Received: 5 April 2023 | Revised: 21 June 2023 | Accepted: 26 June 2023

DOI: 10.1002/edn3.446

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

WILEY

ORIGINAL ARTICLE

Paleometagenomics reveals environmental microbiome response to vegetation changes in northern Siberia over the millennia

Amedea Perfumo[1,2] | Uğur Çabuk[1,3] | Luise Schulte[1] | Jérémy Courtin[1] | Lars Harms[4] | Kathleen R. Stoof-Leichsenring[1] | Ulrike Herzschuh[1,3,5]

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

[2]Department of Physics, Freie Universität Berlin, Berlin, Germany

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

[4]Computing and Data Centre, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

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

Correspondence
Amedea Perfumo, Polar Terrestrial Environmental Systems, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Potsdam, Germany.
Email: amedea.perfumo@awi.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: 462858357

Abstract

The study of environmental ancient DNA provides us with the unique opportunity to link environmental with ecosystem change over a millennial timescale. Paleorecords such as lake sediments contain genetic pools of past living organisms that are a valuable source of information to reconstruct how ecosystems were and how they changed in response to climate in the past. Here, we report on paleometagenomics of a sedimentary record in northern Siberia covering the past 6700 years. We integrated taxonomic with functional gene analysis, which enabled to shed light not only on community compositions but also on eco-physiological adaptations and ecosystem functioning. We reconstructed the presence of an open boreal forest 6700 years ago that over time was gradually replaced by tundra. This vegetation change had major consequences on the environmental microbiome, primarily enriching bacterial and archaeal ammonia oxidizers (e.g., *Nitrospira*, *Nitrosopumilus*, and *Ca.* Nitrosocosmicus) in the tundra ecosystem. We identified a core microbiome conserved through time and largely consisting of heterotrophic bacteria of the *Bacteroidetes* phylum (e.g., *Mucilaginibacter*) harboring numerous functional genes for degradation of plant-biomass and abiotic and biotic stress resistance. Archaea were also a key functional guild, involved in nitrogen and carbon cycling, not only methanogenesis but possibly also degradation of plant material via enzymes such as cellulases and amylases. Overall, the paleo-perspective offered by our study can have a profound impact on modern climate change biology, by helping to explain and predict the ecological interplay among multiple ecosystem levels based on past experiences.

KEYWORDS
environmental microbiome, functional genes, paleoecosystem, paleometagenomics, sedimentary ancient DNA (sedaDNA)

1 | INTRODUCTION

Ancient DNA from lake sediments (sedaDNA) is a valuable archive of past biodiversity that can help us uncover dynamics of ecosystem change with high accuracy. Used in combination with organism-specific DNA markers and metabarcoding, sedaDNA has proven successful in addressing scientific questions such as climate-driven vegetation changes, within-lake biotic/abiotic processes, trophic interactions, migration routes and colonization of humans and animals, etc. (Capo et al., 2021). A new emerging tool in sedaDNA research is untargeted shotgun metagenomics. Paleometagenomics makes use of next-generation sequencing (NGS) techniques to interrogate the genetic signatures left in nature by past organisms (flora, fauna, and microorganisms) and reconstruct biota dynamics over geological timescale. This approach has already contributed high impact knowledge in unraveling the ecological interplay among kingdoms, mainly between plants and megafauna under the influence of climatic conditions (Kjær et al., 2022; Murchie et al., 2021; Wang et al., 2021; Zimmermann et al., 2023). Studies of this kind are still limited, however, it is becoming clear that integrating data of all organisms is essential to attain a comprehensive reconstruction and meaningful understanding of past ecosystems (Dussex et al., 2021). Furthermore, it allows to combine taxonomic with functional information such as on stress adaptive traits, regulatory and signaling processes, and biogeochemical cycles, thus opening the possibility to study also ecosystem functioning over time and in response to environmental change. Despite of this, identification of functional genes in sedaDNA datasets has been seldom reported. For example, it was possible to demonstrate increase in degradation of complex biopolymers such as pectin and cellulose by a lake microbiome during historical hemp fiber production across the past 2000 years, helping to elucidate the anthropogenic impact on the ecosystem (Iwańska et al., 2022). In another study on marine sedaDNA, microbial catabolic genes for the degradation of fatty acids, proteins, polymers, and methanogenesis were found positively correlating with organic matter during interglacial time, reflecting the impact of climate oscillations (Orsi et al., 2017).

Prokaryotes (bacteria and archaea) have been hardly included in long-term sedaDNA studies, even though (i) they are the most abundant and diverse living organisms on earth, (ii) they support the existence of all higher trophic life forms and the functioning of the ecosystem by driving biogeochemical cycles, and (iii) they are the first and fastest responders to environmental changes. This is largely due to the difficulty to conclusively assign an ecological role to many prokaryotes (e.g., ubiquitous and multi-functional taxa), but also to discriminate whether their DNA signal is a remnant of dead microorganisms deposited over time or it originates from the pool of active microorganisms (e.g., anaerobes) that inhabit the sediments (Capo et al., 2022). Only recently, there has been growing interest in exploring further the use of microbial ecotypes as paleo-indicators. Garner et al. (2020) compared bottom sediment (pre-industrial age) with modern surface water metagenomes to determine the preservation potential of lake microorganisms and attempt disentangle ancient DNA from indigenous sediment background. Others have described changes in microbial communities in response to climatic conditions. For example, shifts in archaeal community were found to relate with the raise of air and lake water temperature and sediment organic matter during the Younger Dryas/Holocene transition (Ahmed et al., 2018). Similarly, a distinctive increase in degrader bacteria and fungi potentially involved in the cycling of soil organic matter was found in forests during warmer interglacial climate in northern Siberia (Courtin et al., 2022).

In this work, we investigate the environmental microbiome inferred from the sedaDNA of a thermokarst lake (named CH12) in northern Siberia over the past 6700 years. Thermokarst lakes are widespread in arctic regions where they form upon permafrost thaw. Despite their shallow depths (<4 m), they are generally stratified, which leads to the presence of oxic and anoxic habitats. The input of organic matter originated from permafrost thaw stimulates microbial degradation, which results in methanogenic conditions, especially in the lake bottom. Thus, thermokarst microbial communities are reported to be similar to communities in other freshwaters but enriched in methane cycling microorganisms. This ecosystem is therefore an important source of greenhouse gases with high relevance for climate feedback loops (Liebner & Welte, 2020). Previous studies on lake CH12 sediments (pollen and plant-targeted metabarcoding) revealed that, likely due to climate cooling, the vegetation in the area experienced a gradual transition, with a spatial decline of open larch forests and a spatial expansion of shrubs, and that ca. 3000 years ago the modern-type shrub tundra took over and became established (Epp et al., 2018; Klemm et al., 2016; Schulte et al., 2021). As a follow-up, here we set out to elucidate what consequences this climate and vegetation change had on the microbial community. We analyzed the sedaDNA shotgun metagenome of four samples from the early stage of lake formation (ca. 6700 years ago) until modern time (ca. 60 years ago) confirming the transition from forest to tundra (ca. 3000 years ago). We integrated taxonomic with functional gene analysis, so we could reconstruct changes not only in the community composition but also with regard to biological activities, physiological adaptations to environmental conditions, cross-kingdom (e.g., plant-microbe) interactions, etc. Our findings demonstrate the important contribution of environmental microbiomes to advance our understanding of past ecosystem dynamics, show the tremendous potential of functional paleometagenomics and overall strengthen the use of an integrative approach in paleoecology research.

2 | MATERIALS AND METHODS

2.1 | Samples and metagenomic data

Sedimentary ancient DNA (sedaDNA) analyzed in this study was retrieved from lake CH12 (mean radius 100 m, maximum depth 14.3 m) in the Khatanga region, north-central Siberia (72.399° N,

102.289° E) (Figure S1). The lake catchment vegetation consists of tundra dominated by dwarf shrubs (Ericaceae), *Salix* spp. especially along the shoreline, grass (Cyperaceae and Poaceae), and herbs (*Dryas* spp.) (Klemm et al., 2016). The sediment core used for DNA extraction, processing under stringent contamination control in a dedicated ultraclean paleogenetic laboratory, library preparation, and sequencing has been described in detail previously (Schulte et al., 2021). Briefly, DNA was extracted with DNeasy PowerMax Soil Kit (Qiagen, Germany), libraries were prepared following an optimized protocol for ancient DNA (Gansauge et al., 2017) and sequencing was performed with Illumina HiSeq 2500, 2×125 base pairs. In this study, we used sequencing reads generated by Schulte et al. (2021) to analyze prokaryotic (bacteria and archaea) community composition and functionality along with Viridiplantae community in four samples dating ~6700 calibrated years before present (cal-BP), ~5400 cal-BP, ~1900 cal-BP, and ~60 cal-BP (BP) (chronology has been described in Klemm et al., 2016). An overview of the dataset is provided in Table S1.

