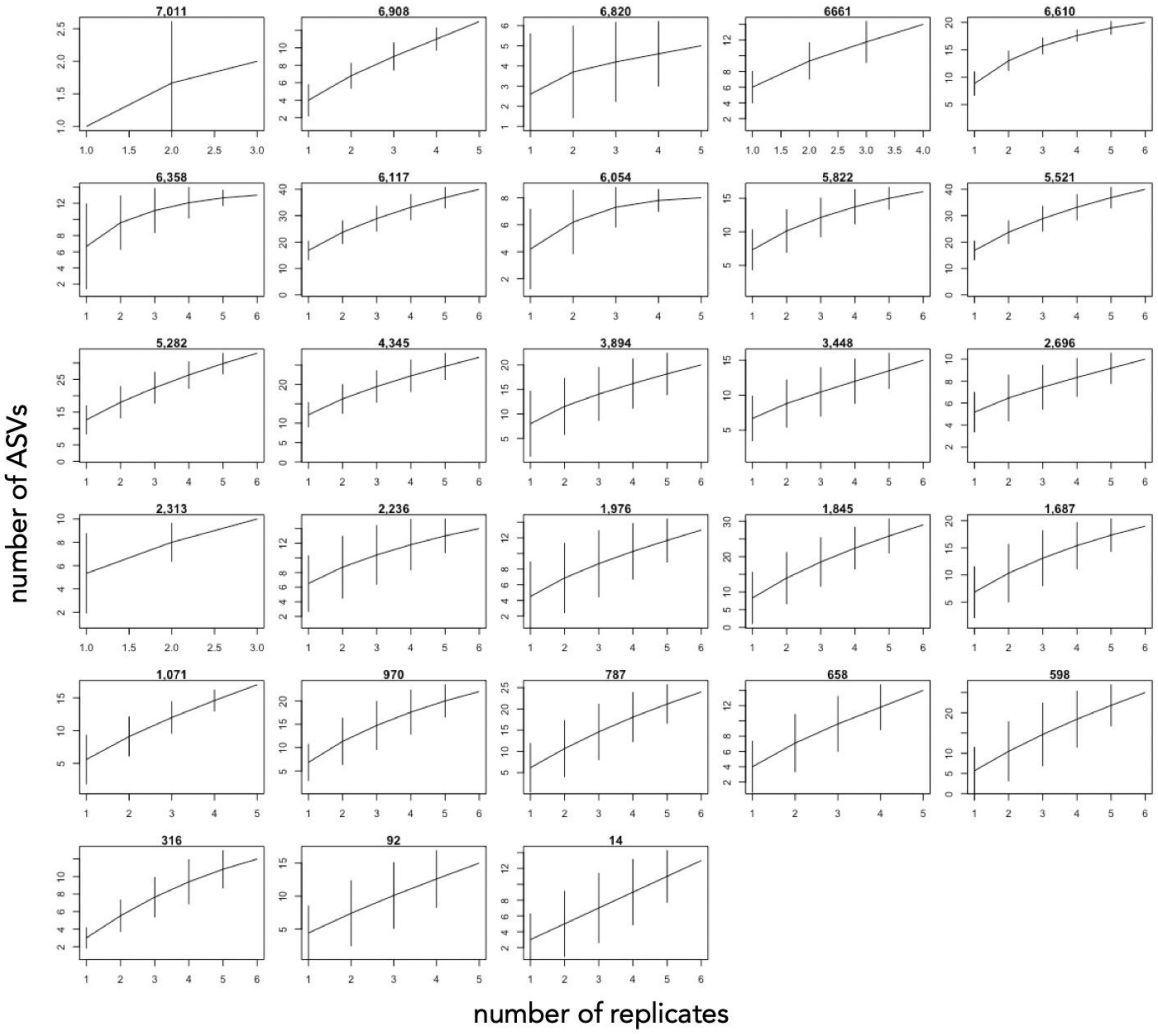
P.A. Seeber, **B. von Hippel**, H. Kauserud, U. Löber, K.R. Stoof-Leichsenring, U. Herzschuh, L.S. Epp. Evaluation of lake sedimentary ancient DNA metabarcoding to assess fungal biodiversity in Arctic paleoecosystems. Environmental DNA (2022), 10.1002/edn3.315.



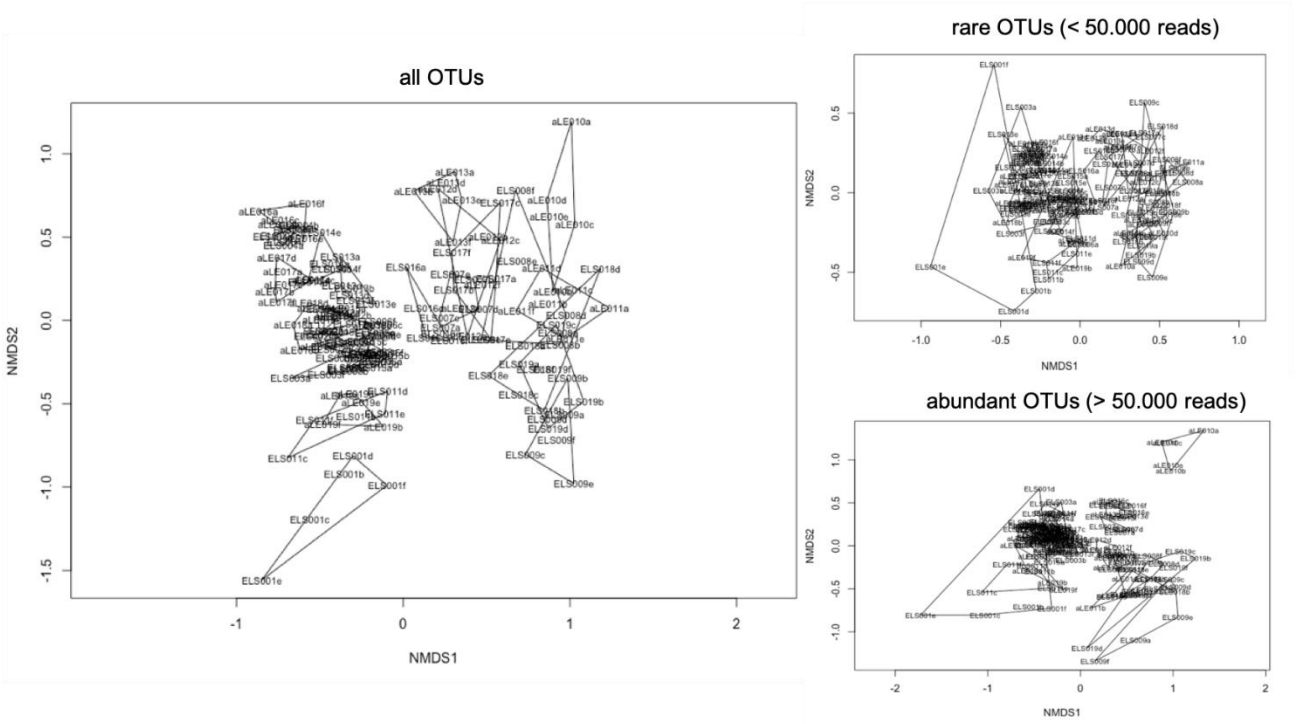
Supplementary Figure S1. Numbers of orders, families, and genera at best-identity cutoffs of 95%–100% (core CH12, before further filtration steps).



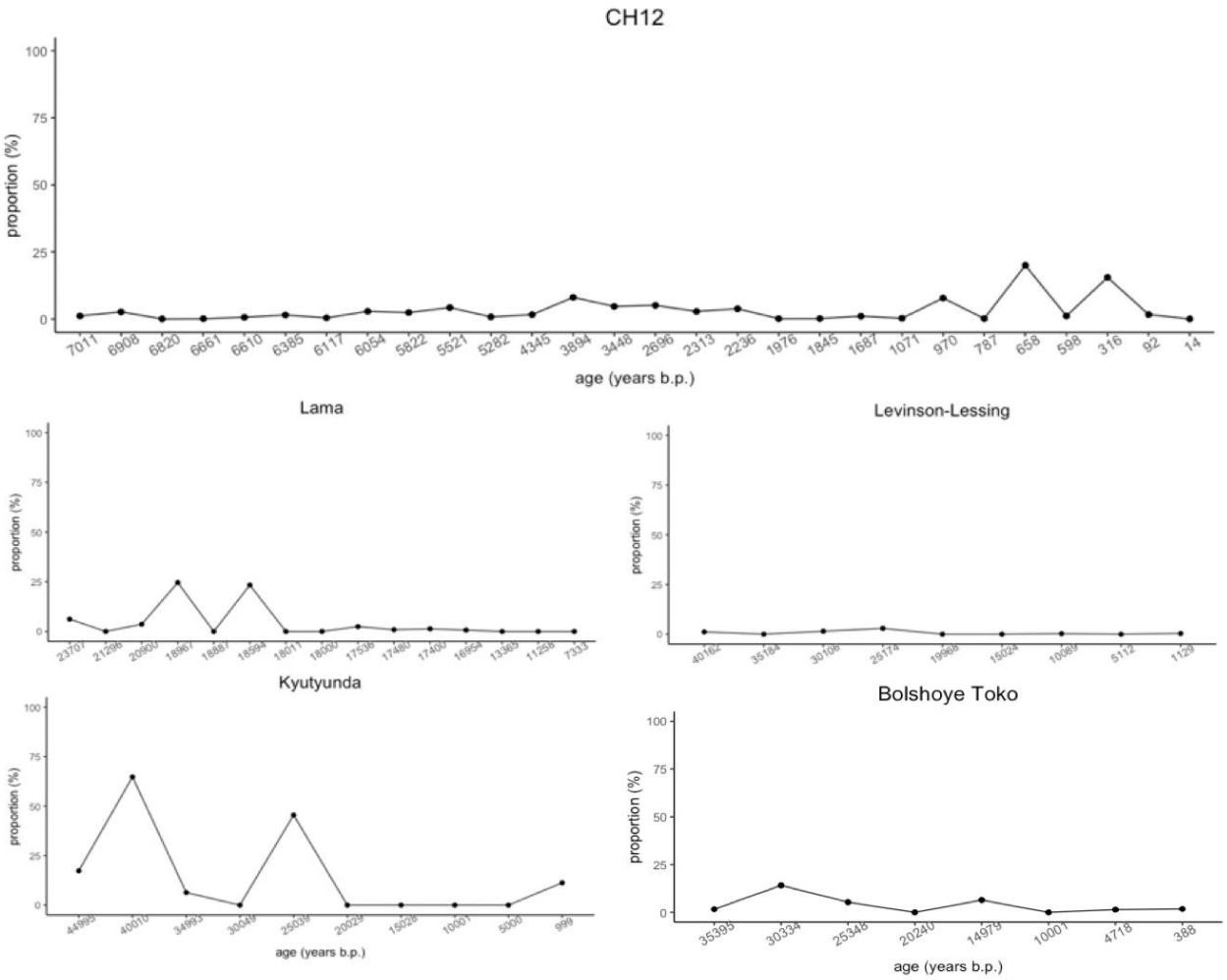
Supplementary Figure S2. Rarefaction curves of PCR replicates of core CH12. Shown are the numbers of OTUs as a function of the number of sequencing reads.



Supplementary Figure S3. Saturation curves of number of OTUs per cumulative number of replicates in the sediment core of lake CH12. The age of each sample is indicated above the respective graph.



Supplementary Figure S4. Cluster visualization based on Bray-Curtis distances of all OTUs (left), rare OTUs (< 50,000 reads; top right), and abundant OTUs (> 50,000 reads; bottom right).



Supplementary Figure S5. Cumulative proportions of mold genera (*Aspergillus*, *Cladosporium*, *Mucor*, and *Penicillium*) over time in the five sediment cores.

7.2 Appendix to manuscript II

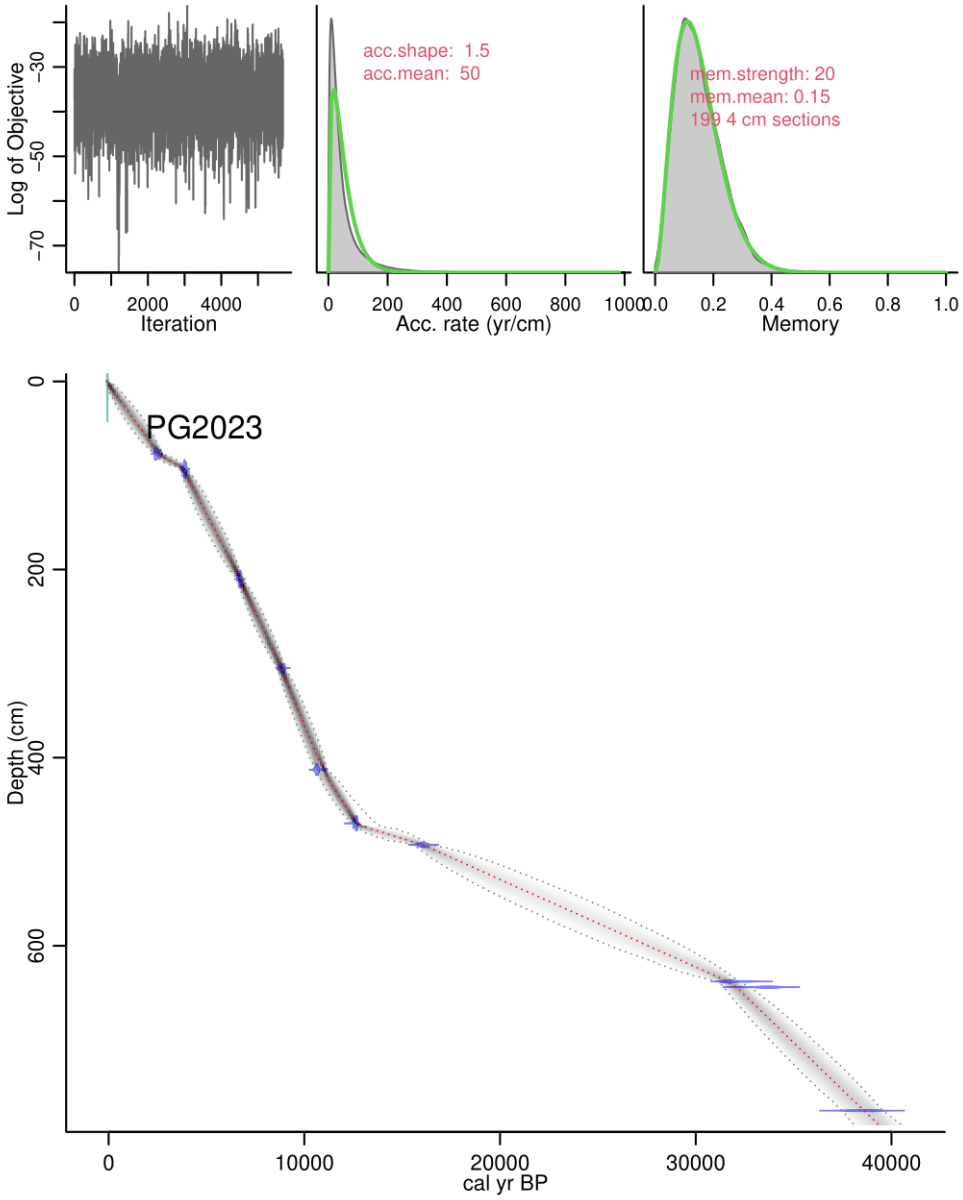
Supplemental information for

von Hippel, B., Stoof-Leichsenring, K.R., Schulte, L., Seeber, P., Epp, L.S, Melles, M., Herzschuh, U. Long-term fungus–plant covariation from multi-site sedimentary ancient DNA metabarcoding. *Quat. Sci. Rev.* **295**, 107758 (2022).

Supplement 1: Available radiocarbon ages (years before 1950 CE) from Biskaborn et al. (2016) next to slightly corrected composite depths, dating error, and method applied in the Poznan radiocarbon laboratory. RES insoluble humin fraction; SOL alkali-soluble humic acids fraction; TOC total organic carbon. To perform age-depth modelling we used RES values from bulk sediment samples.

^{14}C Lab ID	Sample ID	^{14}C age (yrs)	^{14}C error (yrs)	Depth below sediment surface (cm)	Sample type	Method
Poz-49481	PG2023-2_50,5-51	2405	35	76.75	bulk	RES
Poz-49472	PG2023-2_71,5-72	3585	30	93.25	bulk	RES
Poz-49471	PG2023-2_188-188,5	5900	40	210.25	bulk	RES
Poz-49470	PG2023-3_79-79,5	8000	50	304.75	bulk	RES
Poz-49474	PG2023-3_187-187,5	9420	50	412.75	bulk	RES
Poz-49483	PG2023-3_244-244,5	10620	60	469.75	bulk	RES
Poz-49482	PG2023-4_78,5-79	13360	100	492.75	bulk	RES
Poz-50559-	PG2023-5_25-26	16870	260	638	bulk	SOL
Poz-50560	PG2023-5_25-26	27820	300	638	bulk	RES
Poz-50557	PG2023-4_231-232	29180	350	644	bulk	RES
Poz-50558-	PG2023-5_25-26	29990	380	638	bulk	TOC
Poz-50555-	PG2023-4_231-232	33500	500	644	bulk	TOC
Poz-49484	PG2023-5_162,5-163	33770	350	775.25	bulk	RES
Poz-50556-	PG2023-4_231-232	33900	700	644	bulk	SOL

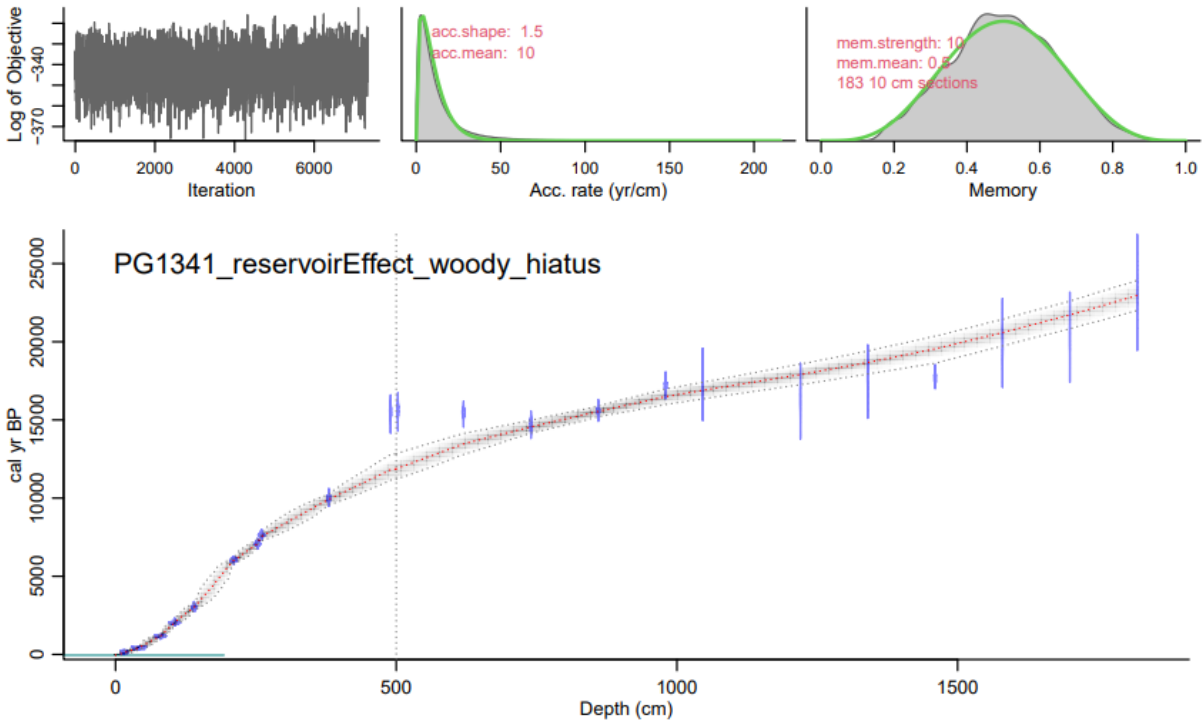
Supplement 2: Bacon age-depth model based on radiocarbon age determinations (bulk sediment samples, RES values, Supplement 1) from sediment core PG2023 retrieved in 2010 from Lake Kyutyunda. This is a refined version of the age-depth correlation for this sediment core published by Biskaborn et al. (2016), based on the IntCal20 calibration curve (Reimer et al., 2020) and compiled in the R package bacon (Blaauw and Christen, 2011).



Supplement 3: Radiocarbon ages from the Lama PG1341 core which were used to calculate the age-depth model alongside dated woody remains (*: data from Andreev et al. (2014)). The reservoir effect for this core is 4460 years and was subtracted from the determined ^{14}C ages before the age-depth modelling.

^{14}C Lab ID	Sample ID	^{14}C age (yrs)	^{14}C error (yrs)	Depth below sediment surface (cm)	Sample type	Method
6794	PG1341-4AR_15-16	4597	25	15-16	Bulk	C14
6795	PG1341-4AR_42-43	4845	25	42-43	Bulk	C14
6796	PG1341-4AR_80-81	5698	25	80-81	Bulk	C14
6797	PG1341-4AR_105-106	6530	26	105-106	Bulk	C14
6036	PG1341-4_140	7333	62	140-141	Bulk	C14
UTC8876	Woody remains	5255*	48	211	Woody remains	C14
UTC8877	Woody remains	6200*	60	253	Woody remains	C14
6037	PG1341-5_260	11258	83	260-261	Bulk	C14
6038	PG1341-5_380	13365	100	380-381	Bulk	C14
6039	PG1341-5_503	17536	154	503-504	Bulk	C14
6040	PG1341-6_620	17400	104	620-621	Bulk	C14
6041	PG1341-6_740	16954	101	740-741	Bulk	C14
6042	PG1341-7_860	17480	105	860-861	Bulk	C14
6043	PG1341-7_980	18594	120	980-981	Bulk	C14
6044	PG1341-8_1046	18887	397	1046-1047	Bulk	C14
6045	PG1341-8_1220	18011	354	1220-1221	Bulk	C14
6046	PG1341-9_1340	18976	395	1340-1341	Bulk	C14
6047	PG1341-9_1460	18967	119	1460-1461	Bulk	C14
6048	PG1341-10_1580	20900	501	1580-1581	Bulk	C14
6049	PG1341-10_1700	21296	531	1700-1701	Bulk	C14
6050	PG1341-11_1820	23707	683	1820-1821	Bulk	C14

Supplement 4: Age-depth model for Lake Lama. The age-depth model was established using the package bacon() in R. The reservoir effect is 4460 years. We used two dated woody remains alongside 19 radiocarbon-dated bulk sediment samples to set up the age model. A change in the sedimentation rate after a depth of 500 cm is visible, leading to a hiatus in the age model.

