scientific data

Check for updates

OPEN

DATA DESCRIPTOR

# A combined microbial and biogeochemical dataset from high-latitude ecosystems with respect to methane cycle

Maialen Barret[1] ✉, Laure Gandois[1], Frederic Thalasso[2], Karla Martinez Cruz[3,4], Armando Sepulveda Jauregui[3,5], Céline Lavergne[6], Roman Teisserenc[1], Polette Aguilar[6], Oscar Gerardo Nieto[2,7], Claudia Etchebehere[8], Bruna Martins Dellagnezze[8], Patricia Bovio Winkler[8], Gilberto J. Fochesatto[9], Nikita Tananaev[10,11], Mette M. Svenning[12], Christophe Seppey[12,13], Alexander Tveit[12], Rolando Chamy[14], María Soledad Astorga España[3], Andrés Mansilla[3], Anton Van de Putte[15], Maxime Sweetlove[15], Alison E. Murray[16] & Léa Cabrol[17,18] ✉

**High latitudes are experiencing intense ecosystem changes with climate warming. The underlying methane ($CH_4$) cycling dynamics remain unresolved, despite its crucial climatic feedback. Atmospheric $CH_4$ emissions are heterogeneous, resulting from local geochemical drivers, global climatic factors, and microbial production/consumption balance. Holistic studies are mandatory to capture $CH_4$ cycling complexity. Here, we report a large set of integrated microbial and biogeochemical data from 387 samples, using a concerted sampling strategy and experimental protocols. The study followed international standards to ensure inter-comparisons of data amongst three high-latitude regions: Alaska, Siberia, and Patagonia. The dataset encompasses different representative environmental features (e.g. lake, wetland, tundra, forest soil) of these high-latitude sites and their respective heterogeneity (e.g. characteristic microtopographic patterns). The data included physicochemical parameters, greenhouse gas concentrations and emissions, organic matter characterization, trace elements and nutrients, isotopes, microbial quantification and composition. This dataset addresses the need for a robust physicochemical framework to conduct and contextualize future research on the interactions between climate change, biogeochemical cycles and microbial communities at high-latitudes.**

[1]Laboratoire d'Ecologie Fonctionnelle et Environnement, Université de Toulouse, CNRS, Toulouse, France. [2]Biotechnology and Bioengineering Department, Center for Research and Advanced Studies (Cinvestav), Mexico City, Mexico. [3]University of Magallanes, Punta Arenas, Chile. [4]Environmental Physics Group, Limnological Institute, University of Konstanz, Konstanz, Germany. [5]Center for Climate and Resilience Research (CR)2, Santiago, Chile. [6]HUB AMBIENTAL UPLA, Universidad Playa Ancha, Valparaíso, Chile. [7]Unidad Académica de Ecología y Biodiversidad Acuática, Instituto de Ciencias del Mar y Limnología, UNAM, Mexico City, México. [8]Microbial Ecology Laboratory, BioGem Department, Biological Research Institute Clemente Estable, Montevideo, Uruguay. [9]Department of Atmospheric Sciences, University of Alaska Fairbanks, Fairbanks, Alaska, USA. [10]Melnikov Permafrost Institute, Yakutsk, Russia. [11]North-Eastern Federal University, Yakutsk, Russia. [12]Department of Arctic and Marine Biology, UiT The Arctic University of Norway, Tromsoe, Norway. [13]Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany. [14]Pontifical Catholic University of Valparaíso, Valparaiso, Chile. [15]BEDIC, OD Nature, Royal Belgian Institute of Natural Sciences, Brussels, Belgium. [16]Division of Earth and Ecosystem sciences, Desert Research Institute, Reno, NV, USA. [17]Aix-Marseille University, Univ Toulon, CNRS, IRD, M.I.O. UM 110, Mediterranean Institute of Oceanography, Marseille, France. [18]Millenium Institute "Biodiversity of Antartic and Subantarctic Ecosystems" (BASE), Santiago, Chile. ✉e-mail: maialen.barret@toulouse-inp.fr; lea.cabrol@mio.osupytheas.fr