2.2 | Taxonomic and functional annotation

Raw reads were quality checked using FastQC v0.11.9, duplicate reads removed with clumpify in BBtools v38.87 and finally paired and merged with Fastp v0.20.1. Metagenome taxonomic classification was done using Kraken 2 (Wood et al., 2019) with the NCBI nt database (downloaded October 2021) and 0.8 confidence threshold on pre-processed reads (Table S2). Prior to analysis, the dataset for Prokaryotes was checked for contaminants at genus rank and taxa were removed if read counts of the library blanks were equal to the read counts of the samples (Table S3). Viridiplantae dataset was also checked for the presence of contaminants and non-indigenous taxa based on the GBIF Global Biodiversity Information Facility database (Table S4).

For metagenome functional annotation, pre-processing of reads, assembly into contigs, prediction of protein coding sequences, and phylogenetic annotation were performed as it follows. Pre-processed reads were error-corrected using tadpole in BBtools v38.87 and de novo assembled into contigs (>300 bp) using MEGAHIT v1.2.9 (Li et al., 2015) with preset meta-large and default parameter. Contigs quality was assessed with QUAST quality assessment tool for genome assemblies (Gurevich et al., 2013) (Table S5). Protein-coding sequences contained in the assembled contigs were predicted with Prodigal v2.6.3 (Hyatt et al., 2010), clustered using MMseq2 v13.45111 (Steinegger & Söding, 2018) with parameter settings –cov-mod 0 and –clust-mod 0.8 and functionally annotated through eggNOG-mapper v2.1.2 (Cantalapiedra et al., 2021) with Diamond search mode. When not stated otherwise, bioinformatic programs were used with default settings. Functional annotation sources included KEGG (pathways, modules, and orthologs) (Kanehisa & Goto, 2000), Gene Ontology (GO) (Ashburner et al., 2000), EC numbers (enzymes) (Bairoch, 2000), CAZy (carbohydrate-active enzymes) (Drula et al., 2022), and PFAM (protein domains) (Mistry et al., 2021).

MapDamage2 (Jónsson et al., 2013) was used to assess DNA fragment length distribution and base substitution rate as indicative of ancient DNA damage on representative keystone microbial taxa. MapDamage2 analysis on plant taxa was already performed as reported in Schulte et al., 2021.

2.3 | Data analyses

Statistical analyses were performed in R v4.0.5 (R Core Team, 2021) and graphical representations of the results were made with ggplot2 package for R (Wickham et al., 2016). Compositional changes based on taxonomically assigned reads were visualized using relative abundance data. Core bacteria and archaea, defined as those taxa present in all samples with relative abundance >0.1%, were determined at genus level using the microbiome package for R (Lahti et al., 2017) and graphically represented as circular packing chart. To interpret data at ecosystem level, a principal component analysis (PCA) was performed on representative bacterial, archaeal, and plant taxa using the rda function in the vegan package for R (Oksanen et al., 2019). Principal component 1 and 2 (PC1 and PC2) axes scores were visualized in a biplot.

For the analysis of functional genes, we referred to protein coding sequences that were identified mainly in the PFAM database, since it returned the highest number of hits. PFAM results were cross-checked with EC and KEGG databases. Count normalization for target proteins was obtained by the number of predicted protein counts from PFAM divided by the number of predicted protein counts assigned to housekeeping single-copy gene *recA* (Acinas et al., 2021) and visualized as a heatmap. Taxonomic information associated with protein coding sequences were used to infer on the composition of the functional community and its changes over time and graphically represented as a bubble chart.

3 | RESULTS

3.1 | Dataset overview

Reads assigned to microorganisms (bacteria and archaea) dominated lake CH12 metagenomes. Overall, 66 taxa out of 1034 (representing 0.94% of the total reads assigned) were identified as contaminants and removed. They mainly included taxa previously identified in negative controls from multiple studies such as *Acidovorax*, *Bacillus*, *Sphingomonas*, *Cupriavidus*, and *Comamonas* (Eisenhofer et al., 2019) (Table S1). DNA sequences assigned to keystone taxa (*Mucilaginibacter* and *Methylotenera*) were assessed for ancient DNA damage but showed negligible C to T substitution rate (Figure S2a,b). In comparison, previous analysis performed on

plant-assigned reads from the same samples showed clear patterns of base substitution (Schulte et al., 2021).

3.2 | Lake CH12 sedaDNA reflects a strong input of terrestrial biomass

Our analyses focused on the environmental microbiome and vegetation. A first assessment of community composition was made at order level. Within Viridiplantae, terrestrial plants were by far more abundant than aquatic plants (<30% of the total assigned reads) (Figure S3a), suggesting a strong input of terrestrial biomass into the lake from the catchment area. Aquatic taxa were mostly detected in the records between 6700 and 5400 cal-BP and included pondweed plants of the order *Alismatales* (e.g., genus *Potamogeton*) and green algae of the orders *Chlorodendrales*, *Charales*, and *Chlorellales* (Figure S2a). Similarly, prokaryotes showed a predominance of bacteria that are typically associated with soil and plant rhizosphere (Figure S3b). Members of the orders *Sphingobacteriales*, *Pseudomonadales*, *Rhizobiales*, and *Burkholderiales* constituted the large majority (>90% relative abundance) of the microbiome in all samples, with the only exception of the 5400 cal-BP sample, in which *Nitrosomonadales* were dominant (66% relative abundance). Bulk soil bacteria included mainly Actinobacteria (orders *Corynebacteriales*, *Streptomycetales*, *Micrococcales*, and *Micromonosporales*) and Firmicutes (*Bacillales*), and contributed 3%–6% of the entire community (Figure S3b). Aquatic microorganisms represented 0.6%–5% of the total and were mainly cyanobacteria of the orders *Oscillatoriales* and *Nostocales*, *Parachlamydiales* and *Pirellulales* (PVC group), and *Caulobacteriales* (Figure S3b).

Based on these preliminary observations, we concluded that, being strongly affected by terrestrial material input, CH12 sedaDNA was particularly suited to investigate long-term vegetation-microbiome dynamics in the high Arctic over the past 6700 years.

3.3 | Compositional changes of lake catchment's vegetation

First, we set out to determine how the lake's catchment vegetation changed since its early stage of formation upon permafrost thawing 6700 years BP until nearly present time (60 years BP). The most abundant plant families indicated that 6700 years ago, lake's surroundings were covered by *Salicaceae*, *Pinaceae*, and *Equisetaceae* at comparable abundance, together with *Betulaceae* and *Rosaceae* and various forbs and grass to a lower extent (Figure 1; Figure S4). Later in time, *Pinaceae* (*Larix*) and *Betulaceae* trees (*Alnus*) began to decrease, while *Salicaceae* shrubs (*Salix*) increased steadily, becoming the predominant (ca. 76% relative abundance) plant taxon in most recent times. A similar trend occurred for *Betula* shrubs (Figure 1; Figure S4). Thus, a tundra, more resembling modern-time vegetation, consisting primarily of *Salix* together with other shrubs (*Rosaceae*, *Ericaceae*), grass (*Cyperaceae*), and forbs (*Asteraceae*, *Brassicacea*, *Fabaceae*) established ca. 2000–3000 years ago (Figure 1; Figure S4).

3.4 | Taxonomic characterization of the environmental microbiome

Once reconstructed the composition of the plant community over time, we examined the microbial community (bacteria and archaea) with the aim to identify key taxa involved in the ecosystem functioning (e.g., biogeochemical cycles) and interactions with plants. We identified a core microbiome consisting of 24 genera present in all samples with at least 0.1% relative abundance (Figure 2a). Altogether, they accounted for the vast majority of the entire bacterial (93%–96%) and archaeal (89%–96%) communities. The bacterial community dating back to 6700 years BP (corresponding to the earliest stage of the lake formation) was composed to a major extent

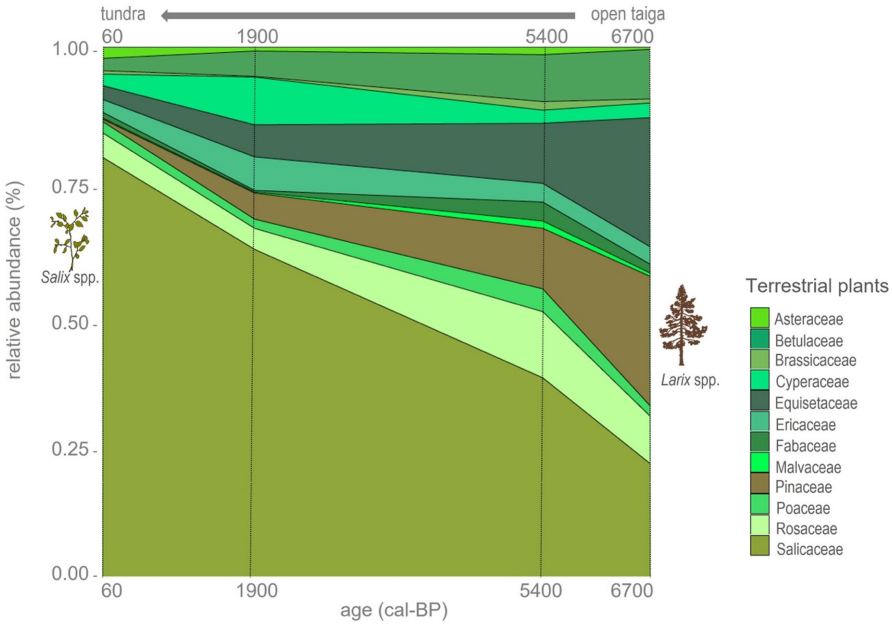

FIGURE 1 Compositional changes of lake catchment's vegetation since its formation until modern times. Stacked area chart visualizing plant community composition based on the most abundant families (top 12, with total read counts >100, representing 91%–97% of the overall community).