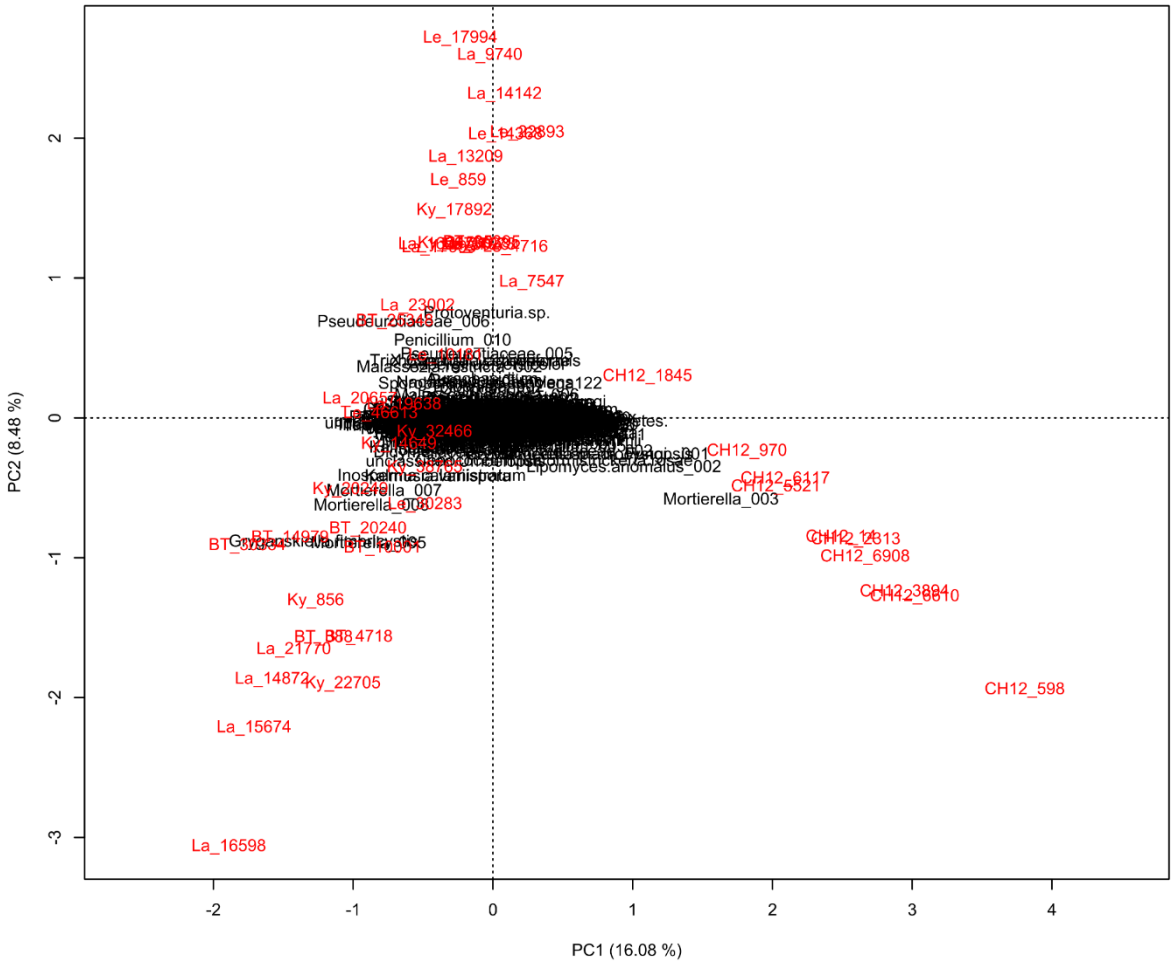


Supplement 5: Obitools Pipeline and Cleaning Steps. The single steps and their commands are listed in the table as well as an explanation of each step. The last column includes the resulting size of the fungal dataset as an example.

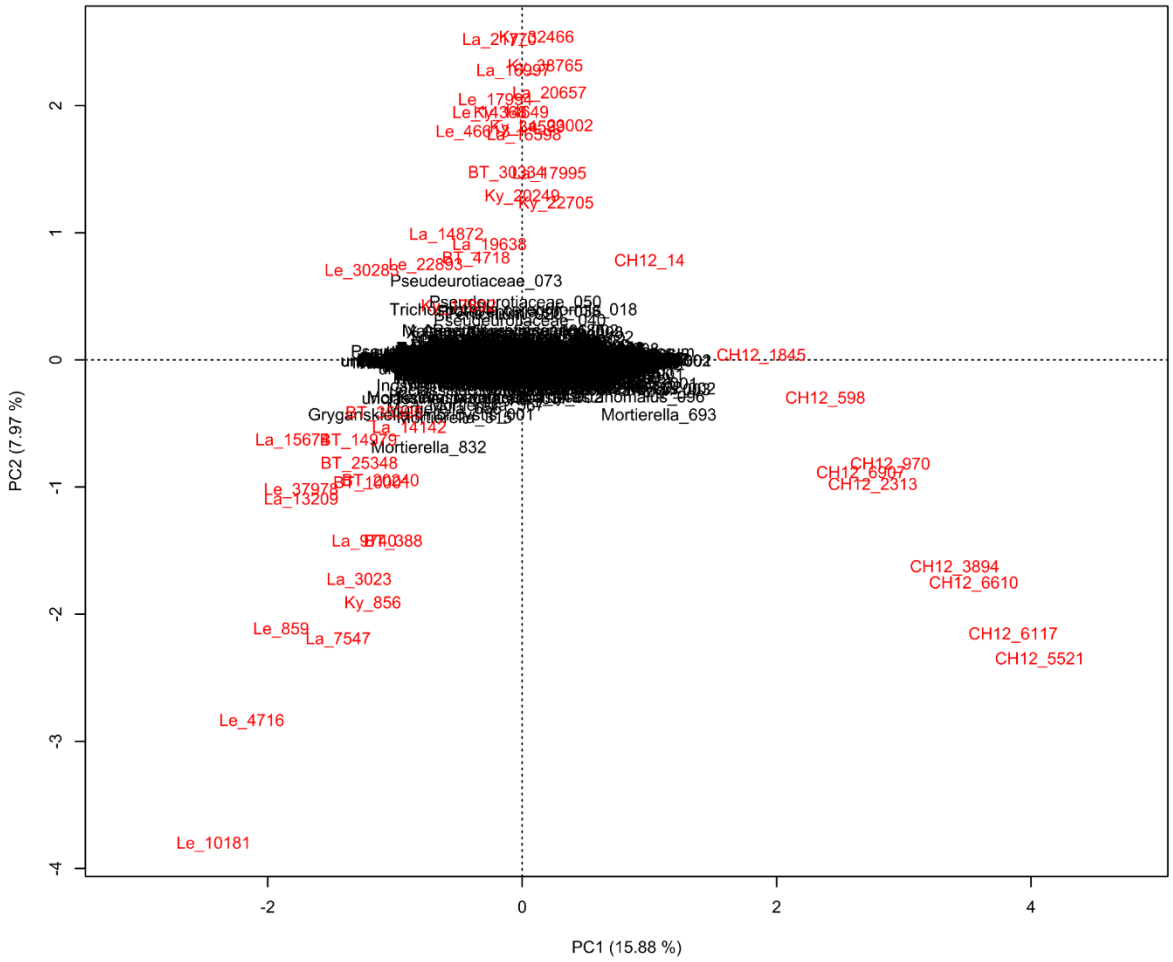
Step	Command	Result (in brackets: the size of the fungal dataset after the subsequent step)
illuminapairedend	illuminapairedend inputR1.fastq -r inputR2.fastq > paired_end.fastq	Paired-ending of sequences (52,213,129)
Obigrep	obigrep -p 'mode!="joined"' paired_end.fastq > paired_end_joined.fastq	Greps out only joined sequences (52,213,129)
Ngsfilter	ngsfilter -t tagfile.txt -u unident.fastq pairedend_joined.fastq > assigned.fastq	Demultiplexing into samples (40,699,830)
obiuniq	obiuniq -m sample assigned.fastq > assigned_unique.fastq	Dereplicate sequence reads (3,294,811)
obigrep	obigrep -l 10 -p 'count>=10' assigned_unique.fastq > assigned_unique_l10_c10.fastq	Delete sequences with length shorter than 10 and counts lower than 10
obiclean	obiclean -s merged_sample -r 0.05 -H assigned_unique_l10_c10.fastq > assigned_unique_l10_c10_clean.fastq	Selects sequences according to head, singleton, and interval; if a sequence is less than 20 times more frequent it will be assigned to the sequence with the higher count under the requirement that the difference is -d number of differences between the sequences (default 1), if you use option -H head sequences will be selected only (26,665)
sumaclust (only for ITS dataset)	./sumaclust -t 0.97 assigned_unique_l10_c10_clean.fastq > assigned_merged_datasets_sumaclust97.fasta	Clustering of sequences into OTUs using a similarity threshold of 97% between cluster centres and member sequences (5411 cluster created)
obigrep (only after sumaclust)	obigrep -p 'cluster_center' assigned_merged_datasets_sumaclust97.fasta >	Extraction of the cluster centres

	assigned_merged_datasets_sumaclust97_centres.fasta	
Ecotag	ecotag -R database.fasta -d database assigned_merged_datasets_sumaclust97_centres.fasta > assigned_merged_datasets_sumaclust97_centres_database.fasta	Taxonomic assignment of the OTUs with the database (either UNITE, embl or ArctBryo)
obiannotate	obiannotate --delete-tag=explain assigned_merged_datasets_sumaclust97_centres_database.fasta > assigned_merged_datasets_sumaclust97_centres_database_anno.fasta	Adds sequence record annotations, deletes the attribute named “explain”
Obitab	obitab -o assigned_merged_datasets_sumaclust97_centres_database_anno.fasta > assigned_merged_datasets_sumaclust97_centres_database_anno.txt	Change fasta to tabular format

Supplement 6: PCA on the fungal data on 716 OTUs. The samples are displayed in red, while the OTUs are displayed in black.



Supplement 7: PCA on the fungal data on 5466 ASVs. The samples are displayed in red, while the ASVs are marked in black.



Supplement 8: Assessment of the vegetation extraction blanks and NTCs with filtered and unfiltered read counts. We show the total counts and their percentage from the unfiltered dataset as well as from the filtered dataset where all contaminants were removed.

Control	total counts (unfiltered)	% total counts	total counts filtered	% total counts
aLE_NTC_1	0	0	0	0
aLE_NTC_2	2122831	16.12	1577	0.014
aLE_NTC_3	0	0	0	0
LE309_P_EB	0	0	0	0
SO129P_EB	3	0	3	0
SO129P_NTC	1	0	1	0
SO130P_EB	1	0	0	0
SO129P_NTC	4	0	3	0
SO134P_EB	118	0	118	0
SO134P_NTC	0	0	0	0
SO137P_EB	23	0	23	0
SO137P_NTC	15	0	14	0
SO139P_EB	3	0	3	0
SO139P_NTC	6395	0.049	3830	0.035
ELS010_EB	0	0	0	0
ELS010_NTC	0	0	0	0
LS141P.10_EB	68	0	68	0
LS141P.11_NTC	23	0	23	0
LS174P.10_EB	23	0	23	0

LS174P.11_NTC	12	0	12	0
LS167P.03_EB	0	0	0	0
LS180P.11_NTC	41	0	41	0
LS145P.10_EB	73	0	68	0
LS145P.11_NTC	80	0	80	0
LS139P.10_EB	359	0	358	0
LS139P.11_NTC	60	0	60	0
LS181P.10_EB	101	0	101	0
LS181P.11_NTC	110	0	109	0
LS183P.10_EB	0	0	0	0
LS183P.11_NTC	70	0	70	0
LS189P.10_EB	17	0	17	0
LS189P.11_NTC	69	0	68	0
LS191P.10_EB	17	0	17	0
LS191P.11_NTC	102	0	102	0
LS162P.10_EB	117	0	116	0
LS175P.10_EB	55	0	55	0
LS175P.11_NTC	69	0	69	0
LS184P.10_EB	29	0	29	0
LS184P.11_NTC	2	0	2	0
LS163P.10_EB	72	0	72	0
LS163P.11_NTC	8332	0.06	2335	0.02
LS176P.10_EB	119	0	111	0

LS176P.11_NTC	122	0	122	0
LS185P.10_EB	23	0	22	0
LS185P.11_NTC	16	0	15	0
LS166P.10_EB	62	0	61	0
LS166P.11_NTC	67	0	63	0
LS186P.10_EB	292	0	21	0
LS186P.11_NTC	11	0	11	0
LS190P.10_EB	709	0	54	0
LS190P.11_NTC	30	0	29	0
LS167P.10_EB	174	0	174	0
LS167P.11_NTC	118	0	117	0
LS178P.10_EB	109	0	107	0
LS178P.11_NTC	87	0	87	0
LS187P.10_EB	50	0	50	0
LS187P.11_NTC	133	0	132	0
LS193P.12_NTC	106	0	105	0

Supplement 9: ASVs from the vegetation data with abundance in the blanks of more than 10 %. The table shows the scientific name of the ASV and how often it appears in the controls versus the samples with read numbers and percentages.

Scientific name	x in controls	reads control	% in controls	% in samples	reads samples	x in samples
Saxifraga oppositifolia_001	1	392	89.7	10.3	45	3
Convallaria majalis_001	1	112	77.78	22.22	32	4
Saliceae_003	1	180	66.18	33.82	92	26
Apiaceae_001	1	2303	50.78	49.22	2232	11
Menyanthes trifoliata_001	11	47	41.96	58.04	65	27
Pooideae_001	1	1577	24.46	75.54	4871	7
Saliceae_004	1	68	23.05	76.95	227	35
Ceratophyllum demersum_008	3	6	19.35	80.65	31	8
Ceratophyllum demersum_007	6	8	14.81	85.19	54	17
Ceratophyllum demersum_005	5	7	12.5	87.5	49	13
Polygonoideae_002	1	3	12	88	22	11
Myriophyllum sibiricum_001	10	96	11.36	88.64	749	48

Supplement 10: Assessment of the fungal extraction blanks and NTCs with filtered and unfiltered read counts.

Control	total counts (unfiltered)	% total counts	Cont. rem.	% total counts
BH040P.10_EB	40280	0.48	40280	0.49
BH039P.11_NTC	36169	0.43	36169	0.44
BH008P.10_EB	32470	0.39	32470	0.39
BH147P.14_NTC	29299	0.35	29299	0.35
BH149P.09_EB	35579	0.43	29232	0.35
BH035P.11_NTC	14620	0.17	14620	0.18
BH148P.12_EB	13025	0.16	13025	0.16
BH031P.11_NTC	12132	0.15	12132	0.15
BH051P.10_EB	8096	0.10	8096	0.10
BH151P.12_EB	7199	0.09	7196	0.09
BH152P.10_NTC	5274	0.06	5274	0.06
BH151P.04_EB	4676	0.06	4676	0.06
BH033P.11_NTC	5144	0.06	2750	0.03
sample.EB2c	2066	0.02	2066	0.02
BH155P.09_EB	1676	0.02	1676	0.02
BH151P.13_NTC	1232	0.01	1232	0.01
BH164P.11_NTC	1227	0.01	1227	0.01
BH154P.04_EB	1113	0.01	1113	0.01
BH147P.13_EB	430	0.01	422	0.01
BH042P.11_NTC	217	0	217	0

sample.EB1d	57	0	57	0
BH036P.10_EB	15	0	15	0
sample.NTC4	12	0	12	0
sample.EB2d	11	0	11	0
BH037P.11_NTC	10	0	10	0
sample.NTC5	6	0	6	0
BH153P.10_EB	4	0	4	0
BH007P.11_NTC	4	0	4	0
sample.EB1f	4	0	4	0
sample.EB2f	4	0	4	0
sample.EB2e	2	0	2	0
sample.EB3a	2	0	2	0
sample.NTC3	2	0	2	0
BH003P.10_EB	1	0	1	0
BH003P.11_NTC	1	0	1	0
BH032P.11_NTC	1	0	1	0
BH009P.10_EB	1	0	1	0
sample.EB1c	1	0	1	0
sample.EB3b	1	0	1	0
BH010P.10_EB	1	0	1	0
BH147P.10_EB	1	0	0	0
BH040P.11_NTC	0	0	0	0

BH008P.11_NTC	0	0	0	0
BH150P.10_EB	0	0	0	0
BH046P.10_EB	0	0	0	0
BH046P.11_NTC	0	0	0	0
BH007P.10_EB	0	0	0	0
BH149P.10_NTC	0	0	0	0
BH034P.11_NTC	0	0	0	0
BH041P.11_NTC	0	0	0	0
BH042P.10_EB	0	0	0	0
BH047P.10_EB	0	0	0	0
BH047P.11_NTC	0	0	0	0
BH049P.10_EB	0	0	0	0
BH049P.11_NTC	0	0	0	0
BH039P.10_EB	0	0	0	0
BH009P.11_NTC	0	0	0	0
sample.EB1b	0	0	0	0
sample.EB1e	0	0	0	0
sample.EB3e	0	0	0	0
sample.NTC6	0	0	0	0
BH041P.10_EB	0	0	0	0
BH162P.10_EB	0	0	0	0
BH162P.13_EB	0	0	0	0

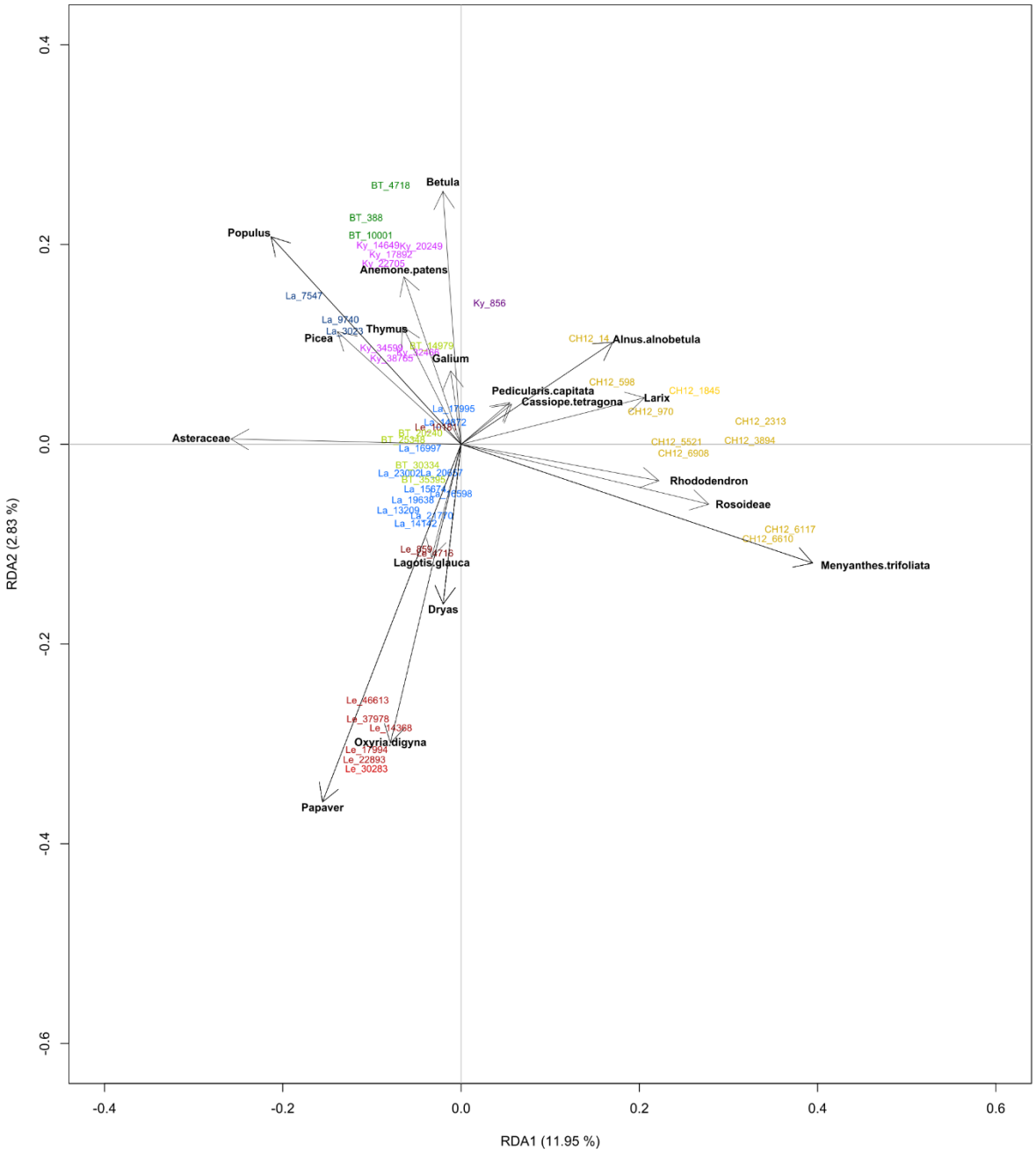
BH162P.14_NTC	0	0	0	0
BH152P.09_EB	0	0	0	0
BH155P.10_NTC	0	0	0	0
BH154P.12_EB	0	0	0	0
BH154P.13_NTC	0	0	0	0
BH153P.18_NTC	0	0	0	0
BH158P.10_EB	0	0	0	0
BH158P.13_EB	0	0	0	0
BH153P.17_EB	0	0	0	0
BH158P.14_NTC	0	0	0	0
BH159P.10_EB	0	0	0	0
BH159P.17_EB	0	0	0	0
BH159P.18_NTC	0	0	0	0
BH150P.17_EB	0	0	0	0
BH150P.18_NTC	0	0	0	0
BH148P.04_EB	0	0	0	0
BH148P.12_EB	0	0	0	0
BH148P.13_NTC	0	0	0	0
BH161P.16_NTC	0	0	0	0
BH165P.14_NTC	0	0	0	0
a _EB1	0	0	0	0
a _EB2	0	0	0	0

a _NTC	0	0	0	0
b _EB2	0	0	0	0
b_NTC	0	0	0	0
g_NTC	0	0	0	0
h_NTC	0	0	0	0
BH032P.10_EB	0	0	0	0
BH035P.10_EB	0	0	0	0
BH033P.10_EB	0	0	0	0
BH036P.11_NTC	0	0	0	0
BH037P.10_EB	0	0	0	0
BH051P.11_NTC	0	0	0	0
BH007P.11_NTC	0	0	0	0
BH031P.10_EB	0	0	0	0
BH034P.10_EB	0	0	0	0
BH010P.11_NTC	0	0	0	0
BH038P.10_EB	0	0	0	0
BH038P.11_NTC	0	0	0	0

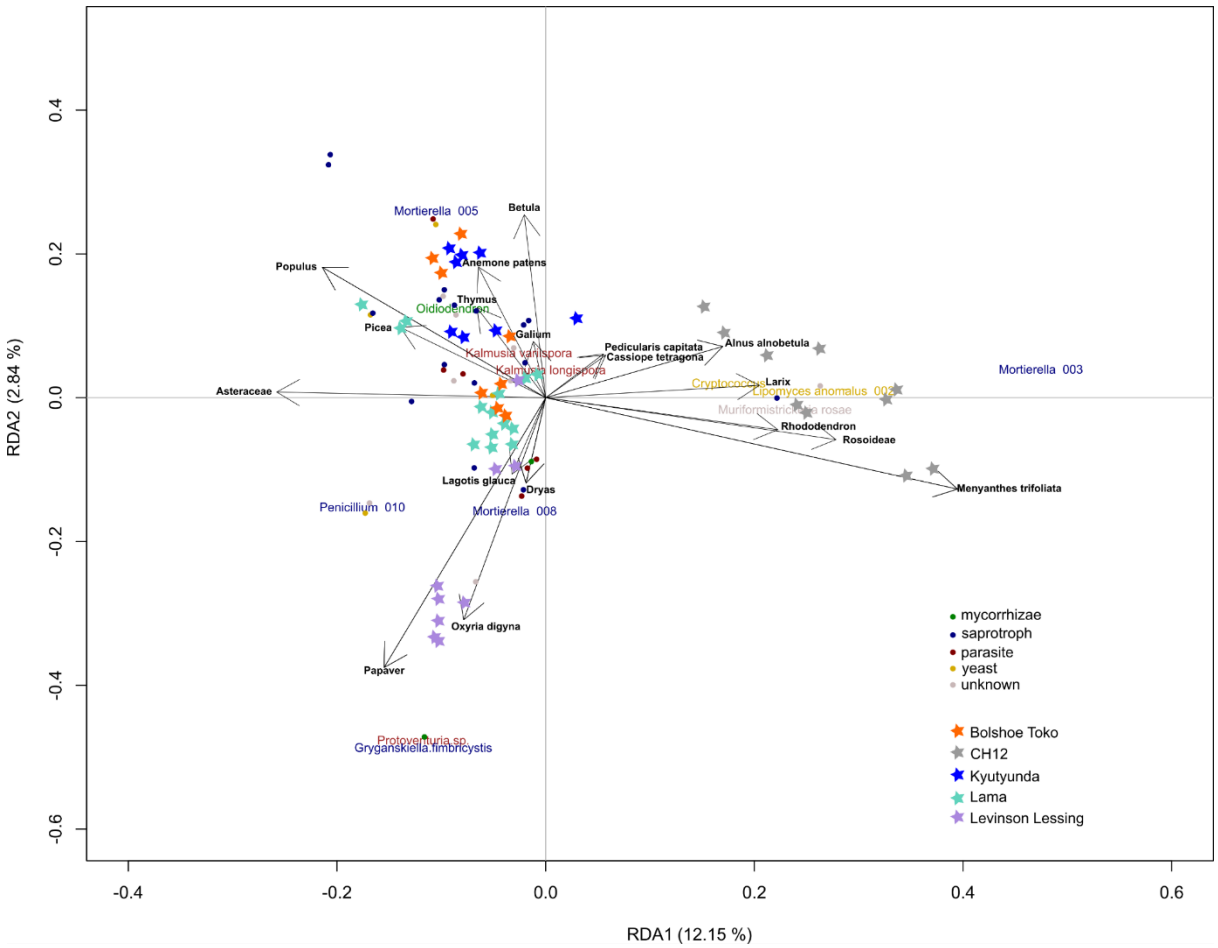
Supplement 11: OTUs from the fungi data with abundance in the blanks of more than 10 %. The table shows the scientific name of the ASV and how often it appears in the controls versus the samples with read numbers and percentages.