## Background & Summary

Almost half of worldwide $CH_4$ emissions originate from natural ecosystems, among which aquatic ecosystems, i.e. wetlands and other inland waters, are major contributors[1]. At high latitudes, where the highest density of freshwater ecosystems is found[2], extreme variability is occurring as a result of climate change[3]. In the Arctic, air temperatures are expected to rise twice as fast as the global average[4]. Higher air temperature speeds up permafrost thawing, making available a larger amount of sequestered organic matter for microbial degradation and mineralisation. Permafrost thawing can also lead to anoxic conditions in the resulting soil, peat and/or water bodies, potentially enhancing $CH_4$ emissions. Any disturbance of natural $CH_4$ cycle can constitute a strong positive feedback on global climate, considering the strong radiative effect of this greenhouse gas[5,6]. For these reasons, documenting microbially-mediated $CH_4$ emissions in high-latitude ecosystems is essential to determine tipping points on the positive feedback to climate warming. Climatic projections may also impact significantly the $CH_4$ emissions from Sub-Antarctic environments. Ecosystems in the Magellanic ecoregion[7] have been understudied compared to their northern counterparts[8] and $CH_4$ cycling has been rarely investigated[9,10]. However, this region is of major importance since it is an expansive, unique continental area between 45 and 55 °S, it is very sparsely populated and the ecology of the region has been highly conserved (Cape Horn Biosphere Reserve, UNESCO).

Net atmospheric $CH_4$ emissions from terrestrial and aquatic ecosystems reflect the balance between $CH_4$ production, transport, and oxidation in these ecosystems. Methane is primarily produced by methanogenic Archaea through anaerobic decomposition of organic matter. Methanogenesis can occur via the acetoclastic, hydrogenotrophic or methylotrophic pathways, the contribution of each depending on temperature, substrate availability, and microbial interactions[11]. This being the only biological $CH_4$ source has been challenged by recent evidences for $CH_4$ production in oxic conditions[12,13], contributing for about 20% of $CH_4$ emissions from lakes[14,15]. However, the global significance of oxic $CH_4$ production, its underlying metabolic pathways, and the identity of microorganisms actively involved in this process are not fully constrained yet[16]. Methane emissions are strongly mitigated by microbial oxidation. For example, in aquatic ecosystems, 51–100% of the $CH_4$ produced in deep sediments can be oxidized in the water column, before reaching the atmosphere[17,18]. Methane oxidation is carried out in oxic conditions by aerobic methane-oxidizing bacteria (MOB) belonging to Alphaproteobacteria, Gammaproteobacteria, and Verrucomicrobia[19]. On the other hand, anaerobic oxidation of methane (AOM) has been also identified as a major process in aquatic[20,21] and terrestrial[22,23] ecosystems, and attributed to anaerobic methane-oxidizing Archaea (ANME)[24], bacteria from the NC10 phylum[25] or Gammaproteobacteria MOB active in anoxic zones[20,26].

At the landscape scale, these processes occur with different magnitude depending on ecosystem type and associated microtopography. Soils are primarily $CH_4$ sinks due to the drawdown of even very low levels of atmospheric $CH_4$[27,28]. Lakes and wetlands are recognized as major $CH_4$ emission sources, despite high flux variability between and within aquatic ecosystems[1,29]. Among wetlands, organic-rich peatlands are estimated to cover more than $3.7 \times 10^6$ $km^2$ in northern high latitudes[30] and to emit 30 Tg $CH_4$ $yr^{-1}$ [31]. These important $CH_4$ contributors are complex ecosystems with a variety of hydrologic regimes, productivity levels, vegetation covers, and variability in $CH_4$ emissions. Especially, lower emissions are generally observed in fens compared to bogs, in association with different $CH_4$ production pathways[32]. Temperature, water-table level, and permafrost state also influence $CH_4$ emissions[33]. In permafrost landscapes, $CH_4$ emissions from 'wet features', characteristic of degrading permafrost (e.g. ponds, hollows, thaw lakes, internal lawns, and collapse scars), are usually higher than from 'dry features', characteristic of intact permafrost (e.g. pingos, polygonal peatlands, hummocks, palsa, or peat plateau)[34]. Physicochemical characteristics such as pH, nitrogen, and phosphorus availability, and carbon source quality and quantity also drive $CH_4$ emissions[35–37]. In the context of global warming, the expected variations of geochemical and physicochemical factors may affect the microbial $CH_4$-cycle.

The present study focuses on the $CH_4$ cycle in three high-latitude regions (Alaska, Patagonia, Siberia). Three ecosystem types (soils, wetlands, and lakes) have been investigated in each region during summer, with a systematic evaluation of their different habitats. This data set addresses the recent call of the global community of microbial scientists for the integration of microorganisms in mainstream climate change research addressing carbon fluxes[38,39]. Here, a thorough analysis of microbial community diversity and structure, carried out through functional gene quantifications and high throughput 16S rRNA gene amplicon sequencing, has been coupled with the physicochemical characterisation of habitats and measurement of atmospheric $CH_4$ and $CO_2$ fluxes. The physicochemical analysis included quantification of nutrients and trace elements, as well as stable isotopic signature of carbon species ($CH_4$, dissolved organic and inorganic carbon) to track $CH_4$ production and oxidation pathways. This database offers the possibility to expand the geographical scope of microbial ecology, biogeographic, and/or biogeochemical studies (either related to C cycling or other cycles) towards high latitude ecosystems. Moreover, this database is of particular interest for the earth system science community in order to parameterize relevant surface and sub-surface biogeochemical processes that can be further used to refine climate models or global models.