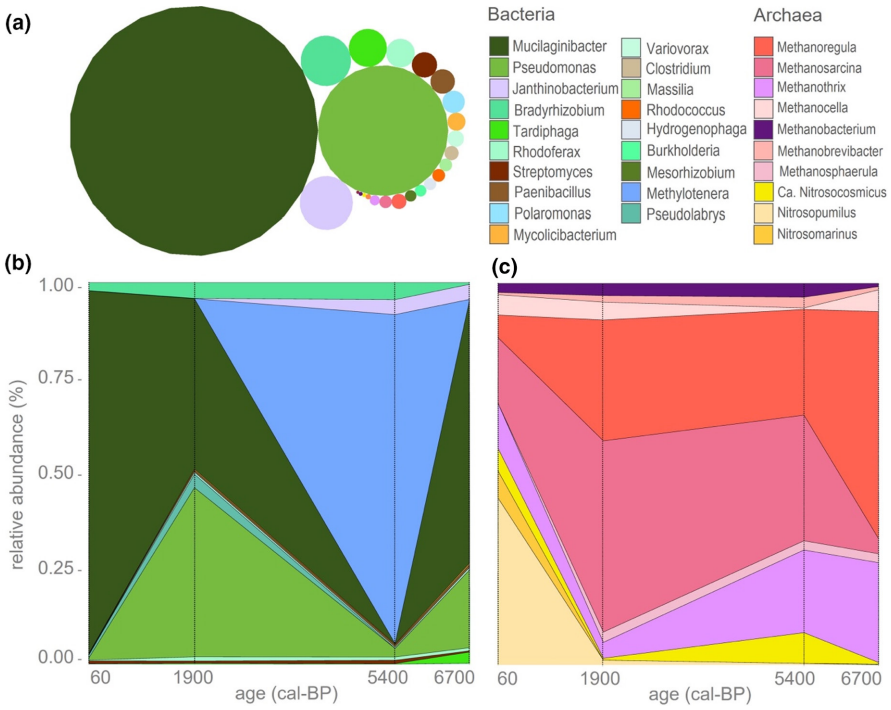

FIGURE 2 Characterization of the paleomicrobiome, core taxa, and compositional changes over time. Visualization of the core microbiome consisting of bacterial and archaeal genera present in all samples (relative abundance >0.1%). Bubble size is proportional to the total read counts per taxon (e.g., *Mucilaginibacter* $n = 140{,}390$, 69% of core taxa; *Pseudomonas* $n = 38{,}471$, 19% of core taxa; all other taxa ≤3%) (a). Stacked area charts showing the compositional changes (relative abundance %) of the main bacterial (b) and archaeal (c) taxa during the past 6700 years.

by *Mucilaginibacter* (closest to the wetland soil species *M. xinganensis*; 66% relative abundance) together with *Pseudomonas* (closest to the fluorescens group; 19% relative abundance) and other minor taxa such as *Janthinobacterium* (4%) and *Tardiphaga* (3% relative abundance). The community between 1900 and 60 years BP was relatively similar, showing a trend toward the establishing of *Mucilaginibacter* as largely dominant bacterium (91% relative abundance) in modern times. *Bradyrhizobium* also was detected from 1900 years BP onwards (4%–2% relative abundance). However, a major community change was recorded in the 5400 cal-BP sample that was characterized by the highest occurrence (81% relative abundance) of the *Nitrosomonadales* genus *Methylotenera* (closest to the freshwater species *M. versatilis*) (Figure 2b). Other core bacteria were *Rhodoferax*, *Polaromonas*, and *Variovorax* (*Burkholderiales*), and *Streptomyces*, *Paenibacillus*, and *Mycolicibacterium* (Terrabacteria) (Figure 2a).

The archaeal community between 6700 and 1900 cal-BP consisted mainly of methanogenic archaea of the genera *Methanoregula* (up to 55% relative abundance), *Methanosarcina* (up to 48% relative abundance), and *Methanothrix* (up to 24% relative abundance). Instead, the youngest sample 60 cal-BP was characterized by a substantial increase in ammonia oxidizing archaea (AOA), especially *Nitrosopumilus*, *Candidatus* Nitrososcosmicus, and *Candidatus* Nitrosomarinus accounting for 51% of the entire archaeal community (Figure 2c).

3.5 | Vegetation-microbiome changes and functional implications

Next, we set out to elucidate how key taxa from different kingdoms, that is, plants, bacteria, and archaea, contributed to shape the ecosystem over time. Principal component analysis (PCA) showed a clear separation of the 6700 cal-BP record from the 5400 cal-BP and the two youngest samples (1900 and 60 cal-BP), most similar with each other, mapping in the same quadrant; the first two principal components (PC1 and PC2) explained 94% of the variation among samples (Figure 3). The left side of the PCA, with the transition from 6700 cal-BP to 1900 cal-BP and 60 cal-BP samples, was representative of changes in the terrestrial ecosystem, overall dominated by *Mucilaginibacter* (phylum *Bacteroidetes*). The right side of the PCA, instead, indicated changes in the lake ecosystem that occurred 5400 years BP driven by freshwater *Methylotenera* (order *Nitrosomonadales*), possibly following a cyanobacterial bloom by *Nostocales* (Figure 3). The sample at 6700 years BP was characterized by the woody plant-associated bacterium *Tardiphaga* (family *Bradyrhizobiaceae*) clustering with *Larix* and *Alnus* trees and the archaea *Methanoregula* and *Methanothrix*. The 1900 cal-BP and 60 cal-BP samples were instead mostly identified by clusters of tree-shrubs and ammonia oxidizers, for example *Betula*, *Methanosarcina* and Ca. Nitrososcosmicus (AOA), *Salix* and *Nitrospira* (AOB). *Nitrosopumilus*, another AOA, was the strongest determinant of the environmental microbiome in the 1900-60-year-time. The denitrifier *Bradyrhizobium* (family *Bradyrhizobiaceae*) was also a distinctive taxon of this cluster (Figure 3).

As a first step to understand the ecological success and dominance of *Mucilaginibacter* over the millennia, we analyzed functional genes assigned to the closest taxonomic level (provided with the eggNOG annotation), that is, phylum *Bacteroidetes*, supported by the fact that the reads assigned to *Mucilaginibacter* accounted for 95% of the total reads of the phylum *Bacteroidetes*. We identified a large range of stress-resistance genes, including resistance to cold and heat, osmotic pressure, heavy metals, and aromatic hydrocarbons,

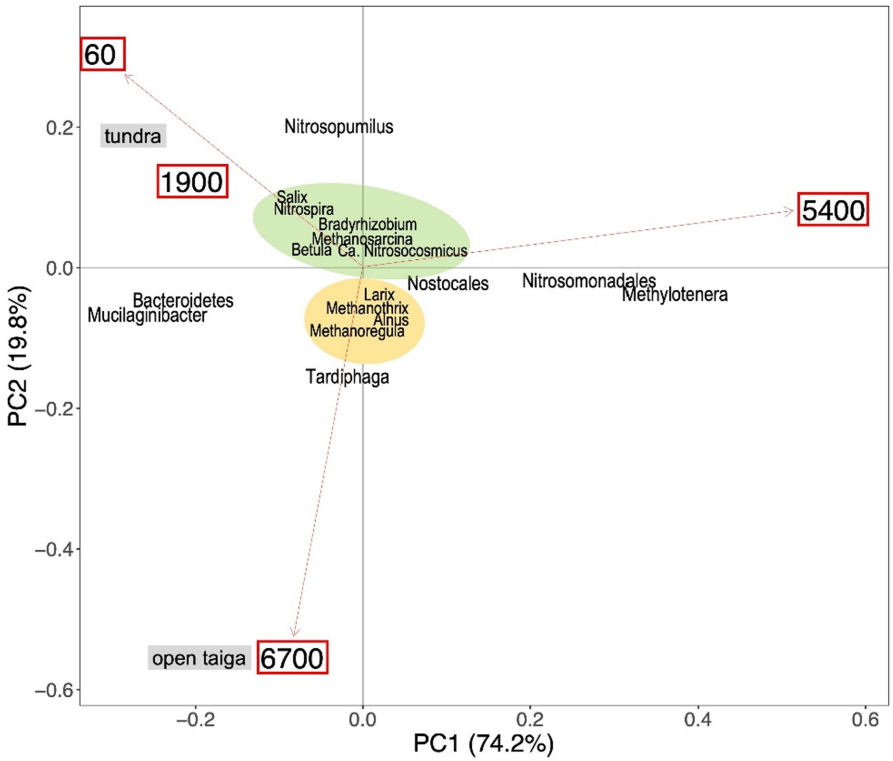

FIGURE 3 Principal Component Analysis (PCA) biplot of representative taxa of bacteria, archaea, and plants. Selected taxa included eight bacteria, five archaea, and four plants. The four samples (6700, 5400, 1900, and 60 cal-BP) were indicated by red boxes. The explained variances of the two principal components (PC1 and PC2) are shown in brackets.