Scientific name	x in controls	reads controls	% in controls	% in samples	reads samples	ecology	x in samples
Malassezia restricta_001	3	26732	31.28	68.72	58741	yeast	37
Mortierella sp._001	5	23865	10.17	89.83	210816	saprotroph	92
Venturia hystrioides	1	19095	13.62	86.38	121134	parasite	5
Malassezia restricta_002	2	18773	97.71	2.29	439	yeast	5
Candida parapsilosis	1	13680	98.2	1.8	251	yeast	1
Malassezia globosa_001	1	8183	97.32	2.68	225	yeast	3
Pichia kudriavzevii	1	6597	70.93	29.07	2704	yeast	1
Aspergillus versicolor	1	3950	21.55	78.45	14381	saprotroph	11
Mortierella sp._004	3	2384	62.00	38.00	1461	saprotroph	1
Saccharomyces	1	1674	44.49	55.51	2089	yeast	1
Piptoporus betulinus	1	1375	10.5	89.5	11715	parasite	2
Aspergillus sp. BF8	1	1090	26.4	73.6	3039	saprotroph	2
Trametes versicolor	1	376	18.64	81.36	1641	saprotroph	5

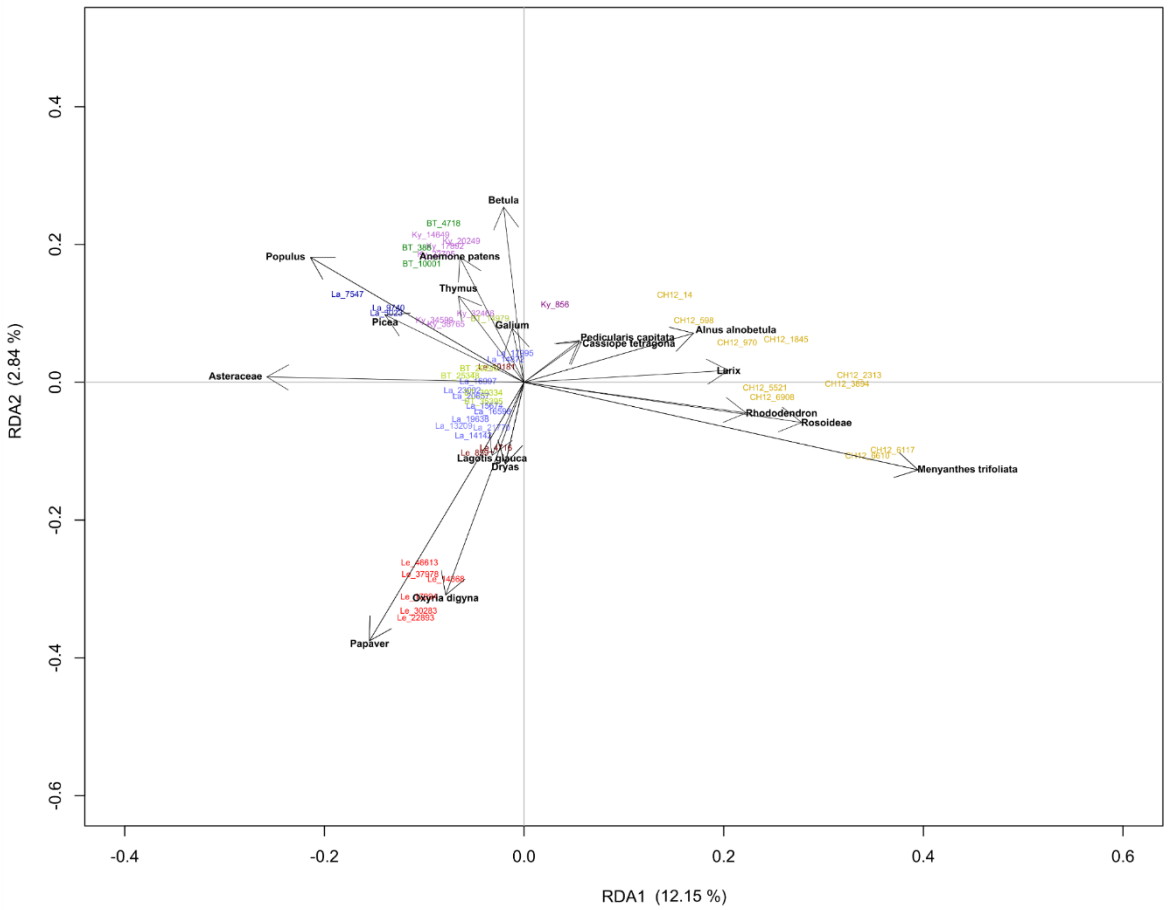
Supplement 12: Supplement to Figure 4 “Fungal and plant co-variation displayed in a redundancy analysis (RDA)”. The sample ages are displayed alongside the vegetation distribution. The samples are colour-coded according to their lake origin as well as their occurrence in a rather forested (darker colour) or tundra (lighter colour) area. The sample names are shortened with the lake name (BT= Bolshoe Toko, Ky = Kyutyunda, La = Lama, Le = Levinson Lessing) and the calibrated year BP.



Supplement 13: Supplement to Figure 4 “Fungal and plant co-variation displayed in a redundancy analysis (RDA)”. In this figure, all OTUs which occur in the blanks were filtered out. The samples are displayed as stars according to their location. The main taxa are colour-coded according to their ecological function.



Supplement 14: Supplement to Figure 4 “Fungal and plant co-variation displayed in a redundancy analysis (RDA)”. In this figure, all OTUs which occur in the blanks were filtered out. The sample ages are displayed alongside the vegetation distribution. The samples are colour-coded according to their lake origin as well as their occurrence in a rather forested (darker colour) or tundra (lighter colour) area. The sample names are shortened with the lake name (BT = Bolshoe Toko, Ky = Kyutyunda, La = Lama, Le = Levinson Lessing) and the calibrated year BP.



References

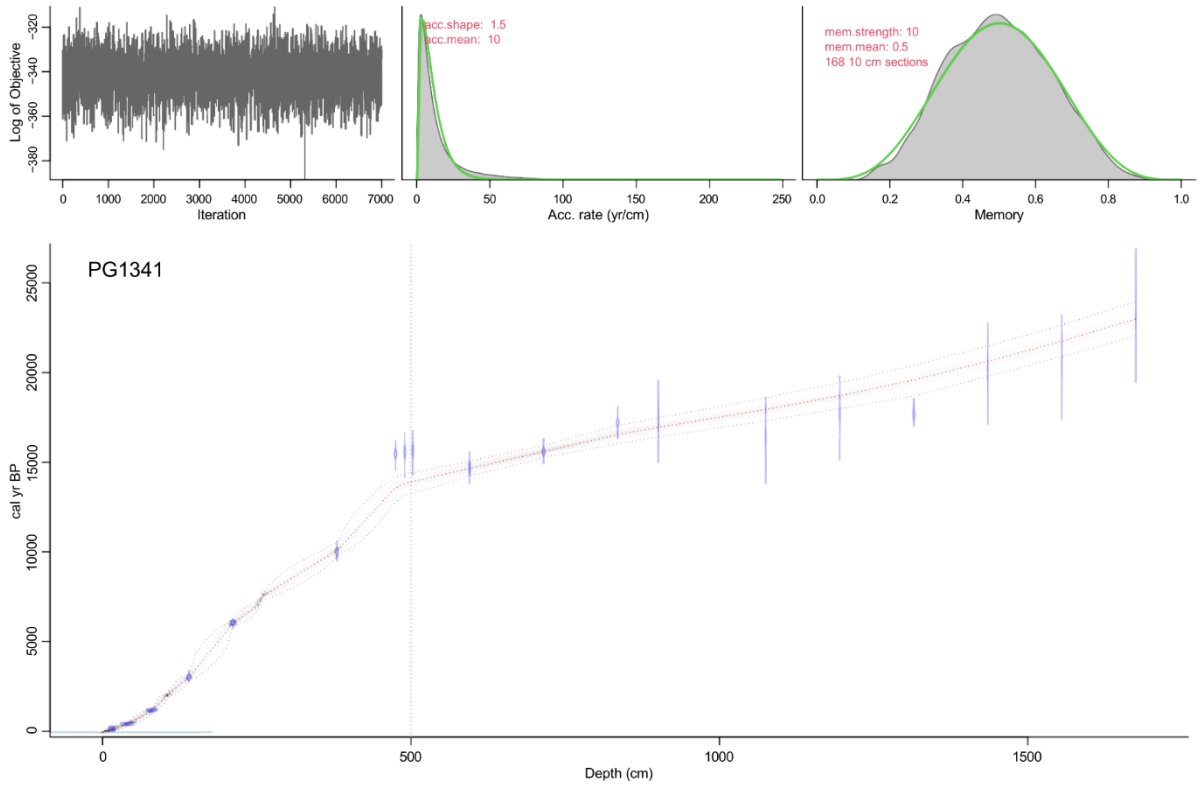
Blaauw, M., Christen, J.A., 2011. Flexible paleoclimate age-depth models using an autoregressive gamma process. *Bayesian Anal.* **6**, 457–474. https://doi.org/10.1214/11-BA618

Reimer, P.J., Austin, W.E.N., Bard, E., *et al.*, 2020. The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757. https://doi.org/10.1017/RDC.2020.41

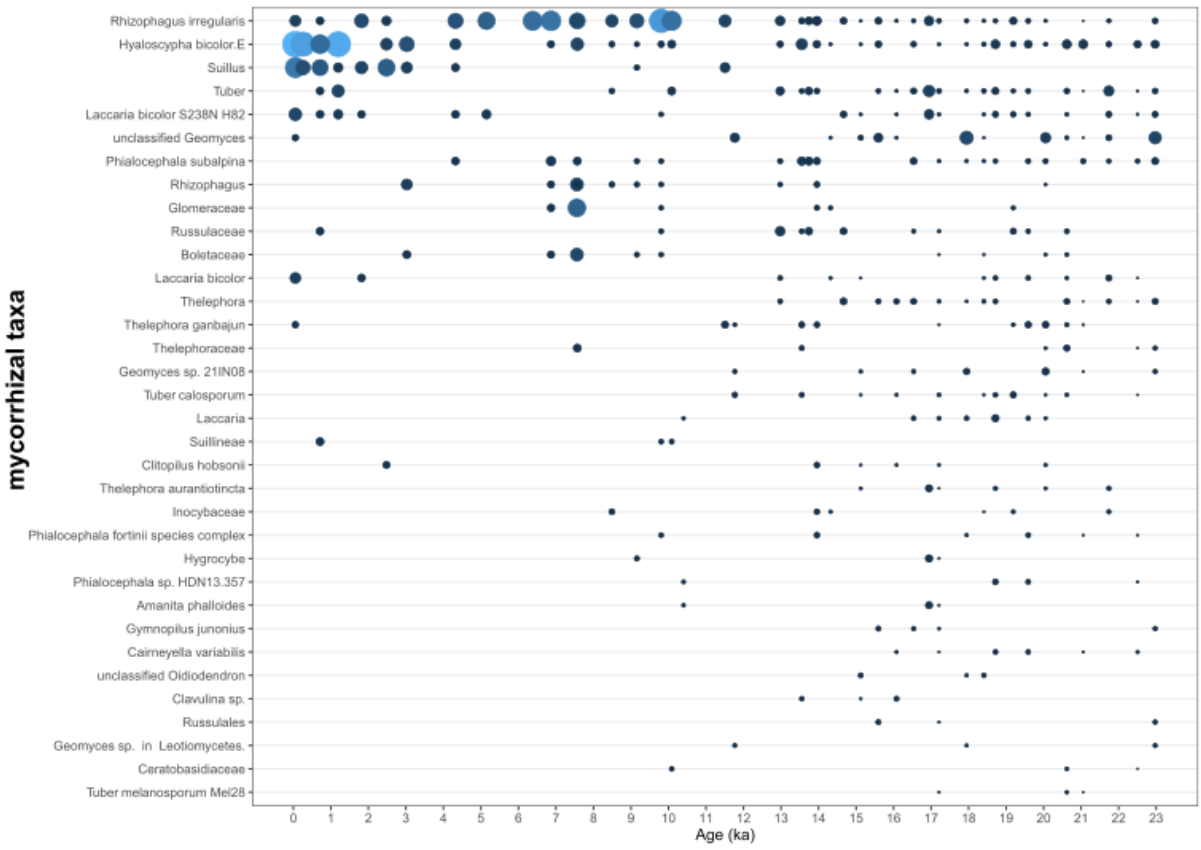
7.3 Appendix to manuscript III

Re-evaluation of the age-depth model of Lake Lama core PG-1341

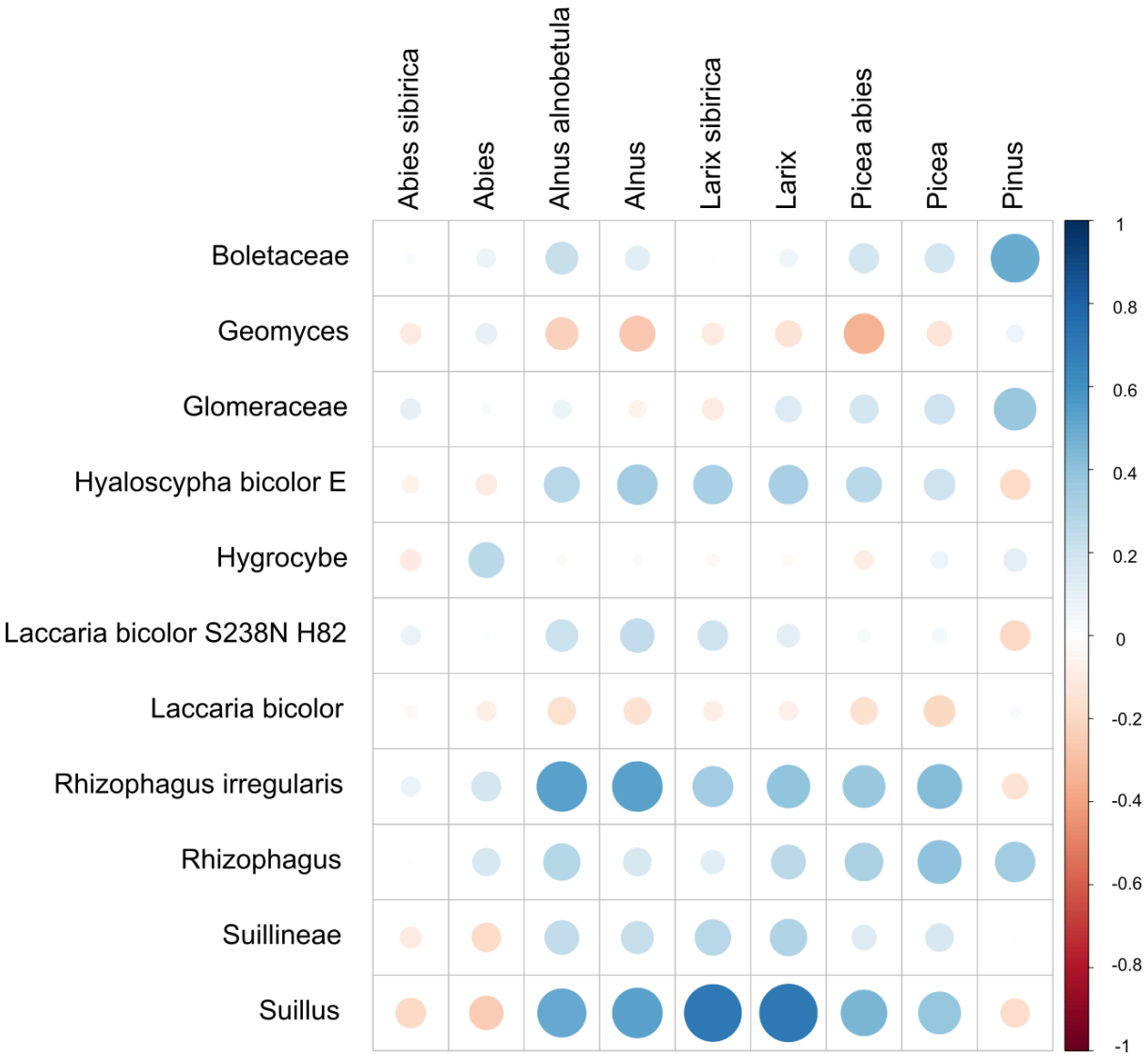
We re-evaluated the age-depth model of the core PG1341, which was previously published in von Hippel et al. (2022). A better correlation of the overlaps using the data for the magnetic susceptibility as well as using the TOC data of Andreev et al. (2004) as comparison to the parallel core yielded a further overlap of 145 cm between the core segments 5 and 6. Therefore, the core is 1.45 m shorter than initially thought. The new age-depth model will be publicly available after acceptance of the manuscript.



Supplement 1: Refined age-depth model of Lake Lama, core PG1341.

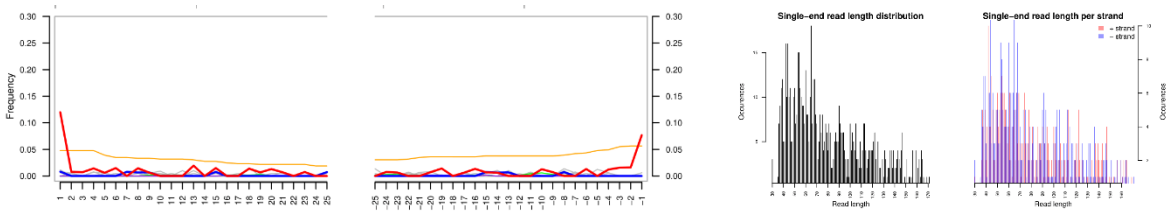
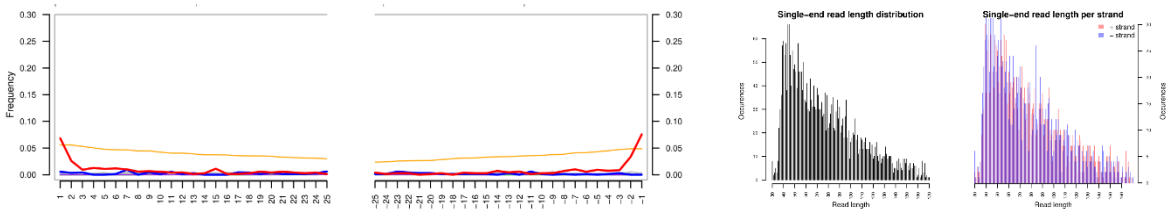


Supplement 2: Mycorrhizal taxa recovered from lake Lama sediment and their respective relative abundance throughout the sediment. The size of the bubbles marks the relative abundance.



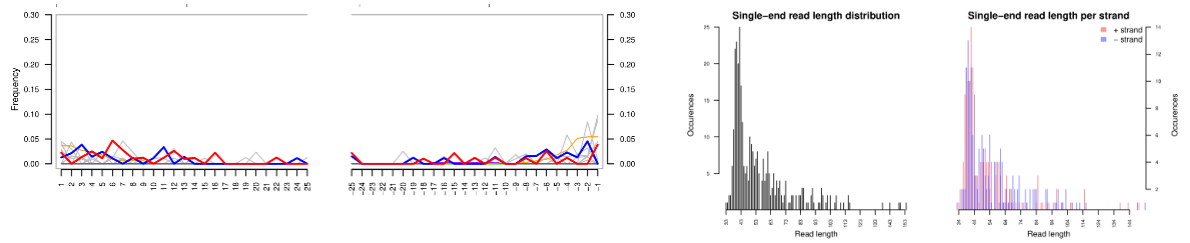
Supplement 3: Co-occurrence analysis between trees and mycorrhizae. Only positively correlated mycorrhizal taxa are displayed. The size and color of the circle represents the degree of correlation.

Supplement 4: Composition and abundance of the bacterial taxa assigned against the Refseq Kraken database (download: 02/2023; https://www.ncbi.nlm.nih.gov/refseq/). Due to the size of the figure, Supplement 4 is provided as a separate file on the CD.

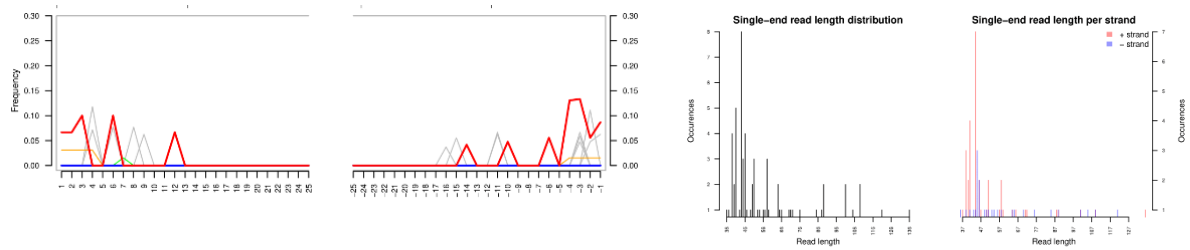
Larix sibirica
 Pleistocene samples (11.5 - 23 cal ka BP; 690 mapped reads)
 Holocene samples (0 - 10.4 cal ka BP; 2.991 mapped reads)

Supplement 5: Post mortem damage patterns for *Larix sibirica* DNA reads. MapDamage plots for overlapping extracted kraken reads aligned against the *Larix sibirica* mitochondrial genome (Accession number MT797187): Left: Misincorporation plots, red: C to T substitutions, blue: G to A substitutions, grey: all other substitutions, orange: soft-clipped bases which are not aligned to the reference and excluded from damage pattern. Right: Read length distributions.