## Methods

**Sites overview and characteristics.** This study focused on three regions located in subantarctic, arctic, and subarctic latitudes. The respective latitudinal and longitudinal ranges covered in this study were: 54.95 to 52.08 °S, and 72.03 to 67.34 °W in Patagonia; 67.44 to 67.54 °N, and 86.59 to 86.71 °E in Siberia; 63.21 to 68.63 °N, and −150.79 to −145.98 °W in Alaska (Figs. 1 and 2). The exact coordinates for each sample were included in the submitted dataset. The field campaigns were conducted in 2016, during the summer for each respective region: January-February in Chilean Patagonia, June-July in Alaska and July-August in Siberia.

For every site included in the present study, a set of nine qualitative environmental and/or ecological site-scale descriptors was selected and adapted from ENVO Environment Ontology[40], which included for

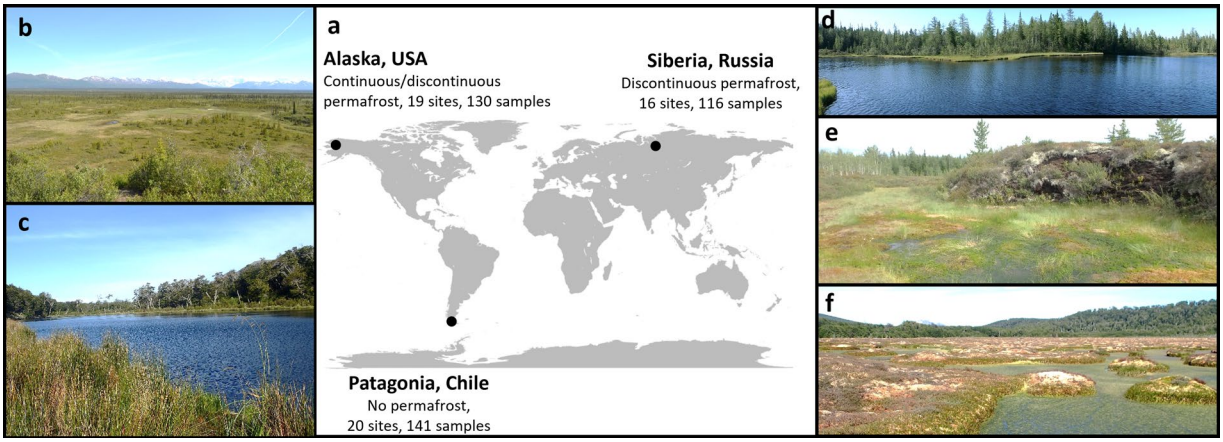
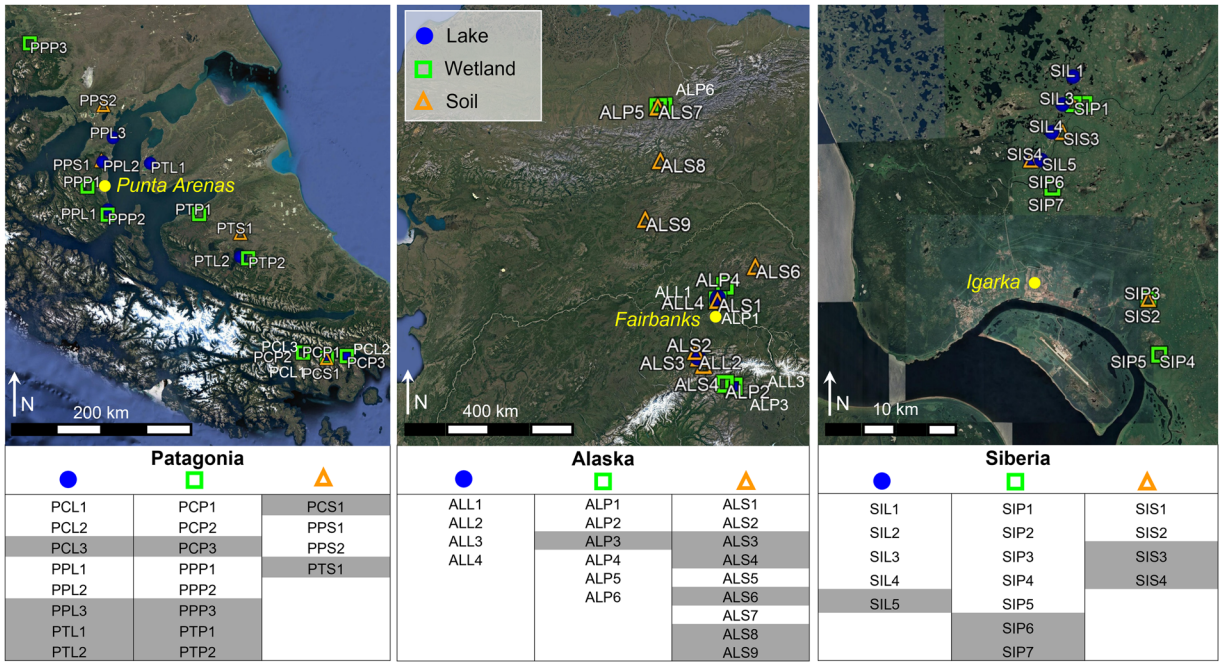