beside universal stress proteins; many genes pointing to a complex extracellular capsule with carotenoid pigments and proteins involved in cell adhesion and biofilm formation; numerous genes involved in antibiotic synthesis (e.g., lactamase, lantibiotics, and penicillins) and multidrug resistance. Regarding metabolism, numerous genes coded enzymes responsible for the degradation of complex biopolymers, for example, a large variety of cellulases, pectinesterases, amylases, xylanases, xylosidases, arabinofuranosidases, chitinases, and lysozymes. Among the also numerous peptidase genes, we identified extracellular triacylglycerol lipases, subtilisin-, and trypsin-like proteases (Table 1 and Table S6).

On the other hand, to *Methylotenera* (contributing 96% relative abundance of the entire *Nitrosomonadales*) were associated several genes coding enzymes involved in single carbon compound (C1) metabolism, including methanol dehydrogenase, formate dehydrogenase, ethanolamine utilization protein, and formaldehyde-activating enzyme (Table S7), suggesting that methylotrophy might have been particularly active 5400 years BP.

Finally, we found a large spectrum of functional genes associated with archaea (Tables 2 and S8). In general, stress-resistance genes were abundant, including universal stress proteins (Usp), heat shock proteins (Hsp) and molecular chaperones, and mannosyl-3-phosphoglycerate synthase for the production of the extremophilic osmoprotectant mannosylglycerate. In addition, we detected various antibiotic biosynthetic pathway genes (e.g., beta-lactamases and penicillins) and antibiotic resistance proteins. With regard to carbon cycling, beside methyl-coenzyme M reductase (methanogenesis) we identified genes for the utilization of methylamines and methanol as a likely substrate for methanogenesis. Enzymes for the degradation of plant polymers, for example, cellulase, amylase, xylosidase, and pectinesterase, were also found in archaea. With regard to nitrogen cycling, we identified genes for nitrogenases in methanogenic archaea and few reads of ammonia monooxygenases in Thaumarchaeota. Furthermore, genes for alkaline phosphatase and dissimilatory sulfite reductase for cycling of phosphorous and sulfur were also detected.

3.6 | Temporal changes of microbial functional genes

Finally, we made a first attempt to identify microbial functionality at community level and reconstruct how it changed over time, focusing on carbon and nitrogen biogeochemical cycles. We detected a variety of genes coding amylase, cellulase, pectate lyase, pectinesterase, cutinase, etc., for the degradation of cellulose, hemicellulose, xylan, pectin, cutin, and other plant biomass (Figure 4a, Table S9). Overall, this gene pool was rather conserved through time, and so it was the associated microbial community, with *Bacteroidetes* as a dominant phylum among lignocellulolytic degraders together with Actinobacteria and Proteobacteria (Figure 4b, Table S10). We detected also high abundance of lysozyme and glycoside hydrolase 108 degrading cell wall peptidoglycan, alginate lyase degrading alginic acid from brown algae and chitinase degrading chitin from fungi and microorganisms (Figure 4a, Table S9), and similarly *Bacteroidetes* were among the main degrader taxa. However, we observed some compositional changes, for example, higher abundance of *Oceanospirillales*' lysozyme and *Sphingomonadales*' glycoside hydrolase 108 in the

TABLE 1 Selected[a] functional genes assigned to *Bacteroidetes* in the sedaDNA metagenomes of lake CH12 relevant to environment adaptation (abiotic/biotic stress resistance) and organic matter degradation.

Biological processes	Functional genes	PFAM	KEGG
Stress resistance	Cold shock protein	PF00313	K03704
	Heat shock protein 20	PF00011	K13993
	Heat shock protein 70	PF00012	K04043
	Lipoyl-dependent peroxiredoxin	PF02566	K04063
	Lycopene beta-cyclase	PF05834	K06443
Antibiotic synthesis and resistance	Beta-lactamase class A	PF13354	K17836
	Lantibiotic biosynthesis protein	PF14028	K20483
	Lantibiotic biosynthesis protein	PF05147	K20484
	Penicillin G amidase	PF01804	K01434
	Quaternary ammonium compound-resistance protein	PF00893	K11741
Biodegradation	Alpha-amylase	PF00128	K01176
	Pectinesterase	PF09492	K01051
	Endo-1,4-beta-xylanase	PF00331	K01181
	1,4-beta-xylosidase	PF04616	K01198
	Alginate lyase	PF05426	K20525
	Chitinase	PF00704	K01183
	Lysozyme	PF01183	K07273

[a]Full list in Table S6.

TABLE 2 Selected[a] functional genes assigned to archaea in the sedaDNA metagenomes of lake CH12 relevant to environment adaptation (abiotic/biotic stress resistance), organic matter degradation, and nitrogen cycle.

Biological processes	Functional genes	PFAM	KEGG
Stress resistance	Universal stress protein	PF00582	—
	Heat shock protein 20	PF00011	K13993
	Heat shock protein 70	PF00012	K04043
	Archaeal chaperone	PF00118	K22447
	Mannosyl-3-phosphoglycerate synthase	PF01025	K03687
Antibiotic synthesis	Beta-lactamase	PF00753	—
	Penicillin G amidase	PF01804	K01434
Methanogenesis	Methyl-coenzyme M reductase	PF02249	K00399
	Methylamine-corrinoid protein Co-methyltransferase	PF05369	K16176
	Trimethylamine-corrinoid protein Co-methyltransferase	PF06253	K14083
Biodegradation	Cellulase	—	K01179
	Cellobiose phosphorylase	PF17167	K00702
	Alpha-amylase	PF03065	K07405
	Alpha-xylosidase	PF16338	K01187
	Pectinesterase	PF13229	K01051
	Chitinase	PF07705	K01183
Nitrogen fixation	Nitrogenase	PF00142	PF00142

[a]Full list in Table S8.

60-year-old record (Figure 4b, Table S10). Genes involved in C1 compound metabolism, utilization of methanol, and detoxification of formaldehyde (e.g., Fae and GFA) were present throughout mainly in methanogenic archaea, but also in *Nitrosomonadales* in the 5400 cal-BP sample, perfectly in line with the high abundance of *Methylotenera* identified earlier by taxonomic annotation (Figure 4b). Methyl-coenzyme M reductase (*mcr*), key-enzyme for methanogenesis, showed a tendency to decrease over time (Figure 4a).