Hyaloscypha bicolor

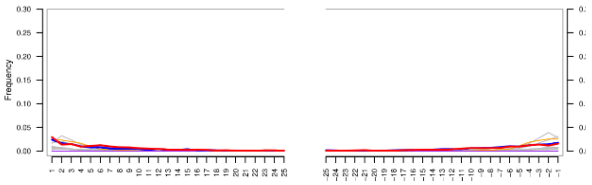
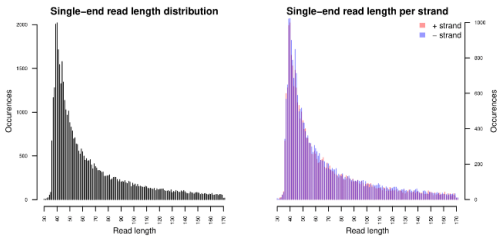
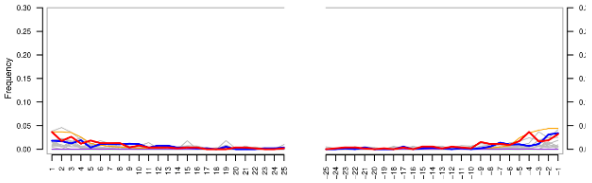
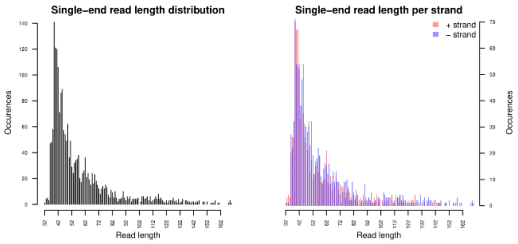


Pleistocene samples (11.5 - 23 cal ka BP; 332 mapped reads)



Holocene samples (0 - 10.4 cal ka BP; 63 mapped reads)

Supplement 6: Post mortem damage patterns for *Hyaloscypha bicolor* DNA reads. MapDamage plots for overlapping extracted kraken reads aligned against the *Hyaloscypha bicolor* genome (Accession number GCF_002865645.1): Left: Misincorporation plots, red: C to T substitutions, blue: G to A substitutions, grey: all other substitutions, orange: soft-clipped bases which are not aligned to the reference and excluded from damage pattern. Right: Read length distributions.

Herminiimonas arsenitoxidans	
	
Pleistocene samples (11.5 - 23 cal ka BP; 44.778 mapped reads)	
	
Holocene samples (0 - 10.4 cal ka BP; 2064 mapped reads)	

Supplement 7: Post mortem damage patterns for *Herminiimonas arsenitoxidans* DNA reads. MapDamage plots for overlapping extracted kraken reads aligned against the *Herminiimonas arsenitoxidans* genome (Accession number GCF_900130075.1): Left: Misincorporation plots, red: C to T substitutions, blue: G to A substitutions, grey: all other substitutions, orange: soft-clipped bases which are not aligned to the reference and excluded from damage pattern. Right: Read length distributions.

Supplement tables: The supplement tables contain information on the metadata of all samples. Besides, they contain a list of all recovered taxa in the subsets. The taxa were checked for contaminants and indicated if applicable. All other taxa yielded the final data subset and were assigned to ecology or element cycle as well as their pH preferences.

Supplement table 1: Metadata of the samples. The table provides information on the sample composition of the sequencing runs and the ages as well as depths of the samples.

Supplement table 2: Recovered plant taxa assigned to their pH preferences.

Supplement table 3: Recovered fungal taxa assigned to their ecology as well as pH preferences

Supplement table 4: Recovered bacterial taxa assigned to their element cycles and pH preferences.

7.4 Manuscript IV

Spatio-temporal microbial associations of plants in cold environments

Status

In preparation

Authors

Barbara von Hippel[1], Kathleen Stoof-Leichsenring[1], Wei Shen[1], Ulrike Herzschuh[1,2,3*]

Affiliations

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

2 Institute of Environmental Science and Geography, University of Potsdam, Germany

3 Institute of Biochemistry and Biology, University of Potsdam, Germany

*corresponding author (Ulrike.Herzschuh@awi.de)

Keywords

ecosystem reconstruction, rhizosphere associations, sedimentary ancient DNA, shotgun sequencing, soil development, soil microbiome

7.4.1 Abstract

Global warming has a major impact on Arctic tundra and taiga ecosystems by causing permafrost thaw and intense forest fires. A further consequence is the warming-induced treeline shift of boreal forests northwards. Alongside treeline migration, also the plant-associated subsoil microorganisms are changing. These northern ecosystems are mostly underlain by permafrost which limits their nutrient supply. Warming leads to deeper thawing-depths, releasing nutrients from previously frozen soil and subsequently, altering the nutrient cycling. This further influences the establishment of plant-microbe associations around the roots. However, the major long-term drivers of such soil community establishment under changing vegetation cover as well as the selectivity of plants towards specific microorganisms remain poorly understood.

We used shotgun sequencing to assess the impact of abiotic and biotic drivers on the establishment of soil communities in permafrost landscapes across Siberia and the Tibetan Plateau. In all locations, the vegetation had the greatest impact on the establishment of fungal as well as bacterial communities, while the temperature had varying influence depending on the bedrock type. We also analyzed patterns in soil microbial ecology across the locations as well as plant-type dependent preferences in soil taxa across the sites. The soil microbiome in the areas with granitic and basaltic bedrock showed greater similarity to each other compared to the sandstone bedrock. The carbon cycling bacteria were to a large extent positively correlated with plant taxa, while sulfur cyclers showed rather little correlations across all cores. In contrast, fungus-plant correlations were generally more site-specific with higher unique local associations of parasites, followed by saprotrophs. All in all, our data help the understanding of healthy plant microbiomes in different soil environments under global warming.

7.4.2 Introduction

The rhizosphere is defined as the dynamic micro-biosphere around the plant roots, where a great variety of microorganisms including fungi and bacteria are responsible for sustaining plant growth and survival (Hiltner, 1904; Hartmann et al., 2008). By excreting organic and inorganic soluble and insoluble metabolites, plants attract and gather microorganisms around their roots (Bais et al., 2006). However, the effect of this so-called rhizodeposition is mainly determined by the release of organic carbon in the soil (Jones et al., 2009b). The contribution of fungi to carbon cycling is mainly by degrading more complex compounds such as lignin (Zhao et al., 2020), while bacteria further utilize and degrade the fungus-derived substrates (Boer et al., 2005).

Besides, plant-growth in terrestrial ecosystems is highly limited by the amount of available nitrogen (Vitousek and Howarth, 1991; LeBauer and Treseder, 2008). In nutrient-poor habitats, mycorrhizal fungus-plant associations in the rhizosphere are inevitable for plant nitrogen supply (Schimel and Bennett, 2004). The nitrogen cycle in the soil is directly impacted by the breakdown of organic matter and the depolymerization of proteins (Jan et al., 2009), as the resulting products are subsequently being used by microorganisms as energy and nutrient sources (Jones et al., 2009a). Furtheron, symbiotic rhizobial bacteria enzymatically fix gaseous nitrogen by converting it to ammonia for better plant-availability (Vitousek et al., 2002). In northern tundra and taiga, the available nutrients are limited by the soil volume, as these ecosystems are typically growing on permafrost ground (Brown et al., 1997) and only nutrients in thawed soil horizons are accessible. Due to reduced decomposition rates resulting from the low temperatures, the overall nitrogen supply to trees is little (Schuur and Mack, 2018). Warming induced thaw of the permafrost enlarges the active layers, increasing the pool of available nitrogen (Salmon et al., 2018). Biological weathering also contributes to the release of nutrients from rock (Uroz et al., 2009) and therefore increases their plant-availability. However, long-term bedrock-dependent nutrient cycling in relation to plant community establishment and the recruitment of a healthy root microbiome have so far not been understood.

Multiple biotic and abiotic environmental factors are known to shape the composition of the rhizosphere taxa. So far, most research focused on understanding parts of the soil community while targeting mostly either bacterial or fungal symbionts. On time scales of a few years, experimental soil warming was found to directly impact the community composition of fungi in the rhizosphere (Solly et al., 2017). For bacterial community composition, a major impact of soil properties like soil texture, water content, and soil type were found, outcompeting the direct influence of biotic factors deriving from the vegetation cover (Vieira et al., 2020). For both bacteria and fungi composition and diversity, the elevation is a known important driver (Zhao et al., 2020; Ren et al., 2021). Nonetheless, the assessment of environmental impacts on either component of these intertwined soil communities is to a large extent lacking. Also, major biotic and abiotic drivers on the long-term establishment of such soil communities remain largely unexplored.

Dynamic processes in the rhizosphere have so far been assessed only on spatial gradients (e.g. Bahram et al., 2013; Silva et al., 2014) or on temporal scales covering a few years (Nuñez et al., 2009; Weber et al., 2019). In such studies, rhizosphere samples are compared to bulk soil to assess the root associated microorganisms. This yields complicated sampling as well as restriction to selected taxa, with the results being biased by, for example, the annual time-point of sample retrieval (Calvaruso et al., 2014), or the plant development stage (Chaparro et al., 2014). Recently, the analysis of sedimentary ancient DNA (sedaDNA) as a proxy to reveal long-term dependencies of microorganisms towards their

plant hosts emerged (von Hippel et al., 2022). Assessing lake sedaDNA offers the unique potential to trace changes in the whole catchment area around the lake without the limitation to a single plant organism. This is possible as the terrestrial DNA from the surrounding of the lake gets continuously transported in the lakes through erosion and further deposits at the bottom of the lake (e.g. Epp et al., 2015; Alsos et al., 2018). So far, the assessment of entire ecosystems from lake sedaDNA has mostly focused on plant-megafauna changes (Wang et al., 2021; Courtin et al., 2022; Kjær et al., 2022). Von Hippel et al. (n.d.) showed for the first time the possibility to reconstruct rhizobial associations as well as soil development relying on lake sedaDNA sequencing data based on bacterial and fungal read counts. Nonetheless, it remains unknown if such reconstructed subsoil microorganisms can be correlated to single plant taxa and if subsequently a reconstruction of single plant rhizospheres is possible.

Here, we assessed the impact of multiple abiotic and biotic drivers on soil development and analyzed dependencies of plants towards specific microorganisms across sites. First, we analyzed shotgun sequencing data from lake sediment cores deriving from three different bedrock areas to assess the drivers of soil community development in the locations. Second, we compared correlations between the soil microbiome and typical plant taxa across the sites. We show that in all sites, vegetation had the greatest impact on the composition of either community, while the impact of time and temperature varied. The taxon-specific microbiome is site-specific with more overlapping taxa shared between higher plants than herbs or grasses. The overlapping soil taxa also shifted from mainly carbon cyclers and yeast being positively correlated with woody taxa in all sites to nitrogen cyclers being dominant positively correlated taxa with herbs across all sites.

7.4.3 Geographical setting and study sites

The studied lakes are located in Siberia, central eastern Russia (Lama and Bolshoe Toko), and on the Tibetan Plateau (Ximencuo). The locations of the lakes are displayed in Figure 1, while further main characteristics are provided in Table 1. The information on the climate data are taken from the Russian Institute of Hydrometeorological Information: World Data Center (2021) if not indicated differently.

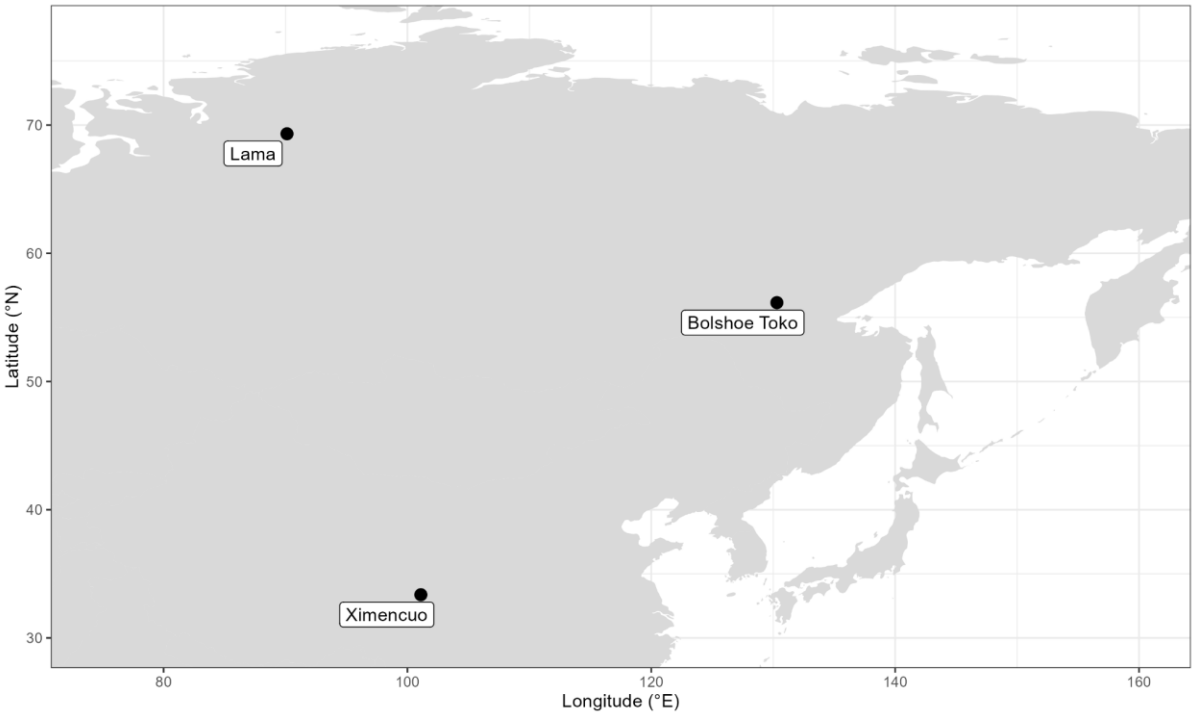


Figure 1: Location of the study sites. The lakes Lama and Bolshoe Toko are located in central Siberia, while the lake Ximencuo is located on the Tibetan Plateau.

Table 1: Main characteristics of the sampled lakes and cores

Lake	Coordinates	Type of vegetation and bedrock	Mean temperature	Dimension	Coring
Lama	69.32 N, 90.12 E; 53 m a.s.l. (Putorana Plateau)	dense taiga with *Picea*, *Larix*, and *Betula*, shrubs such as *Alnus fruticosa*, *Salix*, and *Juniperus communis*, and *dwarf shrubs* (Andreev et al., 2004) bedrock: basalt	July: 13.8 °C January: -28.8 °C Volochanka weather station; 70.97 N, 94.5 E; distance to the lake: 247 km	area: 318 km^2; 80 km x 7 km; maximum depth: 254 m	year: 1997 length: 17.4 depth: 66 m age: 23 cal ka BP (von Hippel et al., 2022 and von Hippel et al., submitted)
Bolshoe Toko	56.15 N, 130.30 E; 903 m a.s.l. (northern slope of eastern Stanovoy Mountain Range	deciduous boreal forests with *Larix cajanderi* and *L. gmelinii* and occurrences of	July: 34 °C January: -65 °C Toko weather station; 56.1 °N, 131.01 °E; distance to the	area: 15.4 km × 7.5 km, maximum depth: 72.5 m	year: 2013 length: 3.8 m depth: 26 m age: 33.8 cal ka

	in southern central Yakutia)	*Picea obovata*, *P. jezoensis*, and *Pinus sylvestris* (Konstantinov, 2000) bedrock: sandstone	lake: 44 km (Konstantinov, 2000)		BP (Courtin et al., 2021)
Ximencuo	33.37 N, 101.10 E; 4030 m a.s.l.	alpine meadows with dwarf shrubs and high alpine cushion and rosette plants (Schlütz, 1999) bedrock: granite	July: 9.8 °C January: -11.4 °C Darlag (Guymai) weather station; 33.79 °N 99.80 °E; distance to the lake: 140 km	area: 50 km^2 maximum depth: 63.3 m	year: 2004 length: 12.8 depth: 50 m age: 39.5 cal ka BP

7.4.4 Material & Methods

7.4.4.1 Sub-sampling of the sediment cores

The cores of Bolshoe Toko and Lama were stored at 4 °C to prevent DNA degradation. Further subsampling of the cores was done in the climate chamber of the German Research Centre for Geosciences (GFZ) where no molecular genetic work is done. The chamber was prepared and cleaned as described by von Hippel et al. (2022). During the sampling, protective clothing as well as face and hair masks were worn to avoid contamination. The surface of the cores was scraped twice with clean knives before sampling. For the sampling itself, four clean knives were used. The samples were placed in sterile 8 ml Sarstedt tubes and kept at -20 °C until DNA extraction.

The samples from Ximencuo were received from a collaboration of Freie Universität Berlin, Germany, and Lanzhou University, China, and further sub-sampled in the dedicated ancient DNA laboratories at AWI Potsdam. The outside of the samples was scraped using sterile scalpels and only the clean inner part was used for extraction.

7.4.4.2 DNA extraction

The DNA was extracted following the manufacturer's instruction, using the DNeasy PowerMax Soil DNA Isolation Kit (Qiagen). To yield higher concentrations of DNA, we added Proteinase K (2 mg mL^{-1}) and DTT (5 M) to the PowerBead solution before adding the samples. Additionally, we incubated the extractions overnight at 56 °C in a rotation oven as described by von Hippel et al. (2022). In total, 96 samples were used for DNA extraction with 26 samples belonging to Lake Bolshoe Toko, 44 to Lake

Lama and 26 to Lake Ximencuo. We concentrated the eluted DNA using the GeneJET PCR Purification Kit (Thermo Fisher Scientific, Germany) using 1 ml of the extracts and concentrated to 50 µl. The contraction of the extracts were determined with a Qubit 4.0 Fluorometer (Thermo Fisher Scientific, Germany) using the Qubit dsDNA BR assay kit. To prevent extensive freeze-thaw cycles, we prepared small aliquots of the DNA. For library built, the concentrations of the extracts were diluted to 3 ng/µl.

7.4.4.3 Single stranded DNA library built

We followed the protocol of Gansauge et al. (2017) for preparing the single-stranded DNA libraries. The ligation of the second adapter was done in a rotating incubator as described by Schulte et al. (2021). Further quantification of the libraries was done using qPCR (Gansauge and Meyer, 2013). The exact protocol was described by Schulte et al. (2021). Then, the libraries were indexed to enable the assignment of the sequencing results. The index PCR reaction is composed of 1x AccuPrime Pfx reaction mix, 2.5 U/µL AccuPrime Pfx Polymerase, 4 µL of P7_X indexing primer (10 µM) and P5_X indexing primer (10 µM), 57 µL of deionized water and 24 µL of the final library. The protocol for the PCR was conducted as follows: 2 min at 95°C, 20 s at 95°C, 30 s at 60°C, 1 min at 68°C and final elongation for 5 min at 68°C. From the qPCR results, we calculated the appropriate number of amplification cycles (steps 2-4) for the following index PCR, which varied between 9 and 14 cycles.

We purified the PCR results using MinElute (Qiagen, Switzerland) following the manufacturer's instructions. The final DNA was eluted in 30 µL elution buffer and the concentration determined with a Qubit 4.0 Fluorometer. Using a TapeStation (Agilent, United States), we performed quality control assessment and also determined the composition of the fragment lengths. 7 sequencing pools were compiled with 3 µL of the libraries and blanks.

We sent the pools APMG-37 and APMG-38 (Lake Lama) to Fasteris SA, Switzerland, where they were run on a NovaSeq device (2x100 bp). The pools for Bolshoe Toko, Ximencuo, and two additional pools for Lama were sent to AWI Bremerhaven, Germany where they were run on a NextSeq 2000 platform (2x100 bp). The sample composition of each sequencing run is part of the supplementary (Supplement Table 1), also including further metadata.

7.4.4.4 Bioinformatic pipeline for the analysis of the sequencing data

The raw sequencing data were quality checked using *fastQC* (version 0.11.; Andrews, 2010) and identical reads were removed (deduplicated) using *clumpify* (BBmap version 38.87, https://sourceforge.net/projects/bbmap/). We merged the paired-end forward and reverse reads with

fastp (version 0.20.1; Chen et al., 2018), including a low complexity filter to filter out reads with low complexity from the dataset. *Kraken2* (Wood et al., 2019) was used for the final taxonomic annotation of the reads against the nt database by NCBI (ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/nt.gz; download: 10/2022 for Lama; download 04/2021 for Bolshoe Toko and Ximencuo, with default k-mer size 35) using a confidence threshold of 0.8. Using the command awk, the kraken report files of each core were converted into a txt-file as further input for R for subsequent data analysis.

7.4.4.5 Data analysis

The analysis of the sequencing data was done in R, version 4.0.3. For further information on the samples, we added metadata including the age and depth of the samples (Supplement Table 1) as well as a file containing the full taxonomic lineage information of the identified taxa based on their taxID. All raw reads for each lake were merged in R.

The initial processing of the data was done as described by von Hippel et al. (n.d.). In brief, for each lake, we created subsets for plants, bacteria, and fungi. For all subsets, we only kept reads of taxa which occurred in at least 3 samples. For the bacteria, only taxa with an overall minimum read count of at least 20 were included. All taxa were checked for their occurrence at the respective location and else filtered out. The vegetation subset for each lake was merged on genus-level before further analysis. For assessing soil development, we worked on genus and species level for the plants.

7.4.4.6 Statistical analysis of the datasets

For the statistical analyses of the datasets, we used R, version 4.0.3. Each dataset was resampled 100 times to the respective base count of the sample mean value of all samples to account for uneven reads, following the GitHub script of Kruse (2019, https://github.com/StefanKruse/R_Rarefaction). The base counts for Lama were 2,556 (plants), 272 (fungi), and 24,285 (bacteria). For Bolshoe Toko, the base counts were 4,575.5 (plants), 247 (fungi), 105,767 (bacteria), and for Ximencuo 305 (plants), 52 (fungi), 59,509 (bacteria). The resampled datasets were used for all further statistical analyses. To account for all variation in the vegetation, we resampled the vegetation data also on species level and used it as the input for the soil development analysis (mean basecount: Lama: 3,533; Bolshoe Toko: 4741; Ximencuo: 349.5)

For the analysis of long-term soil development, we assessed the data as described by von Hippel et al. (n.d.). In brief, “soil development” was defined as either the fungal or the bacterial community. The impact of temperature, vegetation and time on the communities was assessed based on constrained

ordination for a statistical comparison between the respective environmental parameters and the soil community compositions. A redundancy analysis (RDA) was run on the double-square rooted vegetation subset and the significant PC axes were determined using PCAsignificance(). The scores of the first two axes extracted and merged as a dataframe to yield the variable “vegetation”. For the variable “temperature”, we used a reconstruction of temperature variation in the Northern Hemisphere, based on the script of Kruse et al. (https://github.com/StefanKruse/R_PastElevationChange). The respective age of the samples was used as input for the variable “time”.

To assess the correlations between the plants and the soil taxa, we merged the fungi and bacteria as datasets by the respective sample. Each lake was handled separately. The Spearman correlation between the plants and soil taxa was determined using the function cor (package stats; R Core Team, 2021). The resulting correlation matrix was converted in a data frame. Only soil taxa with correlation values of at least 0.4 (respective 0.2 for Table 2) were extracted for each plant taxon. Each co-occurring soil taxon counted as 1 correlation. The correlations of the same fungal ecology or bacterial element cycle were merged to yield overall correlations for the respective soil function. We plotted the final data using ggplot2 (package: tidyverse; Wickham et al., 2019).

7.4.5 Results

7.4.5.1 Compositional changes of representative plant taxa alongside dynamics in fungal ecologies and bacterial element cycling in ancient metagenomic datasets

The 44 samples, 5 extraction blanks, as well as 8 library blanks of lake Lama yielded a total of 1,240,658,912 reads after subsequent filtering. 26 samples, 2 extraction blanks, and 4 library blanks from Ximencuo yielded in 1,675,237,772 reads, while 26 samples, 3 extraction blanks, and 2 library blanks of Bolshoe Toko had 851,075,447 reads. With an identity threshold of 0.8, we recovered overall 1,617 unique terrestrial plant, fungal, and bacterial taxa in the nt database for lake Lama. Ximencuo yielded 861 unique taxa and Bolshoe Toko 1,045. For lake Lama, we worked on 257,710 reads belonging to Viridiplantae being assigned to genus or species level; 3,988,224 reads belonging to bacteria assigned to genus or species level; and 43,040 reads belong to fungi at all taxonomic levels. On the same taxonomic levels, the read counts for Ximencuo yielded 12,134 plant reads, 4,914,223 bacterial reads, and 2,454 fungal reads, while the read counts for Bolshoe Toko yielded 167,353 plant, 6,715,407 bacterial, and 7,220 fungal reads. The composition of the plants, fungi, and bacteria of lake Lama has been described in detail by von Hippel et al. (n.d.). The overall compositional changes are described in the Supplement and Supplemental figures are provided (Supplemental figs 1 - 3).