**Fig. 1** Location of the three areas included in this study (panel **a**). The permafrost state and the number of sites and samples per region is indicated for each area. General views of 5 sites are provided as examples (**b**–**f**). Panel B provides a large view of the ecosystem surrounding the wetland ALP2 (Alaska, exact location indicated by the white circle). Lake PCL1 (panel **c**) is representative of the lakes on Navarino island (Chilean Patagonia). The glacial lake SIL2 is shown in panel d. At site SIP5, the hollow at first plan is surrounded by palsa (hummock, second plan), characterized by dark organic matter and lichen vegetation (panel **e**). The PPP3 peatland shown in panel f is dominated by *Sphagnum magellanicum*, like most peatlands in the area.




**Patagonia**

| | | |
|---|---|---|
| PCL1 | PCP1 | PCS1 |
| PCL2 | PCP2 | PPS1 |
| PCL3 | PCP3 | PPS2 |
| PPL1 | PPP1 | PTS1 |
| PPL2 | PPP2 | |
| PPL3 | PPP3 | |
| PTL1 | PTP1 | |
| PTL2 | PTP2 | |

**Alaska**

| | | |
|---|---|---|
| ALL1 | ALP1 | ALS1 |
| ALL2 | ALP2 | ALS2 |
| ALL3 | ALP3 | ALS3 |
| ALL4 | ALP4 | ALS4 |
| | ALP5 | ALS5 |
| | ALP6 | ALS6 |
| | | ALS7 |
| | | ALS8 |
| | | ALS9 |

**Siberia**

| | | |
|---|---|---|
| SIL1 | SIP1 | SIS1 |
| SIL2 | SIP2 | SIS2 |
| SIL3 | SIP3 | SIS3 |
| SIL4 | SIP4 | SIS4 |
| SIL5 | SIP5 | |
| | SIP6 | |
| | SIP7 | |

**Fig. 2** Maps of sampling sites in Patagonia, Alaska and Siberia, indicating the ecosystem type (lake, wetland, soil). The tables show the complete- (in white) and the partial- (in grey) characterization sites. The exact coordinates of each sample are provided in the data record (See data records section).

example permafrost state, biome, environmental feature and vegetation type (Table 1, Fig. 3). Permafrost state was obtained from the NSIDC permafrost map[41]. The biome, large-scale descriptor based on climate and vegetation criteria, was derived from Olson *et al.*[42]. Temperate forest, boreal forest, and tundra biomes were included. The environmental features that were representative for the three regions were considered: lakes, wetlands, broadleaf/coniferous/mixed forest soils, grassland, tundra, and palsa. All the metadata was included in the submitted dataset. Table 2 summarizes the main types of sampled ecosystems and their main characteristics in the three regions, while Supplementary Table S1 provides the details of each sampling site.

In Alaska, the studied area ranged from the Alaska Range and Fairbanks area (interior, continental climate, 63–65°N, discontinuous permafrost) up to Toolik Field Station (North Slope, arctic climate, 66–69°N, continuous permafrost; Fig. 2). The physiochemistry and $CH_4$ emissions of lakes ALL1 (Killarney lake), ALL2 (Otto lake), ALL3 (Nutella lake), and ALL4 (Goldstream lake) were previously characterized[35]. A number of


**Fig. 3** Description of the qualitative environmental/ecological descriptors used to describe every sample, derived from ENVO Environment Ontology[40].

heterogeneous soil and wetland samples were collected around the studied Alaskan lakes and/or from monitored sites, as detailed in Supplementary Table S1. In the Alaska Range and Fairbanks area, soils were mostly covered by mixed or taiga forests, alpine tundra, and bogs or fens wetlands. In the norther Brooks Ranges mountain system, the landscape was piedmont hills with a predominant soil of porous organic peat underlain by silt and glacial till, all in a permafrost state, characterized mainly by *Sphagnum* and *Eriophorum* vegetation, as well as dwarf shrubs.