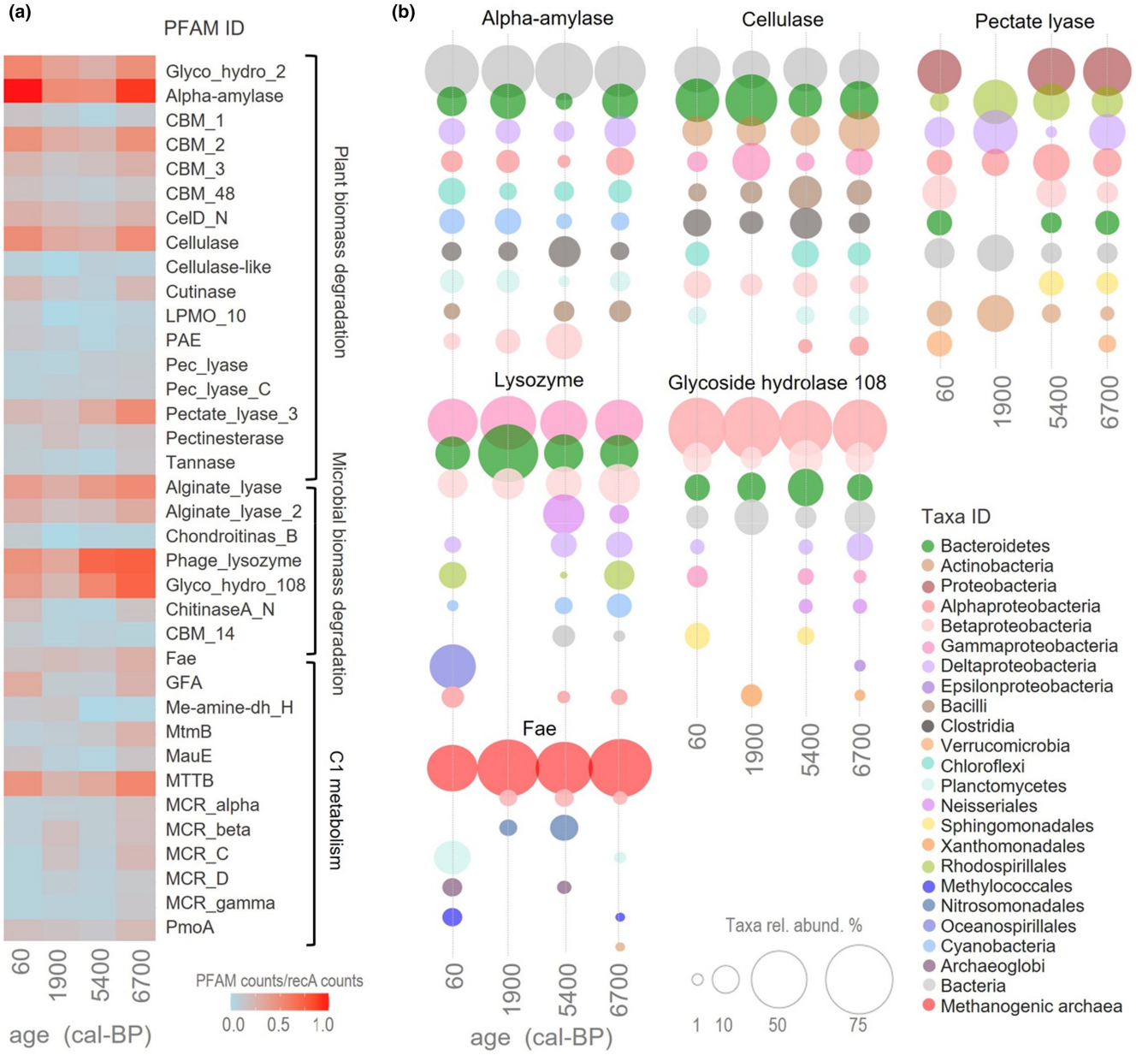

FIGURE 4 Heatmap of key enzymes for plant/microbial biomass degradation and C1 metabolism and corresponding microbial community composition over time (6700 years). A total of 36 enzymes were identified based on PFAM in the sedaDNA metagenomes of lake CH12 and grouped in three main metabolic processes: plant biomass degradation, microbial biomass degradation, and single compounds C1-metabolism. PFAM ID abbreviations: Glyco_hydro_2, Glycosyl hydrolase family 2; Alpha-amyl_C, Alpha-amylase catalytic domain; CBM_1, Fungal cellulose binding domain; CBM_2, Cellulose binding domain; CBM_3, Cellulose binding domain; CBM_48, Isoamylase N-terminal domain; CelD_N, Cellulase N-terminal domain; Cellulase, Glycosyl hydrolase family 5; Cellulase-like, Sugar-binding cellulase; Cutinase, Cutinase; LPMO_10, Lytic polysaccharide mono-oxygenase, cellulose-degrading; PAE, Pectinacetylesterase; Pec_lyase, Pectate lyase; Pec_lyase_C, Pectate lyase; Pectate_lyase_3, Pectate lyase superfamily protein; Pectinesterase, Pectinesterase; Tannase, Tannase, and feruloyl esterase; Alginate_lyase, Alginate lyase; Alginate_lyase_2, Alginate lyase; Chondroitinas_B, Chondroitinase B; Phage_lysozyme, Phage lysozyme; Glyco_hydro_108, Glycosyl hydrolase 108; ChitinaseA_N, Chitinase A N-terminal domain; CBM_14, Chitin-binding peritrophin-A domain; Fae, Formaldehyde activating enzyme; GFA, Glutathione-dependent formaldehyde-activating enzyme; Me-amine-dh_H, Methylamine dehydrogenase heavy chain; MtmB, Monomethylamine methyltransferase; MauE, Methylamine utilization protein; MTTB, Trimethylamine methyltransferase; MCR-alpha, Methyl-coenzyme M reductase alpha subunit, C-terminal domain; MCR_beta, Methyl-coenzyme M reductase alpha subunit, N-terminal domain; MCR_C, Methyl-coenzyme M reductase operon protein C; MCR_D, Methyl-coenzyme M reductase operon protein D; MCR_gamma, Methyl-coenzyme M reductase gamma subunit; PmoA, Methane oxygenase PmoA. PFAM counts were normalized by recA single-copy gene counts relative to the maximum to visualize changes over time (a). Functional microbial community composition for selected PFAM IDs (representative for plant biomass degradation, microbial biomass degradation, and C1-metabolism) over time visualized as a bubble chart with bubble size proportional to the taxon relative abundance. Taxa IDs were derived from eggNOG annotation that provided taxonomic lineage information for each protein coding sequence identified (b).

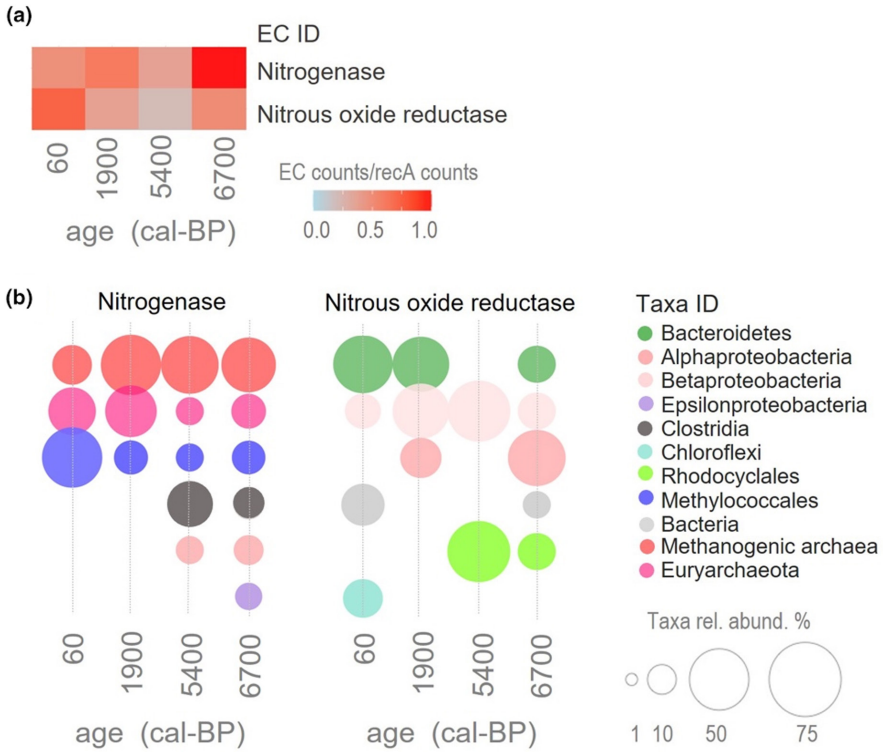

FIGURE 5 Heatmap of key enzymes for nitrogen cycling and corresponding microbial community composition over time (6700 years). Enzymes from nitrogenase (nitrogen fixation) and nitrous oxide reductase (denitrification) were identified based on EC numbers in the sedaDNA metagenomes of lake CH12. EC counts were normalized by recA single-copy gene counts relative to the maximum to visualize changes over time. Taxa IDs were derived from eggNOG annotation that provided taxonomic lineage information for each protein coding sequence identified (a). Functional microbial community composition over time was visualized as a bubble chart with bubble size proportional to the taxon relative abundance (b).

We also detected genes coding key enzymes involved in nitrogen cycling processes, such as nitrogenase (nitrogen fixation), ammonia monooxygenase (nitrification), and nitrous oxide reductase (denitrification). Overall read counts assigned to these enzymes were limited (compared to enzymes involved in carbon cycle) but related well with the ecological role of the taxonomic groups identified. Ammonia monooxygenase counts (assigned to both bacteria and archaea) were particularly low and were not included in quantitative analysis. Nitrogenases (*nif* gene, EC 1.18.6.1) were widespread across archaea and bacteria and mostly abundant in the 6700-year-old sample associated with methanogens. In the younger sample 60 years BP, nitrogenases were only in part related to archaea and approximately half of the nitrogen-fixers was represented by *Methylococcales* (Figure 5a,b). Nitrous oxide reductases (*nosZ* gene, EC 1.7.2.4) showed an increase over time that was particularly marked in the 60-year-old sample (Figure 5a, Table S11). The community of denitrifiers changed composition over time, dominated by Alphaproteobacteria in the sample at 6700 years BP, by Alphaproteobacteria and *Rhodocyclales* in the sample at 5400 years BP and by *Bacteroidetes* together with Betaproteobacteria from 1900 years BP onwards (Figure 5b, Table S12).