7.4.5.2 Impact of abiotic and biotic drivers on soil establishment across geographical locations

We assessed the drivers of soil community establishment between the three lakes using variation partitioning constrained ordination which showed distinct patterns for all locations respective bedrock types (Figure 2).

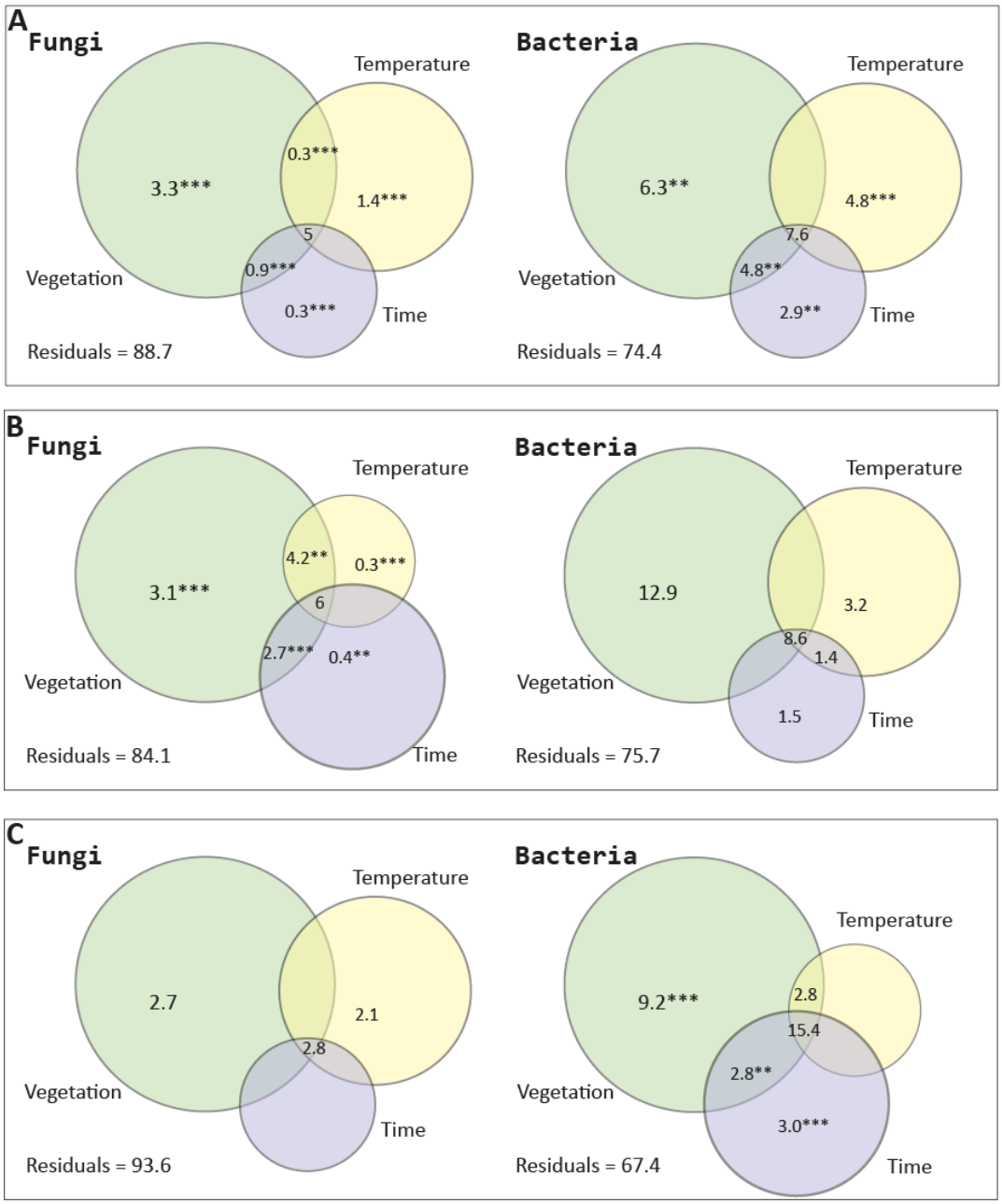


Fig. 2: Variation partitioning of soil fungi and bacteria. The numbers represent the percentage of the respective explained variation (in overlapping areas of the combined variation). A: Lama, B: Bolshoe Toko, C: Ximencuo. We show for all lakes that the vegetation had the greatest unique impact on the establishment of either community. For Lama, the vegetation is followed by temperature, while for Bolshoe Toko and Ximencuo differences between fungi and bacteria were detected. For the bacteria in Bolshoe Toko, the temperature also had the second greatest impact as well as for the fungi in Ximencuo. Vice-versa, for fungi in Bolshoe Toko and bacteria in Ximencuo, time had a greater impact on the community establishment.

For lake Lama, the drivers were described and discussed in detail by von Hippel et al. (n.d.). In brief, the vegetation had the greatest unique impact on the establishment of fungal and bacterial communities, followed by temperature variation, while time after deglaciation had a rather neglectable impact on the establishment of either community (Fig. 2 A). Nonetheless, a great part of the variation remained unexplained.

The establishment of soil communities around lake Bolshoe Toko is also driven by vegetation as the main unique driver (Fig. 2 B). For fungi, the other unique drivers showed neglectable percentages, while a combination of temperature, time, and vegetation also had a great impact on community establishment. For bacterial community establishment, the temperature had a greater impact than the time. Also, for Bolshoe Toko, a great part of the variation remained unexplained.

For Ximencuo, the vegetation also had the greatest unique impact on the establishment of either community out of all tested variables (Fig. 2 C). For the fungi, the temperature also had a greater impact, while time passed since deglaciation was not driving the community establishment. The bacteria in contrast were more driven by the time than the temperature. A large part of the variation here was explained by all combined variables.

7.4.5.3 Relative positive correlations of functional soil taxa with plants across the locations

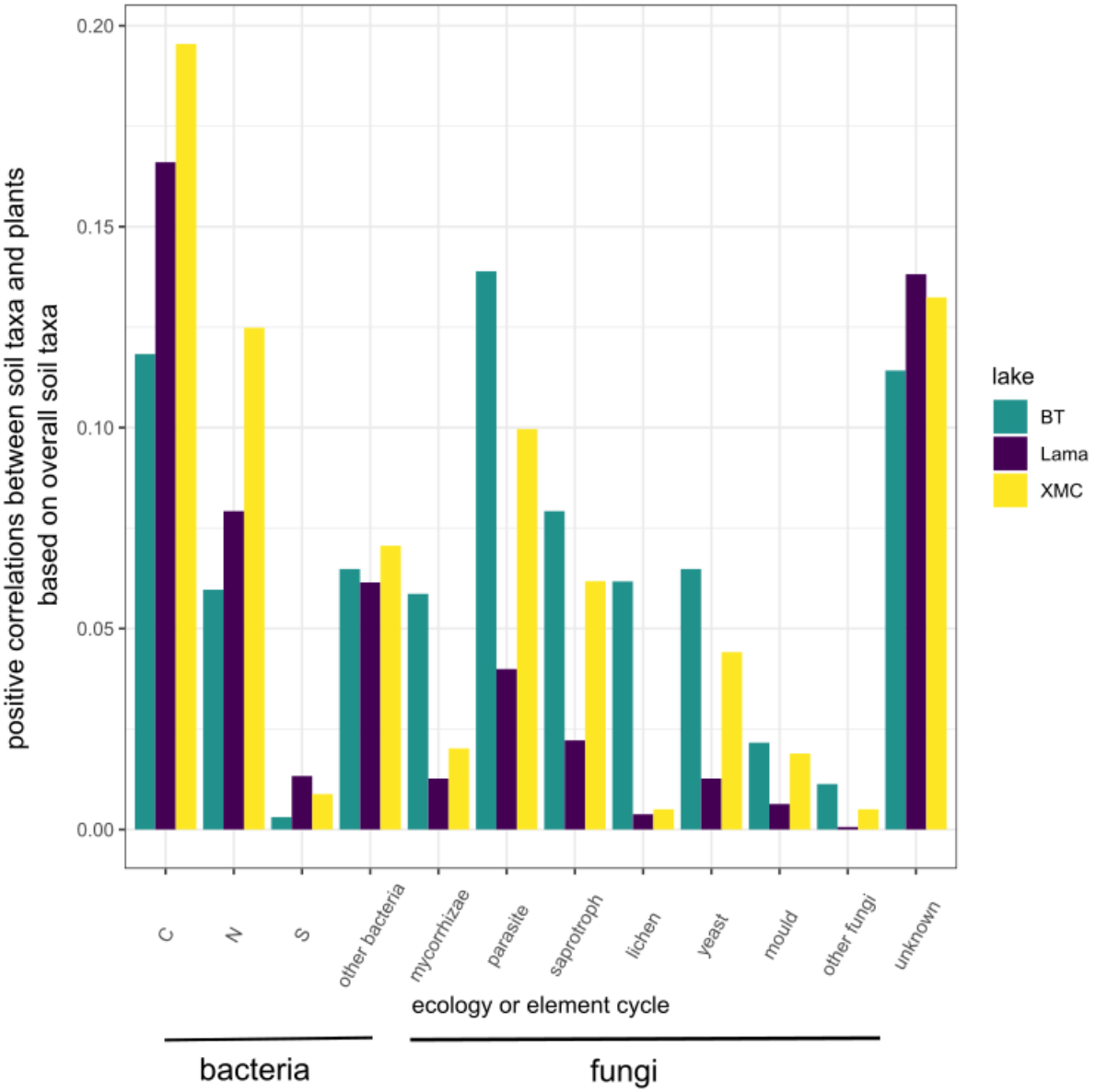


Fig. 3: Soil taxa with positive correlations to plant taxa in relation to the overall occurrence of the respective element cycle or ecological function. The correlations are based on a correlation value of at least 0.4. Compared to the overall occurring taxa of the respective cycle, carbon cyclers show most positive correlations to plant taxa for all cores, followed by nitrogen cyclers. For fungi, most relative positive correlations were detected for parasites, followed by saprotrophs. The names of the lakes are shortened (BT = Bolshoe Toko, XMC = Ximencuo).

We analyzed the soil taxa which are positively correlated with plants for all lakes. This yielded for bacteria in Lama high correlations of carbon cyclers, followed by nitrogen cyclers, while sulfur cyclers showed rather little correlations. For the fungi, the parasites and saprotrophs showed the highest

overall correlations, but also yeast and mycorrhizae were generally highly positively correlated (Fig. 3). Nonetheless, the relative correlations for fungi were lower than for bacteria.

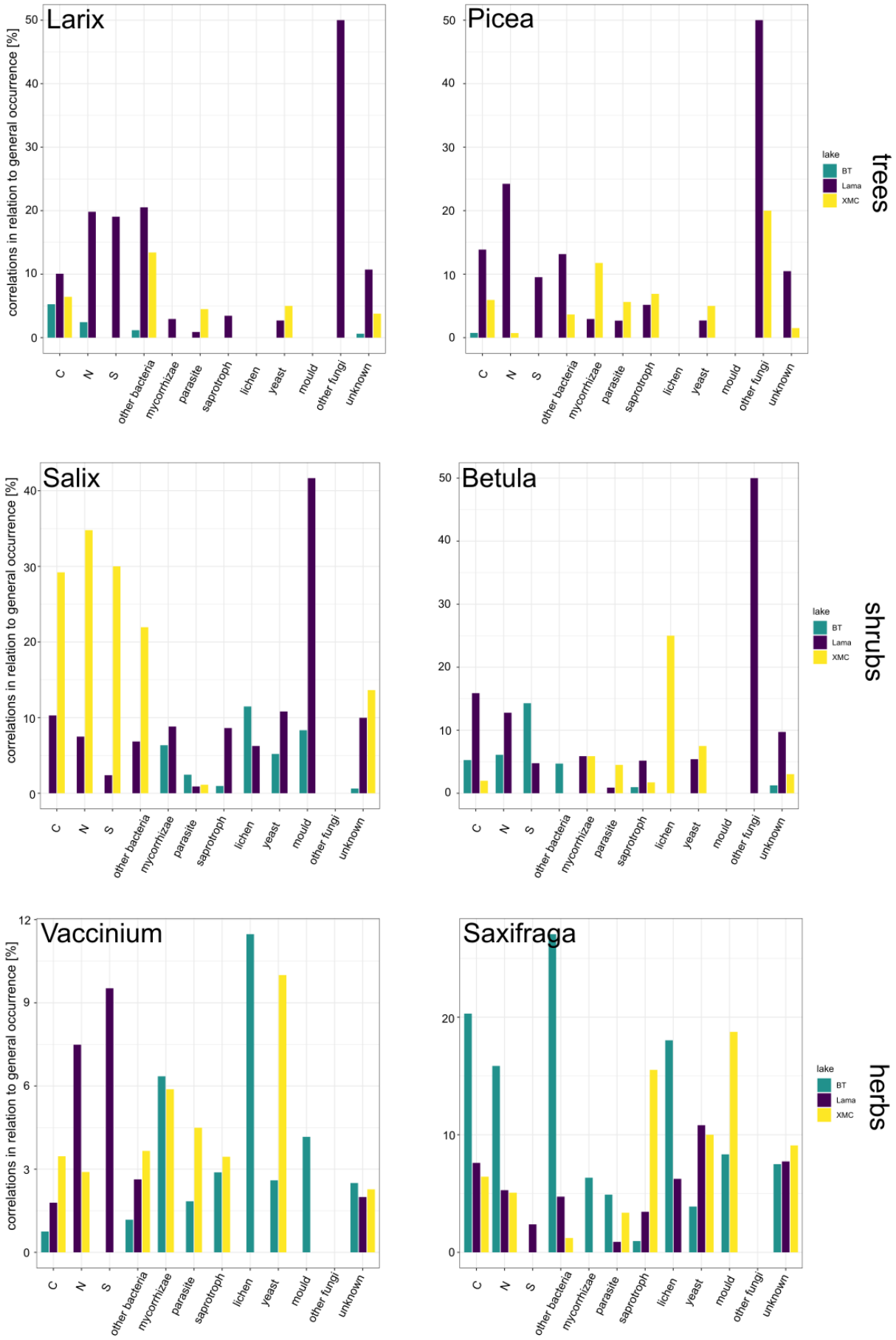
For Bolshoe Toko, the highest correlation for bacteria was found for carbon cyclers, followed by nitrogen cyclers, while sulfur cyclers showed rather neglectable overall correlations. For fungi, the parasites had the highest correlations in the core, followed by saprotrophs and yeast. Compared to the other cores, a higher number of lichen and mycorrhizae correlations was revealed in Bolshoe Toko (Fig. 3).

Ximencuo also had the highest proportion of correlations amongst the carbon cyclers, followed by nitrogen cycling bacteria. For the fungi, the relative correlations followed comparable trends to Lama. Parasites revealed the highest overall proportions, followed by saprotrophs, yeast, and mycorrhizae (Fig. 3).

All in all, Lama and Ximencuo showed similar patterns in their overall positively correlated rhizosphere taxa, while the microbiome in Bolshoe Toko yielded a rather differing pattern.

7.4.5.4 Assessment of the plant taxon-specific microbiome across the locations

We also assessed the microbiomes of typical glacial plant genera which occurred in all three cores (Table 2, Fig. 4). For the trees, we assessed *Larix* and *Picea* as the main boreal tree taxa and *Salix* as well as *Betula* as representative arctic shrubs. *Saxifraga*, which is a typical earlier colonizer after deglaciation (Robbins and Matthews, 2009), and *Vaccinium*c, which is often found as understory of coniferous forests, were the assessed herbs. The grasses *Carex* and *Poa* are also common arctic taxa in glacial landscapes.


Larix
Picea
Salix
Betula
Vaccinium
Saxifraga
correlations in relation to general occurrence [%]
C
N
S
other bacteria
mycorrhizae
parasite
saprotroph
lichen
yeast
mould
other fungi
unknown
lake
BT
Lama
XMC
trees
shrubs
herbs

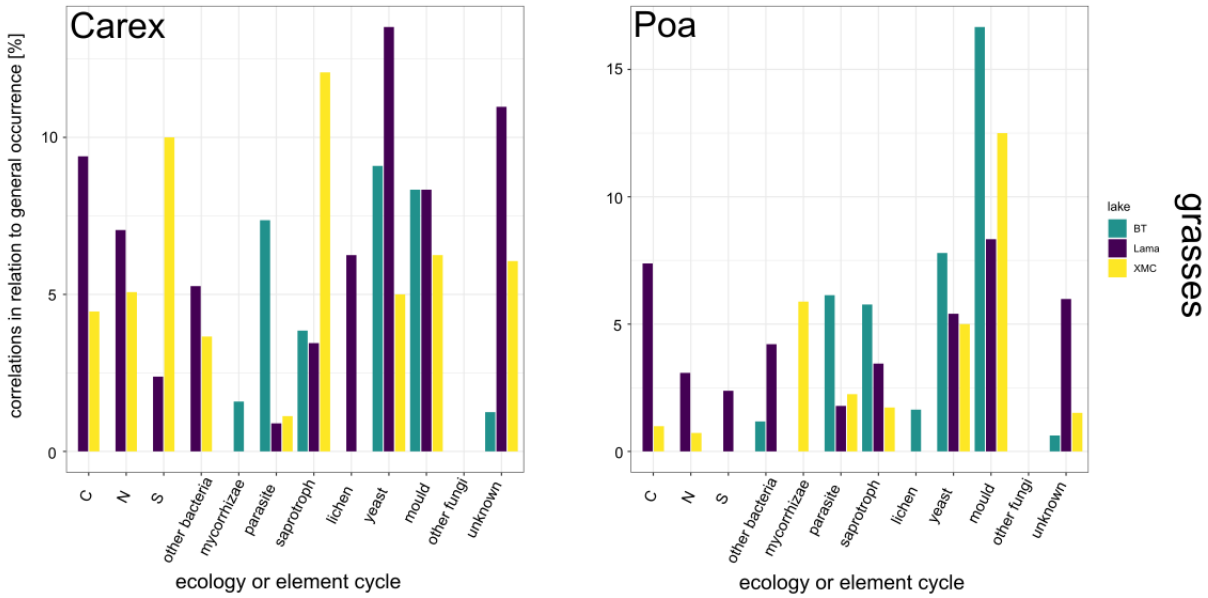


Fig. 4: Plant taxon-specific correlated soil taxa according to their element cycle (bacteria) or ecology (fungi). The assessed plants represent typical glacial taxa for each plant growth form and are abundant in all lakes. Distinct patterns in the correlated soil microbiome for each plant taxon and location were detected. The names of the lakes are shortened (BT = Bolshoe Toko, XMC = Ximencuo).

Table 2: Plant-specific positively correlated soil microbial taxa shared between all locations. The soil taxa in this table are correlated with the respective plant taxa with a rho value of at least 0.2.

Plant	Type	Soil taxa
Larix	tree	*Brettanomyces nanus* (yeast) *Bacillus* (C cycling) *Clostridium* (unknown) *Coniosporium apollinis CBS 100218* (yeast) *Paenibacillus* (C cycling) *Polaromonas sp. JS666* (C cycling)
Picea	tree	*Aplosporella prunicola CBS 121167* (saprotroph) *Bacillus* (C cycling) *Bradyrhizobium icense* (N cycling) *Bradyrhizobium paxllaeri* (N cycling) *Bradyrhizobium sp. CCBAU 051011* (N cycling) *Clostridium* (unknown) *Coniosporium apollinis CBS 100218* (yeast) *Methylocystis rosea* (C cycling) *Monilinia vaccinii corymbosi* (parasite) *Thermothielavioides terrestris NRRL 8126* (unknown)
Salix	shrub	*Brevundimonas* (C cycling) *Trichosporon asahii var. asahii CBS 2479* (yeast)
Betula	shrub	*Caulobacter mirabilis* (C cycling)

		Rhodotorula graminis WP1 (yeast) Saccharomycetales (yeast)
Vaccinium	herb	*Acidovorax antarcticus* (N cycling) *Acidovorax monticola* (N cycling) Burkholderiaceae (N cycling) Glomeraceae (mycorrhizae) Pucciniales (parasite) *Rhodoferax koreense* (unknown) *Variovorax sp. PAM28562* (PGPB) *Variovorax sp. PBL H6* (PGPB) *Variovorax sp. PBS H4* (PGPB)
Saxifraga	herb	*Acidovorax sp. 1608163* (N cycling) *Acidovorax sp. RAC01* (N cycling) *Acidovorax* (N cycling) *Acinetobacter lwoffii* (C cycling) *Acinetobacter* (C cycling) *Hydrogenophaga sp. BPS33* (C cycling) *Hydrogenophaga sp. PBL H3* (C cycling) *Hydrogenophaga sp. RAC07* (C cycling) *Hydrogenophaga* (C cycling) *Malassezia restricta* (yeast) *Nocardioides sp. S5* (N cycling) *Thelephora* (mycorrhizae) *Tuber calosporum* (mycorrhizae)
Carex	grass	*Aspergillus chevaliere* (saprotroph) *Malassezia restricta* (yeast) *Malassezia* (yeast) *Trametes versicolor FP 101664 SS1* (saprotroph)
Poa	grass	*Malassezia restricta* (yeast)

Overall, for *Larix* and *Picea*, similarities in the correlated soil microbial ecologies were detected. For bacteria, the nitrogen cyclers were strongly positive correlated in Lama and Bolshoe Toko, followed by the carbon cyclers (Fig. 4). In contrast, in Ximencuo, the carbon cyclers were prominent for *Larix*, but not for *Picea*, while nitrogen cyclers generally showed rather little correlations (Fig. 4). The highest relative correlation of mycorrhizae for *Larix* and *Picea* was detected in Ximencuo, while no strong correlations were detected in Bolshoe Toko. In Bolshoe Toko, no strong correlation for fungal ecologies were detected for either Pinaceae genus. The assessment of shared correlated soil taxa for *Larix* revealed carbon cyclers (*Bacillus*, *Paenibacillus*) as well as yeast (*Brettanomyces nanus*, *Coniosporium apollinis*). For *Picea*, besides carbon cyclers and yeast, nitrogen cyclers from the genus *Bradyrhizobium* were shared between the cores (Table 2).

For the shrubby taxa *Betula* and *Salix*, the rhizosphere microbiome was rather unique. In Ximencuo, *Salix* showed high correlations with all bacterial cyclers, while *Betula* had rather neglectable correlations. Lichen were strongly positive correlated (Fig. 4). In Bolshoe Toko, *Betula* showed high

correlation values for bacteria, while fungi dominated the correlations in *Salix*. For Lama, the correlations of fungi and bacteria in either plant genus were comparable. For fungi, mycorrhizae and saprotrophs showed prominent correlation values and carbon cyclers for the bacteria (Fig. 4). The assessed soil taxa yielded a rather unique microbiome for all sites with only two taxa overlapping for *Salix* (*Brevundimonas*, C cycling; *Trichosporon asahii var. asahii CBS 2479*, yeast; Table 2) and three taxa shared for *Betula* (*Caulobacter mirabilis*, C cycling; *Rhodotorula graminis WP1*, yeast; Saccharomycetales, yeast; Table 2).

Saxifraga and *Bistorta* showed great differences in their soil microbiome across the sites (Fig. 4). For Ximencuo, yeast and mycorrhizae were strongly correlated with *Vaccinium*, while *Saxifraga* showed greater importance of saprotrophs. The bacteria were dominated by carbon and nitrogen cyclers, while the sulfur cyclers were uncorrelated for these taxa (Fig. 4). In Bolshoe Toko, the herbs had rather unique correlation patterns. Lichen correlations with both plants were strong. Also, mycorrhizal fungi showed importance in the correlation. The bacteria were almost uncorrelated for *Vaccinium*, while *Saxifraga* showed dominance of correlated carbon cyclers, followed by nitrogen cycling bacteria (Fig. 4). In Lama, no strong fungal correlations for *Vaccinium* were detected, while sulfur cycler dominated the correlations, followed by nitrogen and carbon cycler. For *Saxifraga*, yeast and lichen showed stronger correlations, while carbon cyclers were the dominant bacteria for this taxon. The shared soil taxa for *Vaccinium* were dominated by nitrogen cycling bacteria (*Acidovorax*) and plant-growth promoting bacteria (PGPB, *Variovorax*). *Saxifraga* shared *Acidovorax* species (N cycling) and *Hydrogenophaga* (C cycling), but also mycorrhizae (*Tuber calosporum, Thelephora*) were overlapping (Table 2).