In Siberia, the studied area was located in the discontinuous permafrost region surrounding Igarka, on the eastern bank of the Yenisei River (Fig. 2). This region was mainly covered by forest, dominated by larch (*Larix Siberica*), birch (*Betula Pendula*), and Siberian pine (*Pinus Siberica*), and palsa landscapes (frozen peat mounts), the latter being dominated by moss, lichens, Labrador tea and dwarf birch. In degraded areas, thermokarst bogs were dominated by *Sphagnum* spp. and *Eriophorum* spp. Land cover was an indicator of permafrost status, since forested areas reflected a deep permafrost table (>2 m) associated with Pleistocene permafrost, while palsa-dominated landscapes were indicative of the presence of near-surface (<1 m) Holocene permafrost. In this area, most of the lakes were of glacial origin and influenced by permafrost degradation[43] that has been observed for the last 30 years, while some were thermokarst lakes (Supplementary Table S1). Two studies that focused on methane cycling in SIL1 to SIL4 were recently published[18,20]. We sampled organic soils on a degradation gradient from dry palsa to thermokarst bogs[44], as detailed in Supplementary Table S1.

Subantarctic sites were located in three areas in the Southern part of Chilean Patagonia: the Magellanic region around Punta Arenas, Tierra del Fuego, and Navarino Island (Fig. 2). Most of the sampled lakes from Magellanic and Tierra del Fuego regions were of glacial origin, while Navarino Island lakes were peatland lakes, surrounded by peatland and broadleaf forests. Peatlands were characterized by a very low diversity of *Sphagnum* species

| Ecological/environmental descriptors | Bioclimatic variables | Physicochemical characteristics |
|---|---|---|
| **Site-scale descriptors:**<br>Region*<br>Biome*<br>Elevation*<br>Environmental feature*<br>Geographical location*<br>Oxygen stratification*<br>Permafrost state*<br>Total depth of water column<br><br>**Point-scale descriptors:**<br>Latitude*<br>Longitude*<br>Microtopography*<br>Vegetation type*<br><br>**Sample-scale descriptors:**<br>Environmental material*<br>Environmental package*<br>Sample depth* | Annual Mean Temperature*<br>Annual Precipitation*<br>Isothermality*<br>Max Temperature of Warmest Month*<br>Mean Diurnal Range*<br>Mean Temperature of Coldest Quarter*<br>Mean Temperature of Driest Quarter*<br>Mean Temperature of Warmest Quarter*<br>Mean Temperature of Wettest Quarter*<br>Min Temperature of Coldest Month *<br>Precipitation of Coldest Quarter*<br>Precipitation of Driest Month*<br>Precipitation of Driest Quarter*<br>Precipitation of Warmest Quarter*<br>Precipitation of Wettest Month*<br>Precipitation of Wettest Quarter*<br>Precipitation Seasonality*<br>Temperature Annual Range*<br>Temperature Seasonality* | Conductivity*<br>$\delta^{13}$C-DOC ●<br>Dissolved organic carbon ● ●<br>Dissolved oxygen ● ○<br>Dry weight*<br>Organic matter ● ○<br>pH*<br>Redox potential ● ○<br>Suspended organic matter ● ○<br>Suspended particulate matter ● ○<br>Temperature*<br>Total organic carbon ● ○<br>Volatile solids percentage ● ○<br>Water content ● ○<br><br>**Anions and cations:** ● ●<br>ammonium, bromide, calcium, chloride, magnesium, nitrate, nitrite, phosphate, potassium, sodium, fluoride, sulphate<br><br>**Trace elements:** ● ●<br>aluminium, antimony, arsenic, barium, cadmium, caesium, chromium, cobalt, copper, iron, lanthanium, lead, manganese, nickel, rubidium, strontium, titanium, uranium, vanadium, zinc<br><br>**Optical properties:** ● ●<br>Absorbance et 254 nm, Specific UV absorbance, Fluorescence index |
| **GHG cycling** | **Microbial variables** | **Miscellaneous** |
| $CH_4$ emission rate ● ●<br>$CO_2$ emission rate ● ●<br>Dissolved methane ●<br>Dissolved inorganic carbon ●<br>$\delta^2$H-$CH_4$ ●<br>$\delta^{13}$C-$CH_4$ ●<br>$\delta^{13}$C-$CO_2$ ● | Archaeal abundance*<br>Bacterial abundance*<br>Methanogen abundance*<br>Methanotroph abundance*<br>Microbial community composition* | Characterization type*<br>Collection date*<br>Ecosystem nomenclature*<br>Official ecosystem name* |

**Table 1.** Overview of the dataset contained in Mimarks sheet. Units are provided for each variable in the database. Data provided for every sample are indicated with the symbol *, while others are available only for sites with complete characterization (full symbols), or for certain environmental features or packages (soil and sediment samples in brown; water samples in blue). Empty symbols represent data available for partial-characterization sites.