4 | DISCUSSION

Paleometagenomics has opened to the possibility to look back in time and reconstruct how environmental microbiomes changed along with the whole ecosystem. Long-term data are needed to reliably predict how microbial functions and feedback mechanisms will respond to climate change, yet only very few such studies exist. In our work, we analyzed microbiome (bacteria and archaea) changes along with vegetation changes that occurred in the high Arctic over the past 6700 years, integrating taxonomic with functional gene analysis. Plants and microorganisms inhabiting the rhizosphere and surrounding soil engage in a multitude of interactions that range from biogeochemical cycles and nutrient acquisition to protection from abiotic and biotic stress, and that have evolved over the course of millennia (Lambers et al., 2009). Here we demonstrated that the characterization of the environmental microbiome greatly supports and integrates paleovegetation, adding a new essential component to the reconstruction of paleoecosystems. Four main points for in-depth discussion have arisen from our results.

First, there was a clear correspondence between plant and microbial signals. The plant community that we reconstructed based on sedaDNA shotgun metagenomics matched perfectly the one described in earlier studies that used established methods such as pollen, DNA metabarcoding, and hybridization capture (Epp et al., 2018; Klemm et al., 2016; Schulte et al., 2021). Genetic signatures of trees, for example, *Larix* and *Alnus*, indicated that forests were present in the area ca. 6700 years BP and they gradually declined over time, while shrubs, for example, *Salix*, steadily expanded up to establish ca. 2000–3000 years BP a tundra-type landscape, as it is found at present. Such a transition from taiga to tundra was previously explained by climate deterioration during the mid-late Holocene (Marcott et al., 2013; Schulte et al., 2021). The microbial community was dominated by terrestrial organisms, especially plant associated (e.g., *Mucilaginibacter*, *Pseudomonas fluorescens*, *Bradyrhizobium*, *Tardiphaga*, etc.). *Mucilaginibacter* of the *Bacteroidetes* phylum was a keystone taxon, present at the highest abundance through all times. Many species of this genus are widespread in arctic terrestrial environments, where, via a large spectrum of secondary

metabolites, bioactive molecules, and degradative enzymes, they interact with plants, degrade plant-biomass, provide nutrients, and protection against abiotic/biotic stress (Figueiredo et al., 2022; López-Mondéjar et al., 2016; Nissinen et al., 2012). These eco-physiological traits were all highly represented in our *Bacteroidetes/Mucilaginibacter* pool of functional genes and conserved through time, thus providing valuable insights into the evolutionary success and ecological role of these bacteria in boreal ecosystems.

Second, changes in the paleovegetation reflected in changes in the microbial community, with clear implications for the ecosystem functioning. The most evident consequence of the replacement of taiga with tundra over time was the increase in nitrogen cycling taxa. Ammonia-oxidizing-archaea (*Ca.* Nitrosocosmicus and *Nitrosopumilus*) and bacteria (*Nitrospira*) and denitrifiers (*Bradyrhizobium*) were key drivers of the ecosystem from 1900 years BP onwards, when modern-like tundra established. These microorganisms are characteristic of arctic soils, particularly in permafrost-affected wetlands (Alves et al., 2013, 2019; Pessi et al., 2022; Wang et al., 2022), where they play a role in the emissions of the greenhouse gas nitrous oxide (Prosser et al., 2020; Siljanen et al., 2019). Vegetation cover has been indicated as an important factor regulating the dynamics of nitrous oxide emissions in arctic terrestrial ecosystems (Voigt et al., 2020). Similar to our observations, studies on modern treeline gradients showed a shift of soil nitrogen cycling along with vegetation, with higher inorganic nitrogen pools and microbial nitrogen cycling communities, both ammonia oxidizers and denitrifiers, in tundra compared to forest. The absence of trees supporting ectomycorrhizal nitrogen mining was identified as the factor promoting archaeal/bacterial-driven inorganic nitrogen cycling (Clemmensen et al., 2021; Gil et al., 2022). Thus, our paleoreconstruction of plant-microbiome dynamics and their functional consequences for nitrogen biogeochemical cycle finds full support in studies on modern environments, demonstrating to be a reliable new analytical approach with a tremendous potential to untangle complex interplays at ecosystem level. In this context, we also observed an interesting opposite trend of two closely related bacteria of the plant-symbiont (e.g., nitrogen-fixing) *Nitrobacteraceae* family, that is, *Tardiphaga* preferentially occurring in the taiga ecosystem (6700 cal-BP sample) and *Bradyrhizobium* in the tundra ecosystem (from 1900 years BP onwards). *Nitrobacteraceae* have been identified as part of the core microbiome of arctic terrestrial ecosystems (Malard et al., 2019). Studies on modern alpine and arctic habitats showed that *Bradyrhizobium* occurrence was higher in soil, while *Tardiphaga* was more abundant with wood (Probst et al., 2018), pointing to an ecological role as early- to mid-stage colonizer of decomposing leaf-litter and roots (Sorensen et al., 2020). Such a niche differentiation identified in modern systems is consistent with our findings from ancient ones and supports the idea that microorganisms, based on their distinctive eco-physiological traits, can be used as paleo-indicators.

Third, archaea were a key functional guild in these extreme environments, involved not only in biogeochemical cycles but also in interactions with plants. Besides ammonia-oxidizing-archaea active in nitrogen cycling, we also found that nitrogen fixation, the process that converts atmospheric nitrogen to ammonia that is then used by plants as nitrogen source, was driven by methanogenic archaea, for example, *Methanotrix* (*nif* gene). Although poorly considered, studies on modern wetlands have shown that methanogenic nitrogen fixation is an important route for fixed nitrogen to enter anoxic soils in permafrost affected areas, also producing a possible impact on methane production (Bae et al., 2018; Fernandez et al., 2020).

Furthermore, various archaeal functional genes pointed out at interactions with plants, supporting nutrient acquisition (e.g., solubilization of phosphorous via alkaline phosphatase and release of sulfur forms that are bioavailable for plant uptake via dissimilatory sulfite reductase) and resistance to abiotic stress. Archaea also harbored genes for extracellular hydrolytic enzymes for the degradation of plant biomass such as cellulase, pectinase, amylase, and chitinase. There is growing evidence that archaea are members of the plant rhizobiome (together with bacteria and fungi), promoting plant growth by directly or indirectly interacting with plant roots (Jung et al., 2020). The fact that archaea are highly tolerant to abiotic stress and thrive in extreme habitats may suggest that they play an important ecological role in improving plant adaptation to environmental extremes as those found in arctic ecosystems.

Fourth, the signal of *Methylotenera versatilis* in the 5400 cal-BP sample is likely aquatic and originated within the lake, rather than being terrestrial, which would also explain why this sample is so distinct compared to the others. *M. versatilis* (*Methylophilaceae* family) is a methylotrophic bacterium found in lake waters and sediments and is a major player of the carbon cycle in freshwater ecosystems, specialized to use single-carbon compounds such as methanol (Kalyuzhnaya et al., 2012). This was supported in our data by the presence of genes assigned to this bacterial group and involved in methanol metabolism such as methanol dehydrogenase, formate dehydrogenase and formaldehyde-activating enzyme (Chistoserdova, 2011). The dynamics of *Mehylophylaceae* are strongly regulated by the availability of methanol, which in aquatic environments is primarily released by phytoplankton, and generally correlate with cyanobacterial blooms (Salcher et al., 2015). In line with this, we also found an increase in *Nostocales* assigned to *Dolichospermum*, a filamentous nitrogen-fixing cyanobacterium that is widespread in northern latitudes and known to be particularly responsive to water temperature and phosphorous supply (Przytulska et al., 2017). Overall, this may be indicative of changes in phytoplankton-bacterioplankton dynamics in response to environmental conditions.

While bacterial DNA, overall, lacked evidence of damage, it still provided information about past ecosystems since community compositional changes were affected by the environmental conditions prevailing at the time of deposition of these taxa. The same has been reported also by other studies reviewed by Capo et al. (2022). However, to harness the full potential of microorganisms as paleoindicators in sedaDNA studies, it is particularly important to better understand differences in survival and preservation potential of different taxonomic groups.