For the grass taxa in Ximencuo, the sulfur cyclers were the strongest correlated nutrient cyclers for *Carex*, followed by nitrogen cyclers (Fig. 4). Saprotrophs showed high correlations for the fungi, followed by yeast. *Poa* had low correlations for carbon and nitrogen cyclers, but mycorrhizal fungi and yeast were relatively high correlated. In Bolshoe Toko, the correlations of nutrient cycling bacteria for both taxa were neglectable (Fig. 4). Yeast and parasites were the most prominent correlated fungal ecologies, followed by saprotrophic fungi. In Lama, *Carex* and *Poa* showed generally high correlations for the cyclers, dominated by carbon and followed by nitrogen cyclers (Fig.4). For fungi, saprotrophs showed higher correlations than parasitic fungi, while for *Carex*, lichen were also strongly correlated. On taxon level, the grasses had rather little overlap across the sites (Table 2). *Carex* shared overlapping saprotrophs (*Aspergillus chevaliere*, *Trametes versicolor*) and *Malassezia species* (yeast), while *Poa* had the single overlapping taxon *Malassezia restricta* (yeast).

7.4.6 Discussion

7.4.6.1 Site-specific soil development

We show that the vegetation had the highest unique impact on soil community development at all locations. This underlines previous findings of von Hippel et al. (n.d.), who reconstructed soil development from basaltic bedrock using sedaDNA, stating that long-term soil development is rather environmentally driven than a pure trajectory. Earlier onwards, a great influence of vegetation cover on the formation and transition between soil types was already shown using pollen, spores, and geochemical data from sediments of a peat bog (Willis et al., 1997). In addition to these studies, we demonstrate that independently from the bedrock, long-term soil development is environmentally driven with vegetation as the strongest driver.

For Lama, the temperature had a major influence on either soil community, while time since recent deglaciation was neglectable, suggesting that soil development in the area is not a trajectory (von Hippel et al., n.d.). Surprisingly, in Ximencuo, temperature variation only drove the fungal composition. This indicates differences in general soil development between the sites, implying that granite weathering and - as part of it - soil community establishment is only partially impacted by changing climate. For granitic glacier forefields, a great influence of fungi onto rock weathering through the release of a variety of organic acids was demonstrated (Brunner et al., 2011), while also ubiquitous soil bacteria such as *Bacillus subtilis* were found to highly impact basalt weathering (Song et al., 2007). Lichen as well as mycorrhizal fungi were in general of lower abundance in the core of Ximencuo with slight increase of mycorrhizae during the Holocene. This might lead to weathering changes induced by growing mycorrhizae abundance. Combined with the greater impact of temperature variation on fungal communities as indicated above, this might be an implication for a higher relevance of fungal composition compared to bacteria in granitic weathering.

For Bolshoe Toko (sandstone), the temperature impacted only the establishment of bacteria to some extent. The weathering processes yielding to sandstone breakdown are so far not fully understood, though increased weathering with lowering pH alongside the presence of the bacterium *Acidithiobacillus thiooxidans* has been observed (Potysz et al., 2020). This indicates that global warming induced changes in bacterial community compositions might induce pH lowering of the soil and lead to acidification. Soil acidification is further known to induce increasing root respiration and subsequently alter plant composition (Chen et al., 2015). This implies future plant community changes in the area of Bolshoe Toko with ongoing warming as a consequence of bacterial community changes.

However, for either community at all sites, a large proportion of the variation remained unexplained. For grassland communities, Vieira et al. (2020) stated a higher influence of soil than of plant properties on bacterial community establishment. Together with the findings of von Hippel et al. (n.d.), this

indicates that additional further drivers (e.g. nutrient supply (Jiang et al., 2018) or wetness (Voříšková et al., 2019)) have a great impact on soil development and should be taken into further consideration. Also, the complexity of soil development as a process is being underlined.

7.4.6.2 Differences in the bedrock

The overall higher comparability between the microbiomes of plant taxa in Lama and Ximencuo in relation to Bolshoe Toko indicates a higher similarity of soil properties and, as such, nutrient availability between basalt and granite than sandstone. Basalt is a fine-grained igneous rock with high proportions of plagioclase feldspar, olivine and pyroxene. In contrast, the coarse-grained igneous rock granite is composed of quartz, feldspar, and mica (Le Bas and Streckeisen, 1991; Philpotts and Ague, 2009). Sandstone is a sedimentary rock and not igneous, being primarily composed of quartz with proportions of further minerals such as feldspar or mica. Differences between the nutrient availability and concentrations between the rock types would be expected due to the varying rock compositions. As an example, the different bedrock types differ in their silicate composition and subsequently also the mineral compositions of the soils: Granite contains up to 70 - 77 % of silicate, basalt 45 - 53 % and sandstone 50 - 99 %.

Baek et al. (2020) compared the nutrient availability in soil and forest floor between sandstone and basalt and showed that in basaltic areas, the nutrient concentrations in soil are comparably higher. For sandstone, the nutrient content in forest floor was higher. In relation to our data, this indicates that sandstone is lower in nutrient availability and therefore, nutrient recycling from plant material on the forest floor needs to be more efficient, requiring highly specified bacteria. For Antarctica, a dependency of microbial community structure on the bedrock was shown, mainly derived by varying organic carbon content in the soils (Tytgat et al., 2016). For our data, this would suggest closer similarities in the community between sites with similar organic carbon content.

In a comparison of the carbon content of granitic and basaltic soil, no clear differences were found (Castanha et al., 2012). This explains the greater similarity of the positively correlated microbial ecologies of Ximencuo and Lama due to the similarity in nutrient content between the bedrocks. Recently, with increasing ecosystem complexity and biomass, a decreasing nitrogen abundance in tissue was demonstrated for basalt derived soil (Zaharescu et al., 2019). This supports our data showing the need of a drastic increase in nitrogen cycling with *Picea* invasion in the boreal forests around lake Lama. We highlight a bedrock-specific impact, as such major turnover was not detected in Ximencuo with the invasion of boreal forests in the Holocene.

7.4.6.3 Correlation between the lake biota

7.4.6.3.1 General Trends in positively correlated rhizosphere taxa

We assessed the overall numbers in generally positive correlated rhizosphere soil taxa and compared them within as well as between the cores. For fungi, the highest correlation of taxa compared to the number of plant taxa was found for parasites in all cores, followed by the saprotrophs, while mycorrhizae show relatively little correlation compared to the number of plant taxa especially in Lama and Ximencuo. This underlines previous findings on high host specificities, exclusivity, and adaptation of saprotrophs (Zhou and Hyde, 2001) as well as plant parasites (Jiming Li et al., 2020). Moreover, our data shows that the specificity of parasitic fungi in dense boreal forests with high plant richness such as the area around Bolshoe Toko is more pronounced compared to tundra (Ximencuo) or taiga with lower species richness (Lama) (Fig. 3), suggesting that lower species richness as well as harsher growing conditions force the microbes to become more generalist towards their hosts.

We show relatively little correlated mycorrhizal taxa in comparison to the plant taxa with the highest correlations found for Bolshoe Toko. Mycorrhizal fungi are known to show low host-specificity towards their plant hosts due to the relatively little number of mycorrhizal fungi compared to potential host plants (Smith and Read, 2008). Under varying abiotic conditions, changing mycorrhizal associations have been detected, though nonetheless some generalists were identified (e.g. Deslippe et al., 2011; Leski and Rudawska, 2012; Nguyen et al., 2016). However, our data indicate that bedrock as well as local growing conditions have a more pronounced effect on the fungal composition than the specific plant taxa.

7.4.6.3.2 Plant taxa specific microbiome

We showed that the soil microbiome is site-specific as well as plant taxon-specific as relatively little overlap was revealed. Nonetheless, distinct patterns in overlapping taxa for plant groups were observed with woody taxa sharing mostly carbon cycling bacteria and yeast taxa between the sites. Trees and shrubs have a higher contribution to the organic matter in the rhizosphere compared to herbaceous taxa which derives from the plant leaves or needles, branches and barks (Kögel-Knabner, 2002),. This implies a higher need for specified carbon cyclers with the boreal taxa. Also, this suggests that the cycling of other nutrients and fungal saprotrophs, parasites, and mycorrhizae for these plants are highly site-specific and adapted to the local conditions. The overall percentage of nitrogen being stored in broadleaf trees is higher than in needleleaf trees (Berg and Staaf, 1981). Compared to broadleaf taxa, needleleaf trees also possess a slower initial turnover in biomass (Prescott et al., 2000; Osono et al., 2014), which supports the need of increased nitrogen cycling with the coniferous taxa.

However, our data show that the nitrogen cycling taxa for these taxa are relatively site-unique, indicating that a deficiency in a nutrient requires direct local adaptation.

In contrast, for the herbaceous taxa *Saxifraga* and *Vaccinium*, a variety of overlapping correlating nitrogen cycling bacteria as well as mycorrhizal fungi was detected. Additionally, for *Saxifraga* also few shared correlated carbon cycling bacteria were revealed. In all locations, *Saxifraga* is mainly abundant during the Late Glacial, while *Vaccinium* is predominantly occurring in the late Holocene in boreal forests. Lacking the adapted rhizobial taxa of the boreal forest in the Late Glacial, this indicates the additional need of *Saxifraga* for further specified carbon cycling, underlining the need of herbs for a strong, particular rhizobiome during that time. Aside, the overall content of nitrogen in herbaceous foliage is 30 % higher than in trees (Gilliam, 2007; Muller, 2014). This suggests that these taxa are symbiotic with distinct bacteria across sites for faster recycling and subsequent plant-availability of nitrogen. However, Li et al. (2020) proposed a highly plant-genotype-specific as well as environment-dependent rhizobiome for *Vaccinium* species. Additionally, for the herbaceous taxon *Ranunculus glacialis* (early colonizer) a high impact of the altitude in correlation with other abundant plant taxa on the composition of rhizobiome has been shown (Praeg et al., 2019). This might indicate also that the greater overlap in co-occurring taxa for the early colonizer *Saxifraga* derived from the typical glacial vegetation in all sites during the Late Glacial and therefore similar co-occurring plant taxa.

The grass taxa *Carex* and *Poa* possess almost no overlapping positively correlated taxa amongst the sites. Though, both plant genera showed strong positive correlations with *Malassezia* taxa (yeast) at all locations. Potentially, this is an indication for a less complex rhizobiome of grass taxa which would support the study of Li et al. (2023) who showed a less diverse microbiome in *Poa* rhizosphere than in bulk soil samples. Also, this suggests a site-specific and therefore bedrock-specific rhizosphere adaptation of grass taxa, which due to their shallow rooting depths need to be highly adapted to the local environmental conditions for nutrient supply.

All in all, there is also a general difference between the rooting depths of the assessed taxa with the coniferous trees rooting deepest, followed by shrubs, herbs and grasses. With increasing root depths, nutrients from deeper soil layers can be accessed which are unreachable for shallow rooted-plants (Corre-Hellou et al., 2007). This would underline the need for more specific nitrogen cycling bacteria in herbs with shallow roots to provide sufficient nutrients to the plants. Moreover, Bao et al. (2020) showed that soil properties had a higher impact on microbial communities than the vegetation type. This explains the overall very little overlapping correlations in the soil taxa between the plants and sites due to unique soil properties derived from the different bedrocks.

7.4.7 Implications and future directions

Assessing the drivers of rhizosphere establishment and the rhizosphere microbial composition enables the understanding of healthy root systems. The varying impact of temperature on the establishment of soil communities with changing bedrock suggests differing contribution of fungi or bacteria to the respective functioning soil communities and consequently also for plant establishment. Consequently, under future global warming, this would imply growing differences in and importance of the soil communities.

Assessing modern rhizosphere communities to place the palaeo data in a greater context would be a great asset. The chosen sites possess great differences in their altitudes and bedrocks. Subsequently, also their soil properties and as such nutrient availability would be expected to vary. Therefore, it would be advantageous to assess the rhizosphere of these taxa in modern samples at different locations, also deriving from locations with the same bedrock. A comparison within the bedrock as well as amongst the bedrock types will be a great contribution for the understanding of rhizosphere establishment. Long-term, this might even enable selective seeding with important rhizobial microorganisms to sustain plant growth in a future changing world. Knowing important contributors to a healthy plant rhizosphere might help prevent disease outbreaks and counteract future warming-induced plant damages. Further assessment of the roles of fungi and bacteria impacted by natural long-term warming scenarios will be inevitable also in modern agriculture.

Acknowledgements

We are grateful to Janine Klimke for the help with preparing the libraries of lake Bolshoe Toko.

7.4.8 References

Alsos, I.G., Lammers, Y., Yoccoz, N.G., *et al*., 2018. Plant DNA metabarcoding of lake sediments: How does it represent the contemporary vegetation. *PLOS ONE* **13**, e0195403. https://doi.org/10.1371/journal.pone.0195403

Andreev, A.A., Tarasov, P.E., Klimanov, V.A., *et al.*, 2004. Vegetation and climate changes around the Lama Lake, Taymyr peninsula, Russia during the Late Pleistocene and Holocene. *Quat. Int*., HIgh resolution lake sediment records in climate and Environment variability studies:European lake drilling program **122**, 69–84. https://doi.org/10.1016/j.quaint.2004.01.032

Andrews, S., 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Baek, G., Bae, E.J., Kim, C., 2020. Nutrient stocks of Japanese blue oak (*Quercus glauca* Thunb.) stands on different soil parent materials. *For. Sci. Technol.* **16**, 180–187. https://doi.org/10.1080/21580103.2020.1822924

Bahram, M., Kõljalg, U., Courty, P.-E., *et al.*, 2013. The distance decay of similarity in communities of ectomycorrhizal fungi in different ecosystems and scales. *J. Ecol.* **101**, 1335–1344. https://doi.org/10.1111/1365-2745.12120

Bais, H.P., Weir, T.L., Perry, L.G., *et al*., 2006. The role of root exudates in rhizosphere interactions with plants and other Organisms. *Annu. Rev. Plant Biol*. **57**, 233–266. https://doi.org/10.1146/annurev.arplant.57.032905.105159

Bao, L., Cai, W., Cao, J., *et al*., 2020. Microbial community overlap between the phyllosphere and rhizosphere of three plants from Yongxing Island, South China Sea. *MicrobiologyOpen* **9**, e1048. https://doi.org/10.1002/mbo3.1048

Berg, B., Staaf, H., 1981. Leaching, Accumulation and release of nitrogen in decomposing forest litter. *Ecol. Bull*. **33**, 163–178.

de Boer, W., Folman, L.B., Summerbell, R.C., Boddy, L., 2005. Living in a fungal world: impact of fungi on soil bacterial niche development. *FEMS Microbiol. Rev.* **29**, 795–811. https://doi.org/10.1016/j.femsre.2004.11.005

Brown, J., Jr, O.J.F., Heginbottom, J.A., Melnikov, E.S., 1997. Circum-Arctic map of permafrost and ground-ice conditions. Circum-Pac. Map. https://doi.org/10.3133/cp45

Brunner, I., Plötze, M., Rieder, S., *et al.*, 2011. Pioneering fungi from the Damma glacier forefield in the Swiss Alps can promote granite weathering. *Geobiology* **9**, 266–279. https://doi.org/10.1111/j.1472-4669.2011.00274.x

Calvaruso, C., Collignon, C., Kies, A., Turpault, M.-P., 2014. Seasonal evolution of the rhizosphere effect on major and trace elements in soil solutions of Norway spruce (*Picea abies* Karst) and beech (*Fagus sylvatica*) in an acidic forest soil. *Open J. Soil Sci*. **4**, 323–336. https://doi.org/10.4236/ojss.2014.49034

Castanha, C., Trumbore, S.E., Amundson, R., *et al.*, 2012. Mineral and organic matter characterization of density fractions of basalt- and granite-derived soils in montane California, in: An Introduction to the Study of Mineralogy. *IntechOpen*. https://doi.org/10.5772/36735

Chaparro, J.M., Badri, D.V., Vivanco, J.M., 2014. Rhizosphere microbiome assemblage is affected by plant development. *ISME J*. **8**, 790–803. https://doi.org/10.1038/ismej.2013.196

Chen, D., Wang, Y., Lan, Z., *et al*., 2015. Biotic community shifts explain the contrasting responses of microbial and root respiration to experimental soil acidification. *Soil Biol. Biochem*. **90**, 139–147. https://doi.org/10.1016/j.soilbio.2015.08.009

Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Corre-Hellou, G., Brisson, N., Launay, M., *et al*., 2007. Effect of root depth penetration on soil nitrogen competitive interactions and dry matter production in pea–barley intercrops given different soil nitrogen supplies. *Field Crops Res*. **103**, 76–85. https://doi.org/10.1016/j.fcr.2007.04.008

Courtin, J., Andreev, A.A., Raschke, E., *et al*., 2021. Vegetation Changes in Southeastern Siberia During the Late Pleistocene and the Holocene. *Front. Ecol. Evol*. **9**. https://doi.org/10.3389/fevo.2021.625096

Courtin, J., Perfumo, A., Andreev, A.A., *et al*., 2022. Pleistocene glacial and interglacial ecosystems inferred from ancient DNA analyses of permafrost sediments from Batagay megaslump, east Siberia. *Environ. DNA* **4**, 1265–1283. https://doi.org/10.1002/edn3.336

Deslippe, J.R., Hartmann, M., Mohn, W.W., Simard, S.W., 2011. Long-term experimental manipulation of climate alters the ectomycorrhizal community of *Betula nana* in Arctic tundra. *Glob. Change Biol*. **17**, 1625–1636. https://doi.org/10.1111/j.1365-2486.2010.02318.x

Epp, L.S., Gussarova, G., Boessenkool, S., *et al*., 2015. Lake sediment multi-taxon DNA from North Greenland records early post-glacial appearance of vascular plants and accurately tracks environmental changes. *Quat. Sci. Rev*. **117**, 152–163. https://doi.org/10.1016/j.quascirev.2015.03.027

Gansauge, M.-T., Gerber, T., Glocke, I., *et al*., 2017. Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. *Nucleic Acids Res*. **45**, e79. https://doi.org/10.1093/nar/gkx033

Gansauge, M.-T., Meyer, M., 2013. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat. Protoc*. **8**, 737–748. https://doi.org/10.1038/nprot.2013.038

Gilliam, F.S., 2007. The ecological significance of the herbaceous layer in temperate forest ecosystems. *BioScience* **57**, 845–858. https://doi.org/10.1641/B571007

Hartmann, A., Rothballer, M., Schmid, M., 2008. Lorenz Hiltner, a pioneer in rhizosphere microbial ecology and soil bacteriology research. *Plant Soil* **312**, 7–14. https://doi.org/10.1007/s11104-007-9514-z

Hiltner L, 1904. Über neuere Erfahrungen und Probleme auf dem Gebiete der Bodenbakteriologie unter besonderer Berücksichtigung der Gründüngung und Brache. *Arbeiten der Deutschen Landwirtschaftlichen Gesellschaft* **98**, 59–78.

Jan, M.T., Roberts, P., Tonheim, S.K., Jones, D.L., 2009. Protein breakdown represents a major bottleneck in nitrogen cycling in grassland soils. *Soil Biol. Biochem*. **41**, 2272–2282. https://doi.org/10.1016/j.soilbio.2009.08.013

Jiang, Y., Lei, Y., Yang, Y., *et al.*, 2018. Divergent assemblage patterns and driving forces for bacterial and fungal communities along a glacier forefield chronosequence. *Soil Biol. Biochem*. **118**, 207–216. https://doi.org/10.1016/j.soilbio.2017.12.019

Jones, D.L., Kielland, K., Sinclair, F.L., *et al*., 2009a. Soil organic nitrogen mineralization across a global latitudinal gradient. *Glob. Biogeochem*. *Cycles* **23**. https://doi.org/10.1029/2008GB003250

Jones, D.L., Nguyen, C., Finlay, R.D., 2009b. Carbon flow in the rhizosphere: carbon trading at the soil–root interface. *Plant Soil* **321**, 5–33. https://doi.org/10.1007/s11104-009-9925-0

Kjær, K.H., Winther Pedersen, M., De Sanctis, B., *et al*., 2022. A 2-million-year-old ecosystem in Greenland uncovered by environmental DNA. *Nature* **612**, 283–291. https://doi.org/10.1038/s41586-022-05453-y

Kögel-Knabner, I., 2002. The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter. *Soil Biol. Biochem*. **34**, 139–162. https://doi.org/10.1016/S0038-0717(01)00158-4

Konstantinov, A.F., 2000. Environmental Problems of Lake Bolshoe Toko. *Lakes Cold Environ*., Yakutsk: Resource Use, Ecology and Nature Protection Issue 85–93.

Le Bas, M.J., Streckeisen, A.L., 1991. The IUGS systematics of igneous rocks. *J. Geol. Soc*. **148**, 825–833. https://doi.org/10.1144/gsjgs.148.5.0825

LeBauer, D.S., Treseder, K.K., 2008. Nitrogen Limitation of Net Primary Productivity in Terrestrial Ecosystems Is Globally Distributed. *Ecology* **89**, 371–379. https://doi.org/10.1890/06-2057.1

Leski, T., Rudawska, M., 2012. Ectomycorrhizal fungal community of naturally regenerated European larch (*Larix decidua*) seedlings. *Symbiosis* **56**, 45–53. https://doi.org/10.1007/s13199-012-0164-4

Li, Jiming, Cornelissen, B., Rep, M., 2020. Host-specificity factors in plant pathogenic fungi. *Fungal Genet. Biol.* **144**, 103447. https://doi.org/10.1016/j.fgb.2020.103447

Li, J., Mavrodi, O.V., Hou, J., *et al*., 2020. Comparative analysis of rhizosphere microbiomes of Southern highbush blueberry (*Vaccinium corymbosum* L.), Darrow's blueberry (*V. darrowii* Camp), and rabbiteye blueberry (*V. virgatum* Aiton). *Front. Microbiol.* **11**. https://doi.org/10.3389/fmicb.2020.00370

Li, L., Che, Z., Cao, Y., *et al*., 2023. Analyzing the soil microbial characteristics of *Poa alpigena* Lindm. on Bird Island in Qinghai lake based on metagenomics analysis. *Water* **15**, 239. https://doi.org/10.3390/w15020239

Muller, R.N., 2014. Nutrient relations of the herbaceous layer in deciduous forest ecosystems, in: Gilliam, F. (Ed.), The Herbaceous Layer in Forests of Eastern North America. *Oxford University Press*. https://doi.org/10.1093/acprof:osobl/9780199837656.003.0002

Nguyen, N.H., Vellinga, E.C., Bruns, T.D., Kennedy, P.G., 2016. Phylogenetic assessment of global *Suillus* ITS sequences supports morphologically defined species and reveals synonymous and undescribed taxa. *Mycologia* **108**, 1216–1228. https://doi.org/10.3852/16-106

Nuñez, M.A., Horton, T.R., Simberloff, D., 2009. Lack of belowground mutualisms hinders Pinaceae invasions. *Ecology* **90**, 2352–2359. https://doi.org/10.1890/08-2139.1

Osono, T., Azuma, J., Hirose, D., 2014. Plant species effect on the decomposition and chemical changes of leaf litter in grassland and pine and oak forest soils. *Plant Soil* **376**, 411–421. https://doi.org/10.1007/s11104-013-1993-5

Philpotts, A., Ague, J., 2009. Principles of igneous and metamorphic petrology. *Cambridge University Press* (2nd ed.). https://doi.org/10.1017/CBO9780511813429

Potysz, A., Bartz, W., Zboińska, K., *et al.*, 2020. Deterioration of sandstones: Insights from experimental weathering in acidic, neutral and biotic solutions with *Acidithiobacillus thiooxidans*. *Constr. Build. Mater*. **246**, 118474. https://doi.org/10.1016/j.conbuildmat.2020.118474

Praeg, N., Pauli, H., Illmer, P., 2019. Microbial diversity in bulk and rhizosphere soil of *Ranunculus glacialis* along a high-alpine altitudinal gradient. *Front. Microbiol*. **10**. https://doi.org/10.3389/fmicb.2019.01429

Prescott, C.E., Zabek, L.M., Staley, C.L., Kabzems, R., 2000. Decomposition of broadleaf and needle litter in forests of British Columbia: influences of litter type, forest type, and litter mixtures. *Can. J. For. Res*. **30**, 1742–1750. https://doi.org/10.1139/x00-097

R Core Team, 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ren, C., Zhou, Z., Guo, Y., *et al*., 2021. Contrasting patterns of microbial community and enzyme activity between rhizosphere and bulk soil along an elevation gradient. *CATENA* **196**, 104921. https://doi.org/10.1016/j.catena.2020.104921

Robbins, J.A., Matthews, J.A., 2009. Pioneer vegetation on glacier forelands in southern Norway: emerging communities? *J. Veg. Sci*. **20**, 889–902. https://doi.org/10.1111/j.1654-1103.2009.01090.x

Russian Institute of Hydrometeorological Information: World Data Center, 2021. http://meteo.ru/english/climate/temp.php (last access: 18. Juni 2021)

Salmon, V.G., Schädel, C., Bracho, R., *et al.*, 2018. Adding depth to our understanding of nitrogen dynamics in permafrost soils. *J. Geophys. Res. Biogeosciences* **123**, 2497–2512. https://doi.org/10.1029/2018JG004518

Schimel, J.P., Bennett, J., 2004. Nitrogen mineralization: Challenges of a changing paradigm. *Ecology* **85**, 591–602. https://doi.org/10.1890/03-8002

Schlütz, F., 1999. Palynologische Untersuchungen über die holozäne Vegetations-, Klima- und Siedlungsgeschichte in Hochasien (Nanga Parbat, Karakorum, Nianbaoyeze, Lhasa) und das Pleistozän in China (Qinling-Gebirge, Gaxun Nur). Diss. Bot.