| | Patagonia | Alaska | Siberia |
|---|---|---|---|
| Soils | - Broadleaf forest (*Nothofagus*) | - From mixed boreal forest to taiga forest | - Mixed forest (larch, birch, pine) |
| | - Grasslands | - Alpine tundra | - Palsa landscapes (moss, lichens) |
| | - Peatlands (*Sphagnum magellanicum*) | - Boreal tussock tundra | - Thermokarst bogs (*Sphagnum* and *Eriophorum*) |
| | | - Wetlands, including bogs and fens | |
| Lakes | - Peatland lakes | Mixotrophic and oligotrophic lakes formed from either Yedoma- or non-Yedoma permafrost soil | Mostly of glacial origin, influenced by permafrost degradation |
| | - Reservoir | | |
| | - Glacial lakes | | |

**Table 2.** Main types of sampled ecosystems in the three studied regions.

dominated by *S. magellanicum* from hollows up to hummocks. The typical broadleaf forests of the area were dominated by *Nothofagus*. Some grassland soil came from an experimental monitored field site (Supplementary Table S1). Samples collected from Patagonian soils and wetlands have been included in a recent survey of soil geochemical characterization (organic content)[45]. Sediment samples collected in lakes PPL1, PPL2, PCL1, PCL2, PCP2 were also included in a recent study by Lavergne *et al.*[46] which showed that increasing air temperature led to enhanced $CH_4$ production and to an associated metabolic shift in the $CH_4$ production pathway, increasing the relative contribution of hydrogenotrophic methanogenesis compared to acetoclastic methanogenesis, together with consistent microbial community changes.

Surface area for lakes and elevation for all sites were determined using Google Earth Pro. Climate variables (Table 1) for each site were retrieved from WorldClim – Global[47].

**Sampling design.** A specific sampling strategy was defined for each kind of ecosystem, i.e. lakes, soils, and wetlands (Fig. 4), as follows.

In lakes, surface (0–10 cm) sediments and water samples were collected from three replicate points A, B, and C (Fig. 4) corresponding to the deepest zone of the lake, at ~ 2–5 meters of distance from each other. Two sampling depths were considered for the water samples: (i) at the oxycline, and (ii) just above the interface with sediment. Water was sampled using a 2.2 L Van Dorn bottle (Wildco, Mexico). Sediments were sampled using an Ekman dredge.

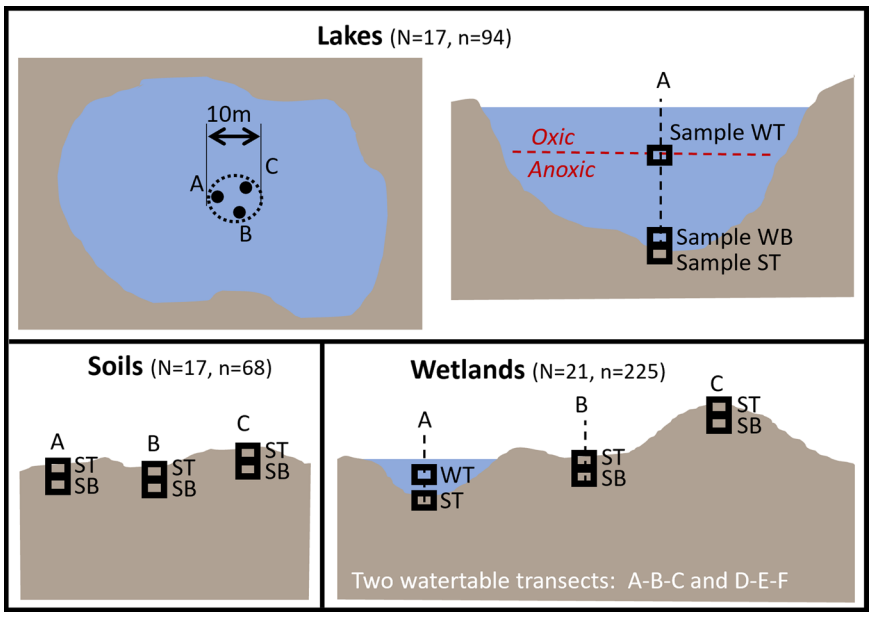



**Fig. 4** Sampling strategy for lake, soil and wetland sites (top, bottom left and bottom right panels, respectively). In lakes, at replicate points A, B and C, the water sample 'WT' was taken at the oxycline, and the water sample 'WB' just above sediment interface. One sediment sample was also collected. At soil sites, at replicate points A, B and C, two soil layers were sampled: 'ST' and 'SB' samples, representing respectively the top and bottom layers. In wetlands, two replicate transects were defined along the microtopography continuum hollow-edge-hummock. In hollows, one water and one solid sample were collected. At edges and hummocks, the same strategy as for soil sites was followed. For each type of sites, the number of sites and the corresponding number of events (*in situ* measurement and/or sampling) is indicated between parentheses.