5 | CONCLUSIONS

In conclusion, we demonstrated that the characterization, both taxonomic and functional, of environmental microbiomes inferred from sedaDNA is highly informative in the context of paleoecosystem reconstruction. Although our dataset has limited resolution (four samples) and covered a relatively short time frame (ca. 6700 years) in the Holocene, we were able to identify several changes in the microbial communities that perfectly integrated and improved the reconstruction of paleoclimate and paleovegetation obtained with traditional approaches. We demonstrated that microbial taxa, functions, and eco-physiological traits are valuable indicators of environmental conditions, encouraging their use as proxies in paleoecosystem reconstruction.

We also provided first insights on the analysis of functional genes recovered from sedaDNA. Genes coding enzymes involved in biogeochemical cycles, physiological adaptations to the environment, cross-kingdom (e.g., plant-microbe) interactions all contributed to explain ecological processes and dynamics of the paleoecosystem. At present, this is still limited to microbial functionality, however, we envisage rapid advance in the field of functional paleometagenomics with new bioinformatic tools becoming available and databases containing more eukaryotic sequences. While paleoecosystem reconstruction has been done so far based on taxonomic data from which functionalities have been inferred, the possibility to directly interrogate the pool of functional genes will allow to retrieve information with unprecedented accuracy.

AUTHOR CONTRIBUTIONS

AP designed the study, analyzed the data, and wrote the manuscript. UC, JC, and LH contributed to the bioinformatics pipeline. LS, KRSL, and UH provided sequencing data and preliminary information on paleovegetation. All authors approved the final version of the manuscript.

ACKNOWLEDGMENTS

AP gratefully acknowledges funding from the German Research Foundation DFG (project number 462858357). Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

No conflicts of interest.

DATA AVAILABILITY STATEMENT

The shotgun sequences are available at the European Nucleotide Archive with Accession Numbers SAMEA6430888 to SAMEA6430893.

ORCID

Amedea Perfumo https://orcid.org/0000-0002-7503-4868
Jérémy Courtin https://orcid.org/0000-0002-7712-4277

REFERENCES

Acinas, S. G., Sánchez, P., Salazar, G., Cornejo-Castillo, F. M., Sebastián, M., Logares, R., Royo-Llonch, M., Paoli, L., Sunagawa, S., Hingamp, P., Ogata, H., Lima-Mendez, G., Roux, S., González, J. M., Arrieta, J. M., Alam, I. S., Kamau, A., Bowler, C., Raes, J., … Gasol, J. M. (2021). Deep ocean metagenomes provide insight into the metabolic architecture of bathypelagic microbial communities. *Communications Biology*, *4*(1), 604.

Ahmed, E., Parducci, L., Unneberg, P., Ågren, R., Schenk, F., Rattray, J. E., Lu, H., Muschitiello, F., Pedersen, M. K., Smittenberg, R. H., Yamoah, K. A., Slotte, T., & Wohlfarth, B. (2018). Archaeal community changes in lateglacial lake sediments: Evidence from ancient DNA. *Quaternary Science Reviews*, *181*, 19–29.

Alves, R. J. E., Kerou, M., Zappe, A., Bittner, R., Abby, S. S., Schmidt, H. A., Pfeifer, K., & Schleper, C. (2019). Ammonia oxidation by the arctic terrestrial thaumarchaeote *Candidatus* Nitrososcosmicus arcticus is stimulated by increasing temperatures. *Frontiers in Microbiology*, *10*, 1571.

Alves, R. J. E., Wanek, W., Zappe, A., Richter, A., Svenning, M. M., Schleper, C., & Urich, T. (2013). Nitrification rates in arctic soils are associated with functionally distinct populations of ammonia-oxidizing archaea. *The ISME Journal*, *7*(8), 1620–1631.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene ontology: tool for the unification of biology. *Nature Genetics*, *25*(1), 25–29.

Bae, H. S., Morrison, E., Chanton, J. P., & Ogram, A. (2018). Methanogens are major contributors to nitrogen fixation in soils of the Florida Everglades. *Applied and Environmental Microbiology*, *84*(7), e02222–e02217.

Bairoch, A. (2000). The ENZYME database in 2000. *Nucleic Acids Research*, *28*(1), 304–305.

Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., & Huerta-Cepas, J. (2021). eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Molecular Biology and Evolution*, *38*(12), 5825–5829.

Capo, E., Giguet-Covex, C., Rouillard, A., Nota, K., Heintzman, P. D., Vuillemin, A., Ariztegui, D., Arnaud, F., Belle, S., Bertilsson, S., Bigler, C., Bindler, R., Brown, A. G., Clarke, C. L., Crump, S. E., Debroas, D., Englund, G., Ficetola, G. F., Garner, R. E., … Parducci, L. (2021). Lake sedimentary DNA research on past terrestrial and aquatic biodiversity: Overview and recommendations. *Quaternary*, *4*(1), 6.

Capo, E., Monchamp, M. E., Coolen, M. J., Domaizon, I., Armbrecht, L., & Bertilsson, S. (2022). Environmental paleomicrobiology: using DNA preserved in aquatic sediments to its full potential. *Environmental Microbiology*, *24*, 2201–2209.

Chistoserdova, L. (2011). Modularity of methylotrophy, revisited. *Environmental Microbiology*, *13*(10), 2603–2622.

Clemmensen, K. E., Durling, M. B., Michelsen, A., Hallin, S., Finlay, R. D., & Lindahl, B. D. (2021). A tipping point in carbon storage when forest expands into tundra is related to mycorrhizal recycling of nitrogen. *Ecology Letters*, *24*(6), 1193–1204.

Courtin, J., Perfumo, A., Andreev, A. A., Opel, T., Stoof-Leichsenring, K. R., Edwards, M. E., Murton, J. B., & Herzschuh, U. (2022). Pleistocene glacial and interglacial ecosystems inferred from ancient DNA analyses of permafrost sediments from Batagay megaslump, East Siberia. *Environmental DNA*, *4*(6), 1265–1283.

Drula, E., Garron, M. L., Dogan, S., Lombard, V., Henrissat, B., & Terrapon, N. (2022). The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Research*, *50*(D1), D571–D577.

Dussex, N., Bergfeldt, N., de Anca Prado, V., Dehasque, M., Díez-Del-Molino, D., Ersmark, E., Kanellidou, F., Larsson, P., Lemež, Š., Lord,

E., Mármol-Sánchez, E., Meleg, I. N., Måsviken, J., Naidoo, T., Studerus, J., Vicente, M., von Seth, J., Götherström, A., Dalén, L., & Heintzman, P. D. (2021). Integrating multi-taxon palaeogenomes and sedimentary ancient DNA to study past ecosystem dynamics. *Proceedings of the Royal Society B*, *288*(1957), 20211252.

Eisenhofer, R., Minich, J. J., Marotz, C., Cooper, A., Knight, R., & Weyrich, L. S. (2019). Contamination in low microbial biomass microbiome studies: issues and recommendations. *Trends in Microbiology*, *27*(2), 105–117.

Epp, L. S., Kruse, S., Kath, N. J., Stoof-Leichsenring, K. R., Tiedemann, R., Pestryakova, L. A., & Herzschuh, U. (2018). Temporal and spatial patterns of mitochondrial haplotype and species distributions in Siberian larches inferred from ancient environmental DNA and modeling. *Scientific Reports*, *8*(1), 1–9.

Fernandez, L., Bertilsson, S., & Peura, S. (2020). Non-cyanobacterial diazotrophs dominate nitrogen-fixing communities in permafrost thaw ponds. *Limnology and Oceanography*, *65*, S180–S193.

Figueiredo, G., Gomes, M., Covas, C., Mendo, S., & Caetano, T. (2022). The unexplored wealth of microbial secondary metabolites: the sphingobacteriaceae case study. *Microbial Ecology*, *83*(2), 470–481.

Gansauge, M. T., Gerber, T., Glocke, I., Korlevic, P., Lippik, L., Nagel, S., Riehl, L. M., Schmidt, A., & Meyer, M. (2017). Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. *Nucleic Acids Research*, *45*(10), e79–e79.

Garner, R. E., Gregory-Eaves, I., & Walsh, D. A. (2020). Sediment metagenomes as time capsules of lake microbiomes. *Msphere*, *5*(6), e00512–e00520.

Gil, J., Marushchak, M. E., Rütting, T., Baggs, E. M., Pérez, T., Novakovskiy, A., Trubnikova, T., Kaverin, D., Martikainen, P. J., & Biasi, C. (2022). Sources of nitrous oxide and the fate of mineral nitrogen in subarctic permafrost peat soils. *Biogeosciences*, *19*(10), 2683–2698.

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, *29*(8), 1072–1075.

Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, *11*(1), 1–11.

Iwańska, O., Latoch, P., Suchora, M., Pidek, I. A., Huber, M., Bubak, I., Kopik, N., Kovalenko, M., Gąsiorowski, M., Armache, J. P., & Starosta, A. L. (2022). Lake microbiome and trophy fluctuations of the ancient hemp rettery. *Scientific Reports*, *12*(1), 8846.