Schulte, L., Bernhardt, N., Stoof-Leichsenring, K., *et al*., 2021. Hybridization capture of larch (*Larix* Mill.) chloroplast genomes from sedimentary ancient DNA reveals past changes of Siberian forest. *Mol. Ecol. Resour*. **21**, 801–815. https://doi.org/10.1111/1755-0998.13311

Schuur, E.A.G., Mack, M.C., 2018. Ecological response to permafrost thaw and consequences for local and global ecosystem services. *Annu. Rev. Ecol. Evol. Syst*. **49**, 279–301. https://doi.org/10.1146/annurev-ecolsys-121415-032349

da Silva, I.R., de Mello, C.M.A., Ferreira Neto, R.A., *et al*., 2014. Diversity of arbuscular mycorrhizal fungi along an environmental gradient in the Brazilian semiarid. *Appl. Soil Ecol.* **84**, 166–175. https://doi.org/10.1016/j.apsoil.2014.07.008

Smith, S.E., Read, D.J., 2008. Mycorrhizal symbiosis, *Academic Press*, Amsterdam Boston (3rd ed. ed.).

Solly, E.F., Lindahl, B.D., Dawes, M.A., *et al*., 2017. Experimental soil warming shifts the fungal community composition at the alpine treeline. *New Phytol*. **215**, 766–778. https://doi.org/10.1111/nph.14603

Song, W., Ogawa, N., Oguchi, C.T., *et al.*, 2007. Effect of *Bacillus subtilis* on granite weathering: A laboratory experiment. *CATENA* **70**, 275–281. https://doi.org/10.1016/j.catena.2006.09.003

Sverdrup, H., Warfvinge, P., 1995. Estimating field weathering rates using laboratory kinetics. *Rev. Mineral. Geochem*. **31**, 485–541.

Tytgat, B., Verleyen, E., Sweetlove, M., *et al*., 2016. Bacterial community composition in relation to bedrock type and macrobiota in soils from the Sør Rondane Mountains, East Antarctica. *FEMS Microbiol. Ecol.* **92**, fiw126. https://doi.org/10.1093/femsec/fiw126

Uroz, S., Calvaruso, C., Turpault, M.-P., Frey-Klett, P., 2009. Mineral weathering by bacteria: ecology, actors and mechanisms. *Trends Microbiol.* **17**, 378–387. https://doi.org/10.1016/j.tim.2009.05.004

Vieira, S., Sikorski, J., Dietz, S., *et al*., 2020. Drivers of the composition of active rhizosphere bacterial communities in temperate grasslands. *ISME J.* **14**, 463–475. https://doi.org/10.1038/s41396-019-0543-4

Vitousek, P.M., Cassman, K., Cleveland, C., *et al*., 2002. Towards an ecological understanding of biological nitrogen fixation, in: Boyer, E.W., Howarth, R.W. (Eds.), The Nitrogen Cycle at Regional to Global Scales. *Springer Netherlands*, Dordrecht, pp. 1–45. https://doi.org/10.1007/978-94-017-3405-9_1

Vitousek, P.M., Howarth, R.W., 1991. Nitrogen limitation on land and in the sea: How can it occur? *Biogeochemistry* **13**, 87–115. https://doi.org/10.1007/BF00002772

von Hippel, B., Stoof-Leichsenring, K.R., Schulte, L., *et al*., 2022. Long-term fungus–plant covariation from multi-site sedimentary ancient DNA metabarcoding. *Quat. Sci. Rev*. **295**, 107758. https://doi.org/10.1016/j.quascirev.2022.107758

von Hippel, B., Stoof-Leichsenring, K.R., Melles, M., Herzschuh, U. n.d. Postglacial bioweathering, soil nutrient cycling, and podzolization from palaeometagenomics of plants, fungi, and bacteria. submitted to *Nature Communications* (06/13/2023)

Voříšková, J., Elberling, B., Priemé, A., 2019. Fast response of fungal and prokaryotic communities to climate change manipulation in two contrasting tundra soils. *Environ. Microbiome* **14**, 6. https://doi.org/10.1186/s40793-019-0344-4

Wang, Y., Pedersen, M.W., Alsos, I.G., *et al*., 2021. Late Quaternary dynamics of Arctic biota from ancient environmental genomics. *Nature* **600**, 86–92. https://doi.org/10.1038/s41586-021-04016-x

Weber, S.E., Diez, J.M., Andrews, L.V., *et al*., 2019. Responses of arbuscular mycorrhizal fungi to multiple coinciding global change drivers. *Fungal Ecol.*, Ecology of Mycorrhizas in the Anthropocene **40**, 62–71. https://doi.org/10.1016/j.funeco.2018.11.008

Wickham, H., Averick, M., Bryan, J., *et al*., 2019. Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686. https://doi.org/10.21105/joss.01686

Willis, K.J., Braun, M., Sümegi, P., Tóth, A., 1997. Does soil change cause vegetation change or vice versa? A temporal perspective from Hungary. *Ecology* **78**, 740–750. https://doi.org/10.1890/0012-9658(1997)078[0740:DSCCVC]2.0.CO;2

Wood, D.E., Lu, J., Langmead, B., 2019. Improved metagenomic analysis with Kraken 2. *Genome Biol. 20*, **257**. https://doi.org/10.1186/s13059-019-1891-0

Zaharescu, D.G., Burghelea, C.I., Dontsova, K., *et al*., 2019. Ecosystem-bedrock interaction changes nutrient compartmentalization during early oxidative weathering. *Sci. Rep.* **9**, 15006. https://doi.org/10.1038/s41598-019-51274-x

Zhao, F., Feng, X., Guo, Y., *et al*., 2020. Elevation gradients affect the differences of arbuscular mycorrhizal fungi diversity between root and rhizosphere soil. *Agric. For. Meteorol*. **284**, 107894. https://doi.org/10.1016/j.agrformet.2019.107894

Zhao, L., Liu, Y., Wang, Z., *et al*., 2020. Bacteria and fungi differentially contribute to carbon and nitrogen cycles during biological soil crust succession in arid ecosystems. *Plant Soil* **447**, 379–392. https://doi.org/10.1007/s11104-019-04391-5

Zhou, D., Hyde, K.D., 2001. Host-specificity, host-exclusivity, and host-recurrence in saprobic fungi. *Mycol. Res*. **105**, 1449–1457. https://doi.org/10.1017/S0953756201004713

7.4.9 Supplement to manuscript IV

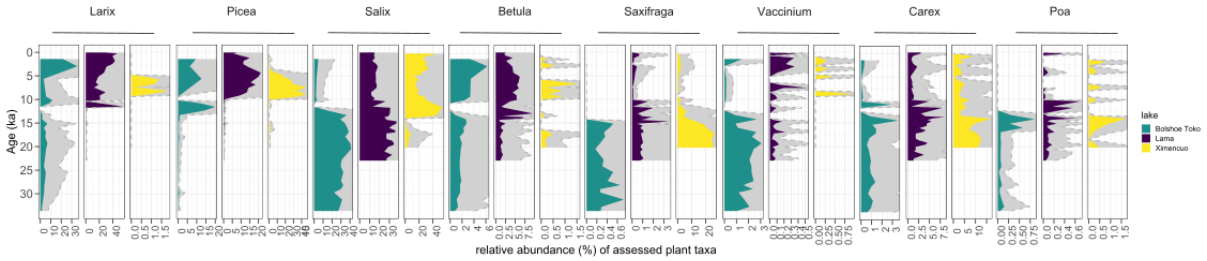
Compositional changes of the plant taxa and soil microbial communities

For Lake Lama, the plant data showed an invasion of *Larix* and *Picea* in the area around the onset of the Holocene (Supplemental Fig. 1). Shrubby *Salix* was present in high abundances throughout the record, while *Betula* abundance increased after the Bølling-Allerød. *Saxifraga* was mainly present until the onset of the Holocene, while *Vaccinium* abundance increased in the late Holocene. *Carex* and *Poa* showed higher abundances until the onset of the Holocene (Supplemental Fig. 1). In lake Bolshoe Toko, *Larix* and *Picea* as well as *Betula* were present throughout the whole record, but showed all a drastic increase in relative abundance around the onset of the Holocene. Shrubby *Salix* was mainly abundant throughout the Late Glacial and declined around the onset of the Holocene. *Saxifraga* and *Vaccinium* were in this core mainly abundant in the Late Glacial (Supplemental Fig. 1). The grass genera *Poa* and *Carex* had their major occurrences before the onset of the Holocene. The record of Ximencuo showed an increase in *Larix* and *Picea* around the onset of the Holocene which was alongside increasing *Betula* abundance. The Pinaceae disappeared after around 5 cal ka BP. *Salix* showed a drastic increase around the Bølling-Allerød and remained broadly abundant. *Saxifraga* was mainly abundant until the onset of the Holocene, while *Vaccinium* appeared in the Holocene. *Poa* and *Carex* were mainly abundant until the Bølling-Allerød warm period around 15 cal ka BP, but remained present throughout the record (Supplemental Fig. 1).

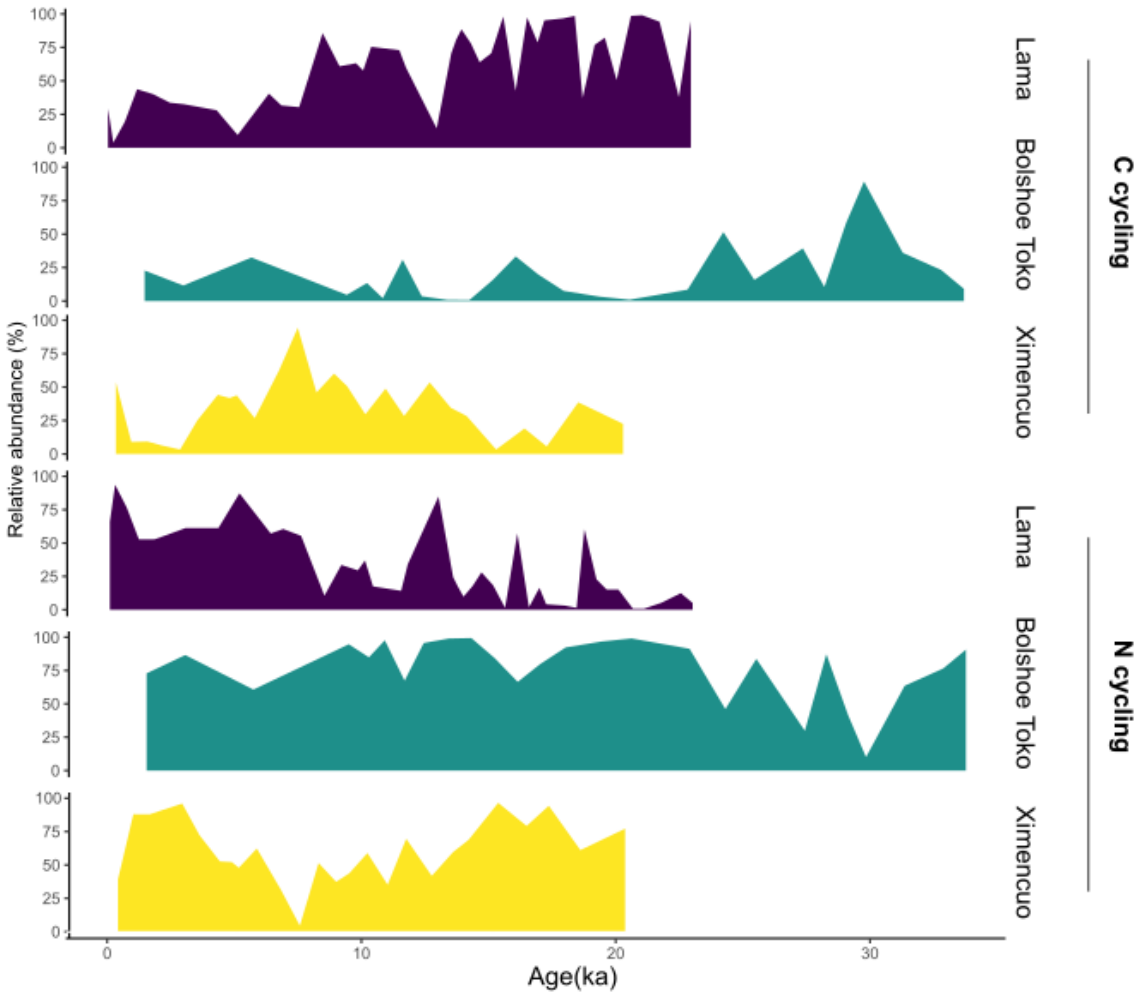
For lake Lama, we showed a relative increase in nitrogen cycling in the Holocene, while the relative abundance of carbon cycling bacteria was declining throughout the core (Supplemental Fig. 2). Overall for Bolshoe Toko, we detected a high relative abundance of N cyclers throughout the record (Supplemental Fig. 2) and peaks in carbon cyclers at 30 cal ka BP before the Last Glacial Maximum. Carbon cyclers were also peaking around the warmer periods of Bølling-Allerød and after the onset of the Holocene. The bacterial data of Ximencuo were dominated by nitrogen cycling taxa throughout the record. Around 7.5 cal ka BP, carbon cycling bacteria showed a drastic peak, coinciding with the peak of the Pinaceae in the record (Supplemental figs 1, 2).

For lake Lama, lichen declined throughout the record, while mycorrhizal taxa were gaining importance with the invasion of Pinaceae (Supplemental figs 1, 3). Parasitic fungi were increasing in their relative abundance after the Bølling-Allerød warm period. Saprotrophic fungi were peaking around the Last Glacial Maximum and in the Late Holocene. Overall yeast abundance was declining with warming (Supplemental Fig. 3). For fungal dynamics in Bolshoe Toko, we detected a drastic decrease in lichen with the increase in Pinaceae. Mycorrhizae were peaking around the Last Glacial Maximum and in

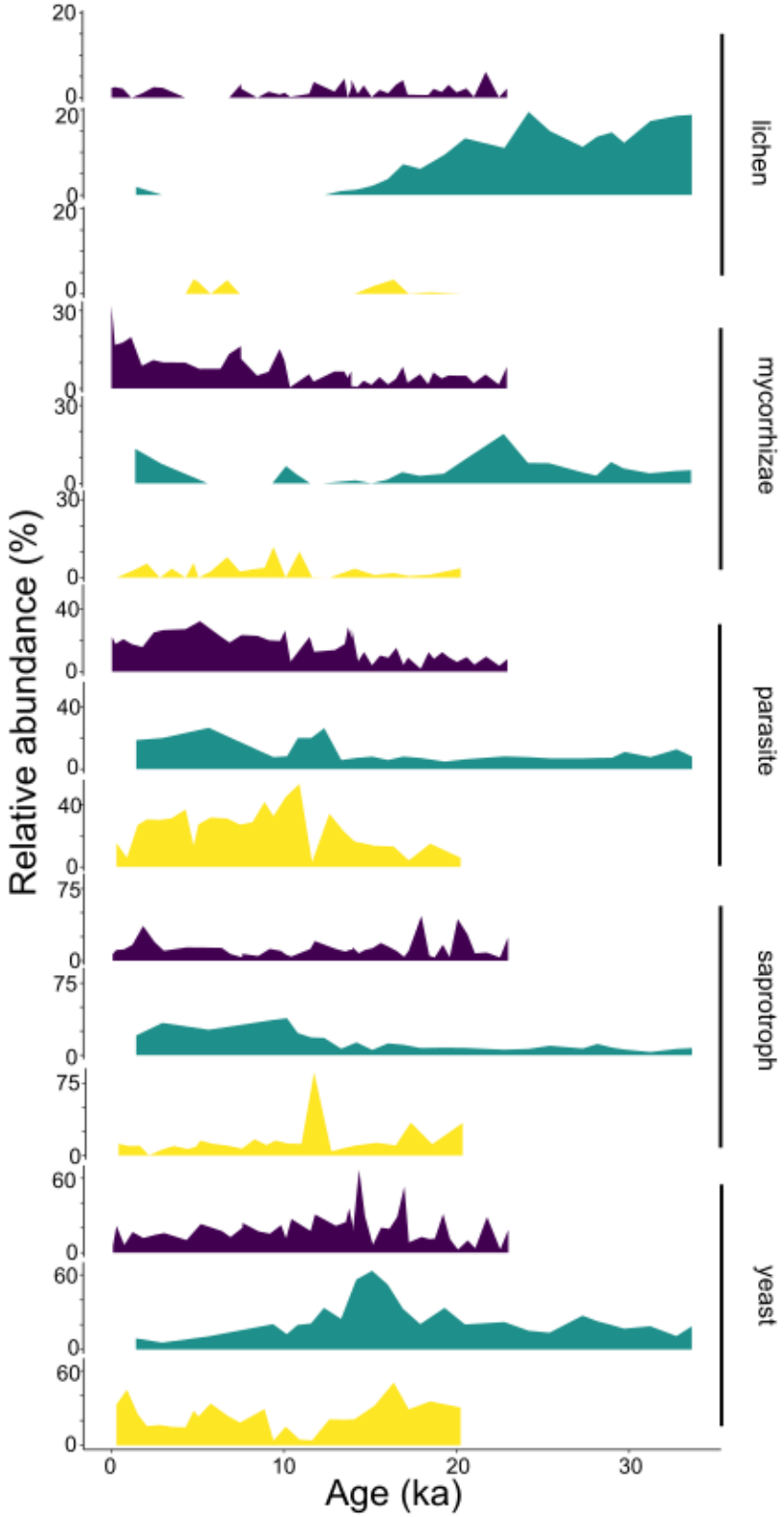
warmer periods of the Holocene. Parasitic fungi as well as saprotrophs showed an increase with the expansion of Pinaceae alongside warming with the onset of the Holocene (Supplemental figs 1, 3). Yeast showed a drastic peak around the Bølling-Allerød and decreased with warming in the Holocene (Supplemental Fig. 3). The fungi record of Ximencuo showed lichen peaks around the Last Glacial Maximum as well as during warmer phases of the Holocene. Mycorrhizae were increasing in their relative abundance after the onset of the Holocene. Parasitic fungi were highly abundant throughout the record with increasing relative abundance during the Holocene. Saprotrophic fungi showed a drastic peak after the Bolling Allerod warm period before the onset of the Holocene. Yeast revealed generally high relative abundances throughout the core, with a decline around the onset of the Holocene alongside Pinaceae invasion (Supplemental Fig. 3).



Supplemental figure 1: Dynamics in the relative abundances of the assessed plant taxa in the cores. The assessed plants represent typical glacial taxa for each plant growth form and are abundant in all lakes. Most assessed plants show comparable patterns in their occurrences for all lakes from the Late Glacial to the Holocene.



Supplemental figure 2: Dynamics in the relative abundances of carbon and nitrogen cycling bacteria in the cores. Lake Bolshoe Toko and Ximencuo were overall dominated by nitrogen cycling, while in Lama, carbon cycling was prominent throughout the Late Glacial. The shift to nitrogen dominance occurred around 7.5 cal ka BP.



Supplemental Figure 3: Dynamics in the relative abundances of fungal ecological functions in the cores. An increase in the relative abundance of parasitic fungi with warming was detected for all cores, while yeast rather declined. All other functional ecology groups showed unique patterns for each core.

Supplemental table 1: Metadata of the samples. The table includes information on the sample composition of the sequencing runs and the ages as well as depths of the samples.

Acknowledgements

First of all, I would like to thank Prof. Ulrike Herzschuh for giving me the opportunity to conduct such fascinating research over the last 3 years, the fruitful discussions, and the possibility to conduct field work in Central Yakutia. Second, I am grateful to Dr. Kathleen Stoof-Leichsenring for giving me the chance to join her research group and for supervising me in all matters concerning lab work and bioinformatics.

I am more than happy that I got to share the office A43-210 for some time with my two friends and colleagues Philip and Ramesh whose support and encouragement carried me through until the finish line. Also, a big thank you to Luise not only for proof-reading this thesis, but whose friendship and R skills helped me a lot in the struggles of my first PhD year.

I also owe gratitude to Claudia Sprengel and Claudia Hanfland from the POLMAR graduate school for their help in most organisational issues regarding my PhD.

I am grateful for the support of many more people at AWI, mainly from the Biodiversity group, especially Sarah, Janine, Amelie, Moein, Amedea, Jérémy, Stefan, Simeon, Izabella, Ugur, Robert, Sarah, and Josefine, and the AWI Basketball crowd for making me laugh and run on Wednesday nights.

I would also like to thank the supervisors of my Master thesis, Dr. Kaya Bork and Prof. Rüdiger Horstkorte, for showing me the beauty in science and PCR and encouraging me to start a PhD.

Starting a PhD with the pandemic made the move to Potsdam not easy which is why I am more than grateful to my friends Mia, Quitterie, and Lea who helped me survive the lockdowns with dinner nights and supported me when the stress was overwhelming. Also, I am grateful to Sarah and Christina for all the chats and to Flor for our friendship back from Argentina, helping me to get settled here.

I am lucky to have made many more friends in Potsdam, supporting me especially in the last bits of the PhD with open arms, food, distractive sport sessions and laughters: thank you to Katha, Wiki, Tobi, Johanna, all bouldering friends from 7a+, and my Ultimate frisbee team.

For the very last time, I want to thank my good friend Pascal who did not only support me during the hard years of studying Biochemistry and pursuing a PhD, but most important, who never stopped believing in me and my skills, even when I myself did. A big hug to many more people for your supportive friendship during the last years: Juli, Sophie, Clara, Lene, Michel, Corinna, Basti, Lisa, Farina, Pia, Finn, Linus, Inge, Julia, Thea – I am more than lucky to call you my friends.