Mineral soil samples were collected from three replicate points A, B, and C (at ~ 2–5 meters of distance from each other), considering two sampling depths for each point (Fig. 4).

In wetlands, microtopography is known to influence organic matter decomposition, $CH_4$ emissions, microbial community structure, and metabolic pathways[48–50]. The sampling strategy covered the three main microtopographic features of wetlands: hollows (i.e. small depressions, ponds, that can be filled with water or not at the time of sampling) (points A and D, Fig. 4); flat edges (or lawns) at the water table level or below, usually water-saturated and characterized by *Sphagnum* moss vegetation (points B and E, Fig. 4); and hummocks (i.e. dryer elevated mounts/raised domes, above the water table level, usually characterized by lichens and shrubs) (points C and F, Fig. 4). Two duplicate transects were considered, i.e. A-B-C and D-E-F transects, collected at ~ 10 meters of distance from each other. At each point, two sampling depths were considered, according to the same strategy as explained below for soils.

For both mineral and organic soils, soil blocks (20 × 20 × 20 cm blocks) were collected with a bread knife or a shovel. If soil layers could be clearly identified, top and bottom samples were defined accordingly and reported in the database. Otherwise, default depths were 0–10 cm for the surface layer and 10–20 cm for the bottom layer.

In addition to ecosystem-scale descriptors, every sample was characterized by point-scale descriptors (latitude, longitude, microtopography and vegetation type) and sample-scale descriptors such as environmental material (water, sediment, organic or mineral soil; Table 1 and Fig. 3). Soil samples were classified between organic and mineral soils using organic matter content (40% threshold) as the discriminating criterion between the two environmental materials[6].

The material and methods used for characterizing these samples *in situ* and in the laboratory are described in the following sections. In some sites (ALP3, ALS3, ALS4, ALS6, ALS8, ALS9, PCL3, PCP3, PCS1, PPL3, PPP3, PTL1, PTL2, PTP1, PTP2, PTS1, SIL5, SIP6, SIP7, SIS3, SIS4), a basic characterization was carried out, due to harsh conditions and limited access. This basic characterisation included restrained set of measured parameters as listed in Table 1, yet enabling to fully fill the objective of this project. All the other sites were fully characterized, including the whole set of measured parameters as listed in Table 1, according to the environmental package (water, sediment, soil).

***In situ* analyses.** *Physicochemical analyses.* At each sampling point and depth in lakes and hollows, dissolved oxygen, temperature, pH, conductivity, and redox potential were measured in water with a multiparametric probe (HI 9828, Hanna Instrument, Mexico). The detection limits for dissolved $O_2$ was 10 µg $L^{-1}$. In soil ecosystems, temperature was measured with an insertion thermometer (Isolab, Laborgerate GmbH).

*Dissolved $CH_4$ and $CO_2$ concentrations.* In lakes, the dissolved $CH_4$ and $CO_2$ concentrations were measured at each replicated sampling point and depth with the membrane-integrated cavity output spectrometry method