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L., & Orlando, L. (2013). mapDamage2. 0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*, *29*(13), 1682–1684.

Jung, J., Kim, J. S., Taffner, J., Berg, G., & Ryu, C. M. (2020). Archaea, tiny helpers of land plants. *Computational and Structural Biotechnology Journal*, *18*, 2494–2500.

Kalyuzhnaya, M. G., Beck, D. A., Vorobev, A., Smalley, N., Kunkel, D. D., Lidstrom, M. E., & Chistoserdova, L. (2012). Novel methylotrophic isolates from lake sediment, description of Methylotenera versatilis sp. nov. and emended description of the genus Methylotenera. *International Journal of Systematic and Evolutionary Microbiology*, *62*(1), 106–111.

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27–30.

Kjær, K. H., Winther Pedersen, M., De Sanctis, B., De Cahsan, B., Korneliussen, T. S., Michelsen, C. S., Sand, K. K., Jelavić, S., Ruter, A. H., Schmidt, A. M. A., Kjeldsen, K. K., Tesakov, A. S., Snowball, I., Gosse, J. C., Alsos, I. G., Wang, Y., Dockter, C., Rasmussen, M., Jørgensen, M. E., … Willerslev, E. (2022). A 2-million-year-old ecosystem in Greenland uncovered by environmental DNA. *Nature*, *612*(7939), 283–291.

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

Lahti, L., Shetty, S., Blake, T., & Salojarvi, J. (2017). *Tools for microbiome analysis in R Version 2.1. 26*.

Lambers, H., Mougel, C., Jaillard, B., & Hinsinger, P. (2009). Plant-microbe-soil interactions in the rhizosphere: an evolutionary perspective. *Plant and Soil*, *321*(1), 83–115.

Li, D., Liu, C. M., Luo, R., Sadakane, K., & Lam, T. W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, *31*(10), 1674–1676.

Liebner, S., & Welte, C. U. (2020). Roles of thermokarst lakes in a warming world. *Trends in Microbiology*, *28*(9), 769–779.

López-Mondéjar, R., Zühlke, D., Becher, D., Riedel, K., & Baldrian, P. (2016). Cellulose and hemicellulose decomposition by forest soil bacteria proceeds by the action of structurally variable enzymatic systems. *Scientific Reports*, *6*(1), 1–12.

Malard, L. A., Anwar, M. Z., Jacobsen, C. S., & Pearce, D. A. (2019). Biogeographical patterns in soil bacterial communities across the Arctic region. *FEMS Microbiology Ecology*, *95*(9), fiz128.

Marcott, S. A., Shakun, J. D., Clark, P. U., & Mix, A. C. (2013). A reconstruction of regional and global temperature for the past 11,300 years. *Science*, *339*(6124), 1198–1201.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., & Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research*, *49*(D1), D412–D419.

Murchie, T. J., Monteath, A. J., Mahony, M. E., Long, G. S., Cocker, S., Sadoway, T., Karpinski, E., Zazula, G., MacPhee, R. D. E., Froese, D., & Poinar, H. N. (2021). Collapse of the mammoth-steppe in central Yukon as revealed by ancient environmental DNA. *Nature Communications*, *12*(1), 1–18.

Nissinen, R. M., Männistö, M. K., & van Elsas, J. D. (2012). Endophytic bacterial communities in three arctic plants from low arctic fell tundra are cold-adapted and host-plant specific. *FEMS Microbiology Ecology*, *82*(2), 510–522.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H., Szoecs, E., & Wagner, H. (2019). *Vegan: Community Ecology Package (Version 2.5-6). The Comprehensive R Archive Network*.

Orsi, W. D., Coolen, M. J. L., Wuchter, C., He, L., More, K. D., Irigoien, X., Chust, G., Johnson, C., Hemingway, J. D., Lee, M., Galy, V., & Giosan, L. (2017). Climate oscillations reflected within the microbiome of Arabian Sea sediments. *Scientific Reports*, *7*(1), 6040.

Pessi, I. S., Viitamäki, S., Virkkala, A. M., Eronen-Rasimus, E., Delmont, T. O., Marushchak, M. E., Luoto, M., & Hultman, J. (2022). In-depth characterization of denitrifier communities across different soil ecosystems in the tundra. *Environmental Microbiome*, *17*(1), 30.

Probst, M., Gómez-Brandón, M., Bardelli, T., Egli, M., Insam, H., & Ascher-Jenull, J. (2018). Bacterial communities of decaying Norway spruce follow distinct slope exposure and time-dependent trajectories. *Environmental Microbiology*, *20*(10), 3657–3670.

Prosser, J. I., Hink, L., Gubry-Rangin, C., & Nicol, G. W. (2020). Nitrous oxide production by ammonia oxidizers: physiological diversity, niche differentiation and potential mitigation strategies. *Global Change Biology*, *26*(1), 103–118.

Przytulska, A., Bartosiewicz, M., & Vincent, W. F. (2017). Increased risk of cyanobacterial blooms in northern high-latitude lakes through climate warming and phosphorus enrichment. *Freshwater Biology*, *62*(12), 1986–1996.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Available at: https://www.R-project.org/

Salcher, M. M., Neuenschwander, S. M., Posch, T., & Pernthaler, J. (2015). The ecology of pelagic freshwater methylotrophs assessed by a high-resolution monitoring and isolation campaign. *The ISME Journal*, *9*(11), 2442–2453.

Schulte, L., Bernhardt, N., Stoof-Leichsenring, K., Zimmermann, H. H., Pestryakova, L. A., Epp, L. S., & Herzschuh, U. (2021). Hybridization capture of larch (Larix Mill.) chloroplast genomes from sedimentary ancient DNA reveals past changes of Siberian forest. *Molecular Ecology Resources*, *21*(3), 801–815.

Siljanen, H. M., Alves, R. J., Ronkainen, J. G., Lamprecht, R. E., Bhattarai, H. R., Bagnoud, A., Marushchak, M. E., Martikainen, P. J., Schleper, C., & Biasi, C. (2019). Archaeal nitrification is a key driver of high nitrous oxide emissions from arctic peatlands. *Soil Biology and Biochemistry*, *137*, 107539.

Sorensen, P. O., Beller, H. R., Bill, M., Bouskill, N. J., Hubbard, S. S., Karaoz, U., Polussa, A., Steltzer, H., Wang, S., Williams, K. H., Wu, Y., & Brodie, E. L. (2020). The snowmelt niche differentiates three microbial life strategies that influence soil nitrogen availability during and after winter. *Frontiers in Microbiology*, *11*, 871.

Steinegger, M., & Söding, J. (2018). Clustering huge protein sequence sets in linear time. *Nature Communications*, *9*(1), 1–8.

Voigt, C., Marushchak, M. E., Abbott, B. W., Biasi, C., Elberling, B., Siciliano, S. D., Sonnentag, O., Stewart, K. J., Yang, Y., & Martikainen, P. J. (2020). Nitrous oxide emissions from permafrost-affected soils. *Nature Reviews Earth & Environment*, *1*(8), 420–434.

Wang, L., Liu, J., Yuan, J., & Wang, N. (2022). The study of soil bacterial diversity and the influence of soil physicochemical factors in meltwater region of Ny-Ålesund, arctic. *Microorganisms*, *10*(10), 1913.

Wang, Y., Pedersen, M. W., Alsos, I. G., De Sanctis, B., Racimo, F., Prohaska, A., Coissac, E., Owens, H. L., Merkel, M. K. F., Fernandez-Guerra, A., Rouillard, A., Lammers, Y., Alberti, A., Denoeud, F., Money, D., Ruter, A. H., McColl, H., Larsen, N. K., Cherezova, A. A., … Willerslev, E. (2021). Late quaternary dynamics of arctic biota from ancient environmental genomics. *Nature*, *600*(7887), 86–92.

Wickham, H., Chang, W., & Wickham, M. H. (2016). Package 'ggplot2'. Create elegant data visualisations using the grammar of graphics. *Version*, *2*(1), 1–189.

Wood, D. E., Lu, J., & Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biology*, *20*(1), 1–13.

Zimmermann, H. H., Stoof-Leichsenring, K. R., Dinkel, V., Harms, L., Schulte, L., Hütt, M. T., Nürnberg, D., Tiedemann, R., & Herzschuh, U. (2023). Marine ecosystem shifts with deglacial sea-ice loss inferred from ancient DNA shotgun sequencing. *Nature Communications*, *14*, 1650.

SUPPORTING INFORMATION

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

How to cite this article: Perfumo, A., Çabuk, U., Schulte, L., Courtin, J., Harms, L., Stoof-Leichsenring, K. R., & Herzschuh, U. (2023). Paleometagenomics reveals environmental microbiome response to vegetation changes in northern Siberia over the millennia. *Environmental DNA*, *5*, 1252–1264. https://doi.org/10.1002/edn3.446