Last but not least, I am grateful to my family and their never-ending support and encouragement in the last three years and to my siblings Katharina, Johannes, and Andreas for always making me laugh.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit mit dem Titel „Long-term bacteria-fungi-plant associations in permafrost soils inferred from palaeometagenomics" selbstständig und unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe. Wörtlich oder sinngemäß übernommenes Gedankengut habe ich als solches kenntlich gemacht. Diese Dissertation wird erstmalig an der Universität Potsdam eingereicht. Die dem Verfahren zu Grunde liegende Promotionsordnung ist mir bekannt.

Potsdam, 29.06.2023

Barbara von Hippel

Damage pattern analysis – Auflagen Doktorarbeit

Summary

The analysis of the DNA damage supports that the main patterns which were assessed in the manuscripts III and IV of this thesis are of ancient origin. The comparison between the datasets HOPS and Kraken yielded overall similar temporal patterns for the most prominent taxa, indicating that either pattern allows a valid interpretation of temporal ecological trends. We compared the datasets as the HOPS pipeline allows to distinguish between ancient and default reads. We found damage patterns for all three kingdoms i.e. plants, fungi, and bacteria with the strongest patterns for plants. More detailed analyses are required to also better understand differences in DNA damage between the assessed kingdoms.

Main

Proofing the ancient origin of metagenomics derived DNA reads is necessary to differentiate reads from modern contaminants as well as to ensure that the ancient organisms did not continuously survive and replicate in the sediment. To do so, a variety of bioinformatic tools has been established in recent years, including the HOPS pipeline (Huebler et al., 2019) and mapdamage (Jonsson et al., 2013). Providing the damage pattern analysis of ancient DNA reads in scientific publications has so far been well established for plants, while research on microorganisms including bacteria and fungi is lacking. In this thesis, I included for the first time an analysis of compositional trends of ancient fungal and bacterial DNA reads derived from shotgun metagenomics (von Hippel et al., unpublished a, b). I assessed their DNA damage using mapdamage (see dissertation Barbara von Hippel), showing that plants provide the typical damage pattern, while the assessment of fungi and bacteria is rather complex. With the help of Dr. Kathleen Stoof-Leichsenring and under supervision of Ulrike Herzschuh, I also assessed the damage using the HOPS pipeline to compare the results. Here, Kathleen Stoof-Leichsenring established the bioinformatic implementation of the HOPS pipeline and the initial scripts for data analysis.

For the damage pattern analysis of the reads, we re-analysed the raw data using the HOPS pipeline. We assessed the outputs of the HOPS and the kraken pipeline for all three subsets. The temporal patterns of the community compositions showed comparable trends between the HOPS and the kraken pipeline for the prominent taxa (Figure 1 and 2). We assessed the composition of the selected taxa scaled up to 100 % for the comparison of the dataset of the two reference databases. The

comparison between the HOPS_damage (at least one damage lesion in the first five bases from either end of the read) and HOPS_default (all reads that fulfill the filtering criteria) reads datasets yields overall similar patterns for the assessed taxa. The comparison between the HOPS_default dataset and the kraken dataset revealed overall comparable trends between the Late Glacial and the Holocene. However, the magnitude of relative abundances varied for some taxa between the used pipelines.

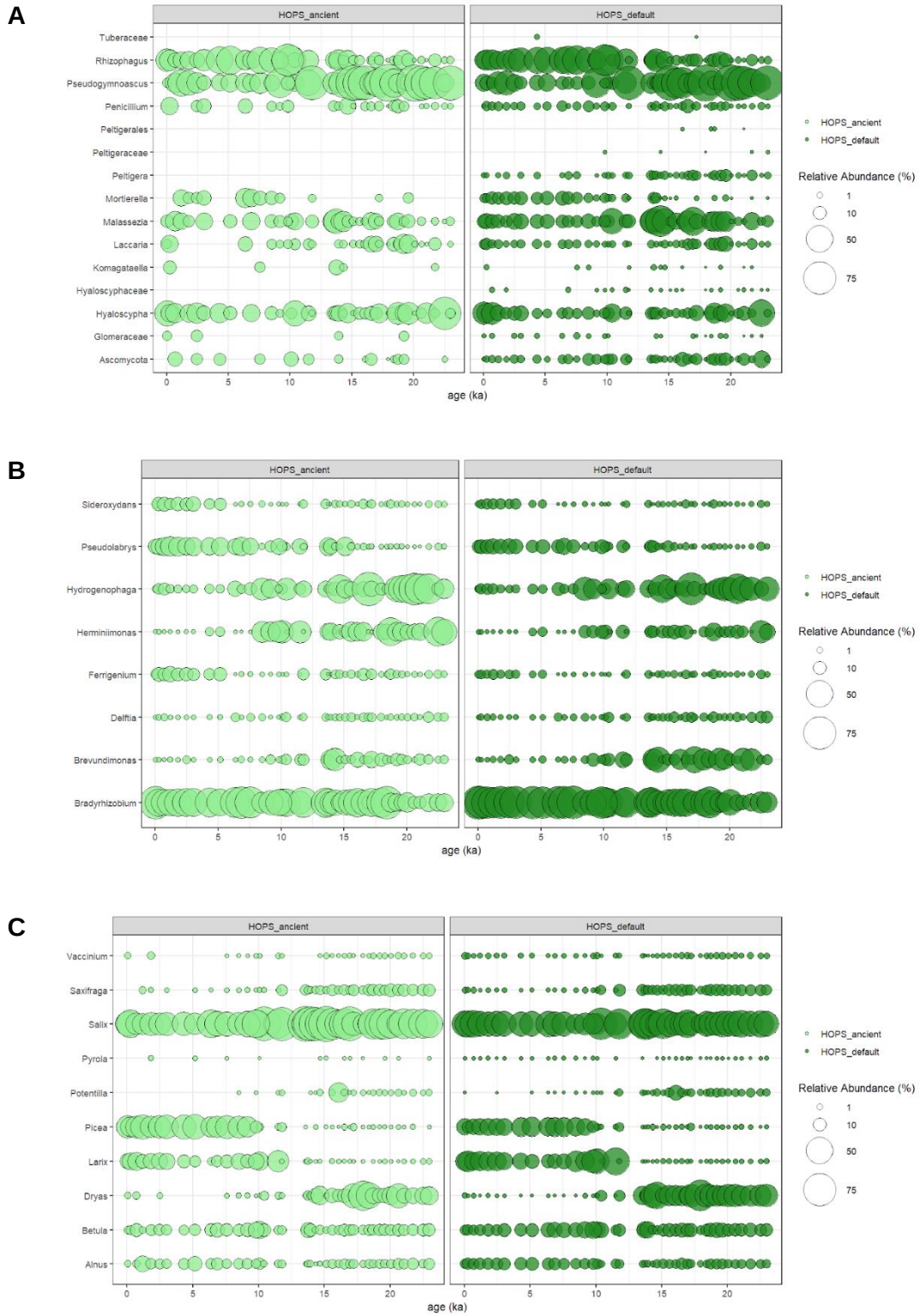


Fig. 1: Comparison between HOPS ancient and HOPS default reads. A: Fungi, B: Bacteria, C: Vegetation.

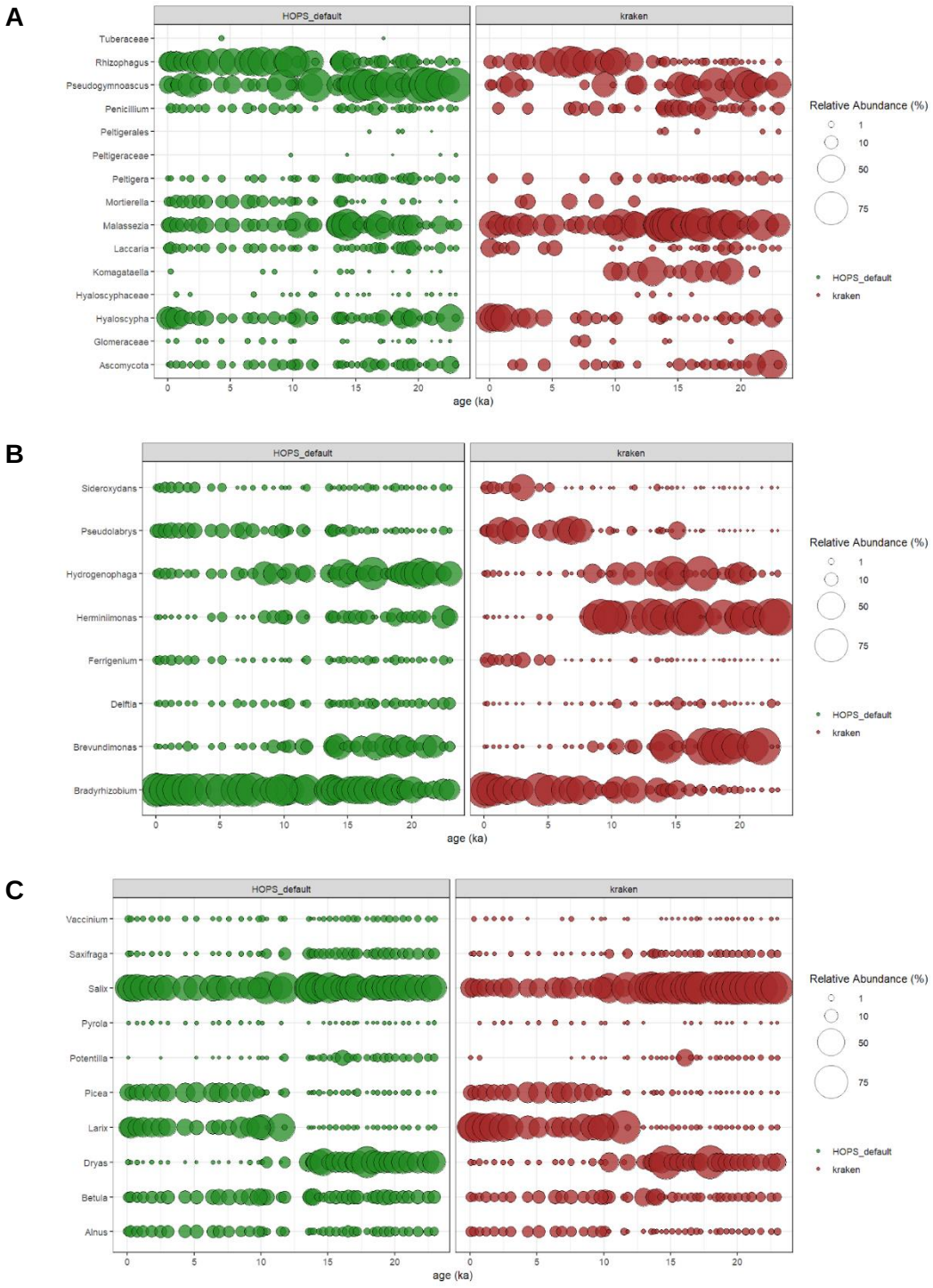


Fig. 2: Comparison between the HOPS (default) and the kraken pipeline. A: Fungi, B: Bacteria, C: Vegetation.

We assessed the statistical comparability of the datasets using procrustes and tested the significance using protest in R Studio, package vegan (Oksanen et al., 2017). The results for all taxa scaled on 100 percent with the procrustes call on the PCA scores of the datasets are displayed in Table 1 below:

Table 1: Statistical comparison between the kraken and the HOPS datasets

procrustes call	procrustes sum of squares	root mean square error	correlation in symmetric procrustes rotation	significance
kraken dataset, HOPS ancient dataset	0.7803	0.1363064	0.4687	0.001
kraken dataset, HOPS default dataset	0.6822	0.127449	0.5637	0.001
HOPS ancient dataset, HOPS default dataset	0.2158	0.07167679	0.8856	0.001

The protest analysis reveals the closest similarity between the ancient and the default dataset (Table 1). This is expectable as the ancient dataset forms part of the HOPS default dataset. Also, the procrustes comparison between the kraken pipeline output and the HOPS default output revealed close similarities between the datasets. All dataset comparisons are statistically significant. With this comparison, we conclude that a damage pattern analysis based on the HOPS pipeline is of use for the validation of the results based on the kraken output.

To evaluate the ancient origin of the assessed reads, we analyzed the damage patterns of the subgroups in the data. To do so, we focused on the data from the Holocene samples due to variations in the sediment properties at a depth of 530 cm (Melles et al., 2006), equalling a depth of 380 cm according to the age-depth model (von Hippel et al., 2022, 2023). This change in the sediment properties indicates changed preservation conditions for the DNA. We summed up all reads in the ancient as well as default datasets for each subset based on sample age. We assessed the damage pattern of selected key taxa for each subset. For fungi in the Holocene, for most samples more than 40% of the reads were assigned on species level, which would yield an overall good taxon-based assessment of the reads. The percentage lowered during the Late Glacial with ongoing sample age. This might derive due to the haplotype variation of fungi (Estensmo et al., 2021; Tedersoo et al., 2022), limited databases (Goodwin et al., 2016; Quince et al., 2009) which reduces the possibilities for an adequate reference genome, or due to expected higher damage for older samples.

To additionally prove the ancientness of the reads, we assessed trends of the C/T substitutions from the first position to position 10 for all taxa in the subsets and placed a trendline (Figure 3). We plotted all reads (Fig. 3A) as well as all reads without zero C/T substitutions in the first position (Fig. 3B), as a

C/T substitution at position 1 is an additional indicator for the ancientness of the reads. For damaged DNA reads, an increase of the damage pattern - measured as the increase in C/T transitions via the deamination of cytosine - at the first positions of the DNA reads is stated (Warinner et al., 2017; Key et al., 2017; Briggs et al., 2007; Kirchner, 2011; Huebler et al., 2019). The further away the position is from the ends of the strands, the less the substitutions become. We recovered these trends for all subsets and proved the ancient origin of the reads.

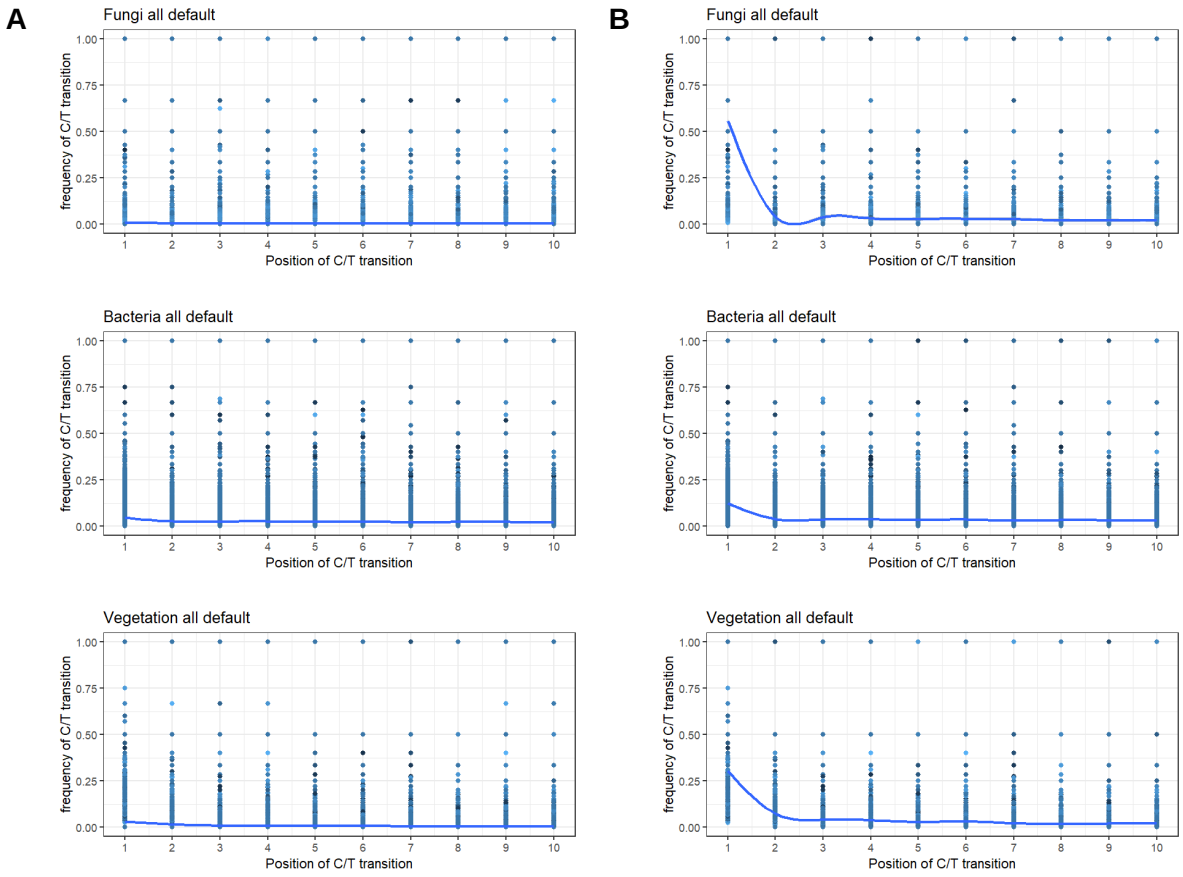


Fig. 3: Frequency of C/T transition at the first 10 positions. A: including all taxa, B: without reads with 0 C/T transitions at position 1.

In general, we show damage patterns for all subsets. Nonetheless, we detected poorer DNA damage patterns mainly for fungi but also bacteria, than for plants. The pattern for fungi is highly biased by reads containing either only C/T substitutions at the respective position or none (Fig. 3A and B), probably derived by the overall very low read counts of fungi.

So far, there is one study assessing DNA damage for fungi derived from Oetzi gut (Oskolkov et al., 2024), assessing the damage of the DNA of *Pseudogymnoascus*. The fungal DNA in Oetzi gut is likely to be better preserved due to permanent freezing than the DNA derived from the lake sediment. Up to

this point, DNA damage was mainly shown for plants (e.g. Crump et al., 2021; Courtin et al., 2022) and mammals (Kistler et al., 2017).

Fungi and plant DNA possess differing splice site combinations (Frey and Pucker, 2020) with some plant species including up to 90 % of repetitive sequences in their genome (Mehrotra and Goyal, 2014). These repetitive sequences are susceptible to DNA damage (Nimeth et al., 2020), which indicates that they could lead to strong damage patterns. In contrast, asco- and basidiomycetes usually have less than 5 % repetitive DNA (Wöstemeyer and Kreibich, 2002). This indicates that an improvement and adaptation of the yet existing damage pattern analysis tools is important when assessing damage of genomes from non-plant kingdoms, such as proposed in Fernandez-Guerra *et al*. (2023), as differently constituted genomes will probably degrade in a different way.

Additionally, the databases for fungi are known to be lacking species (Nilsson et al., 2016). This points out that there are difficulties with finding an appropriate reference genome. A fungal metabarcoding study showed a drastic increase in sequenced fungal genomes over the last 10 years (Seeber et al., 2022), indicating that with deeper sequencing of fungal genomes, the assessment of ancient patterns will be facilitated. The same study also provided information on the mean amplicon length for fungal metabarcoding derived from the same lake and parts of the samples, pointing out shorter amplicons with older sample ages, indicating higher fragmentation.

All in all, this represents only initial results and requires further assessment.

References

Briggs et al., 2007. Patterns of damage in genomic DNA sequences from a Neandertal. *Proceedings of the National Academy of Sciences*, 104(37). 10.1073/pnas.0704665104

Courtin et al., 2022. Pleistocene glacial and interglacial ecosystems inferred from ancient DNA analyses of permafrost sediments from Batagay megaslump, East Siberia. *Environmental DNA*, 4(6). 10.1002/edn3.336

Crump et al., 2021. Ancient plant DNA reveals High Arctic greening during the Last Interglacial. *Proceedings of the National Academy of Sciences*, 118(13). 10.1073/pnas.2019069118

Estensmo et al., 2021. The influence of intraspecific sequence variation during DNA metabarcoding: A case study of eleven fungal species. *Molecular Ecology Resources*, 21(4). 10.1111/1755-0998.13329

Fernandez-Guerra et al., 2023. A 2-million-year-old microbial and viral communities from the Kap København Formation in North Greenland. *bioRxiv*. 10.1101/2023.06.10.544454

Frey and Pucker, 2020. Animal, Fungi, and Plant Genome Sequences Harbor Different Non-Canonical Splice Sites. *Cells*, 9(2). 10.3390/cells9020458

Goodwin et al., 2016. Coming of age: ten years of next-generation sequencing technologies. *Nature*

Reviews Genetics, 17(6). 10.1038/nrg.2016.49
Huebler et al., 2019. HOPS: automated detection and authentication of pathogen DNA in archaeological remains. *Genome Biology*, 20(1). 10.1186/s13059-019-1903-0
Jónsson et al., 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*, 29(13). 10.1093/bioinformatics/btt193
Key et al., 2017. Mining Metagenomic Data Sets for Ancient DNA: Recommended Protocols for Authentication. *Trends in Genetics*, 33(8). 10.1016/j.tig.2017.05.005
Kirchner, 2011. Analysis of High-Throughput Ancient DNA Sequencing Data. In: Ancient DNA: Methods and Protocols. *Humana Press*. 10.1007/978-1-61779-516-9_23
Kistler et al., 2017. A new model for ancient DNA decay based on paleogenomic meta-analysis. *Nucleic Acids Research*, 45(11). 10.1093/nar/gkx361
Mehrotra and Goyal, 2014. Repetitive Sequences in Plant Nuclear DNA: Types, Distribution, Evolution and Function. *Genomics, Proteomics & Bioinformatics*, 12(4). 10.1016/j.gpb.2014.07.003
Melles et al., Eds. 2006. Field evidence for the Late Quaternary climatic and environmental history of the southern Taymyr Peninsula, Central Siberia. In: *Leipziger Geowissenschaften*, 16. Leipzig: Univ.: Inst. f. Geophys. u. Geol.
Nilsson et al., 2016. Top 50 most wanted fungi. *MycoKeys*, 12. 10.3897/mycokeys.12.7553
Nimeth et al., 2020. Alternative splicing and DNA damage response in plants. *Frontiers in Plant Science*, 11. 10.3389/fpls.2020.00091
Oksanen et al., 2017. vegan: Community Ecology Package. R package version 2.4-4.
Oskolkov et al., 2024. Unravelling the ancient fungal DNA from the Iceman's gut. *bioRxiv*. 10.1101/2024.01.24.576930
Quince et al., 2009. Accurate determination of microbial diversity from 454 pyrosequencing data. *Nature Methods*, 6(9). 10.1038/nmeth.1361
Seeber et al., 2022. Evaluation of lake sedimentary ancient DNA metabarcoding to assess fungal biodiversity in Arctic paleoecosystems. *Environmental DNA*, 4(5). 10.1002/edn3.315
Tedersoo et al., 2022. Best practices in metabarcoding of fungi: From experimental design to results. *Molecular Ecology*, 31(10). 10.1111/mec.16460
von Hippel et al., 2022. Long-term fungus–plant covariation from multi-site sedimentary ancient DNA metabarcoding. Quaternary Science Reviews, 295. 10.1016/j.quascirev.2022.107758
von Hippel et al. 2023. Updated age-depth model of sediment core PG1341 from Lake Lama in northern-central Siberia [dataset]. *PANGAEA*. 10.1594/PANGAEA.963262
von Hippel, 2024. Dissertation: Long-term bacteria-fungi-plant associations in permafrost soils inferred from palaeometagenomics
von Hippel et al., unpublished a. Postglacial bioweathering, soil nutrient cycling, and podzolization from palaeometagenomics of plants, fungi, and bacteria
von Hippel et al., unpublished b. Spatio-temporal microbial associations of plants in cold environments.
Warinner et al., 2017. A Robust Framework for Microbial Archaeology. *Annual Review of Genomics and Human Genetics*, 18. 10.1146/annurev-genom-091416-035526
Wöstemeyer and Kreibich, 2002. Repetitive DNA elements in fungi (Mycota): impact on genomic architecture and evolution. *Current Genetics*, 41(4). 10.1007/s00294-002-0306-y