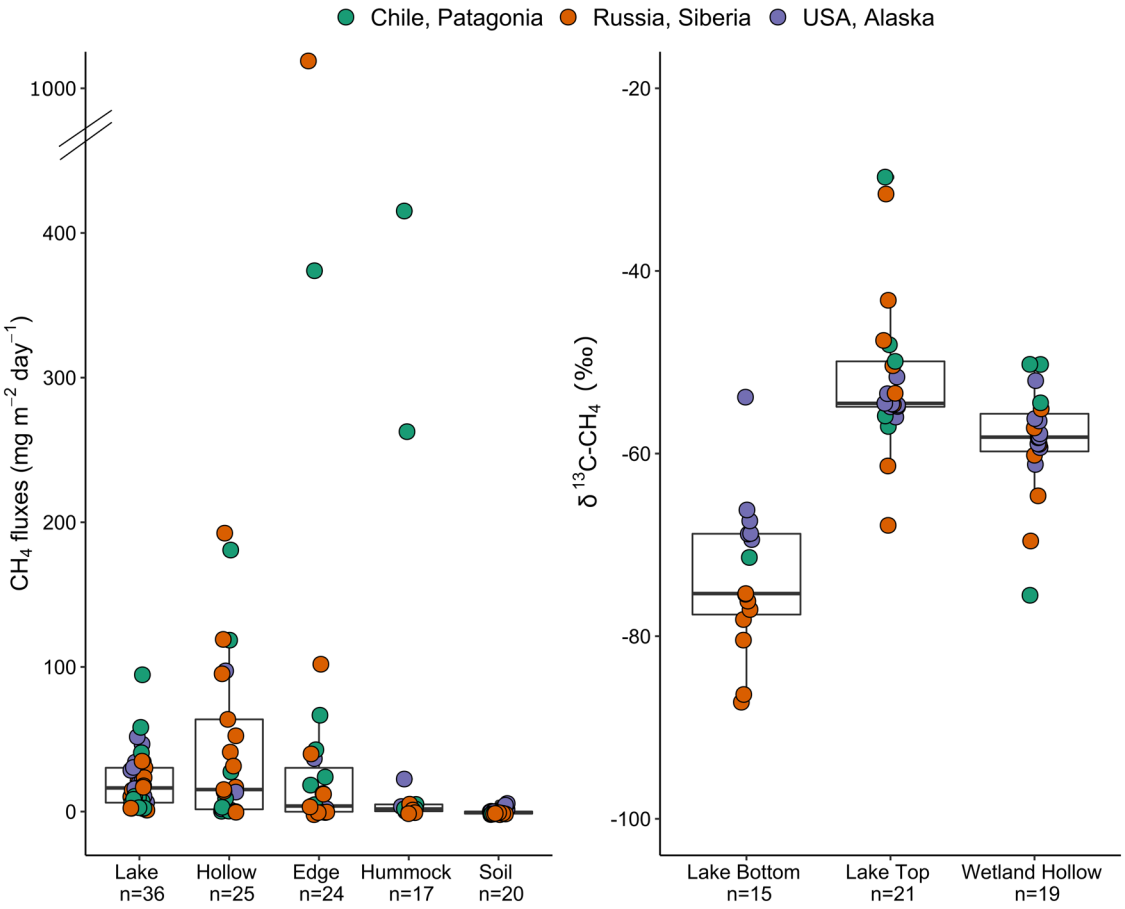



**Fig. 5** Methane emission rates measured during field campaigns (left) and $\delta^{13}$C-$CH_4$ fractionation (right). Methane emission measurements using static chambers were pooled according to meaningful categories that combined the environmental feature, environmental package and microtopography descriptors. The $\delta^{13}$C-$CH_4$ fractionation was measured in water samples only, i.e. in samples collected in the water column of lakes (at oxycline and at the bottom) and in hollows found in wetlands.

using an ultraportable greenhouse gas analyzer (UGGA, Los Gatos Research, USA)[51]. The detection limits for dissolved $CH_4$ and $CO_2$ concentrations were 5 nmol $L^{-1}$ and 4 μmol $L^{-1}$ respectively.

*Atmospheric $CH_4$ and $CO_2$ emission rates.* $CH_4$ and $CO_2$ emission rates were estimated with a static opaque chamber coupled in a loop to the UGGA (Los Gatos Research, USA), following the procedure described previously[9]. Briefly, a 0.102 $m^2$ floating chamber (7.8 L) was placed at the surface of lakes and ponds and a 0.035 $m^2$ chamber (12.3 L) was installed on soil sites. Accumulation of $CH_4$ and $CO_2$ was recorded during 5 min, and flux determined from the slope of $CH_4$ and $CO_2$. Then the chamber was ventilated and closed to perform another flux measurement. At least three replicate measurements were performed at each location (sampling points defined in Fig. 4). The static chamber method used measures the total flux at the surface, i.e. including both diffusive and ebullitive fluxes. As illustrated in Fig. 5a, the highest $CH_4$ emission rates were found in hollows, especially in Siberian peatlands of discontinuous permafrost and lakes.

**Sample processing in the field.** For further analysis, water subsamples were collected into 10 mL glass vials, directly in the field. For $\delta^{13}CCH_4$, $\delta^2HCH_4$, and total organic carbon (TOC) analysis, samples were acidified (HCl 6 N). For dissolved inorganic carbon (DIC) concentration and $\delta^{13}$C-DIC analysis, $HgCl_2$ was added to the samples to stop any biological activity. After fixation by HCl and/or$HgCl_2$, water subsamples were stored at 4 °C in dark conditions. Soil samples were also kept at 4 °C for 24 h maximum before further processing.

**Laboratory methods.** *Moisture and organic matter content.* Soil and sediment samples were dried at 110 °C overnight to determine the dry weight. Organic matter content was assessed via loss on ignition at 550 °C.

*Suspended solids.* Lake and hollow water samples (20 mL to 3 L, until clogging) were filtered on pre-weighted combusted GF/F grade glass microfiber filters (0.7 μm pore size, Whatman). The filters were dried overnight at 105 °C to calculate the total suspended solids (TSS). The filters were then incinerated at 550 °C for 2 hrs to determine the concentration of particulate organic matter (POM).

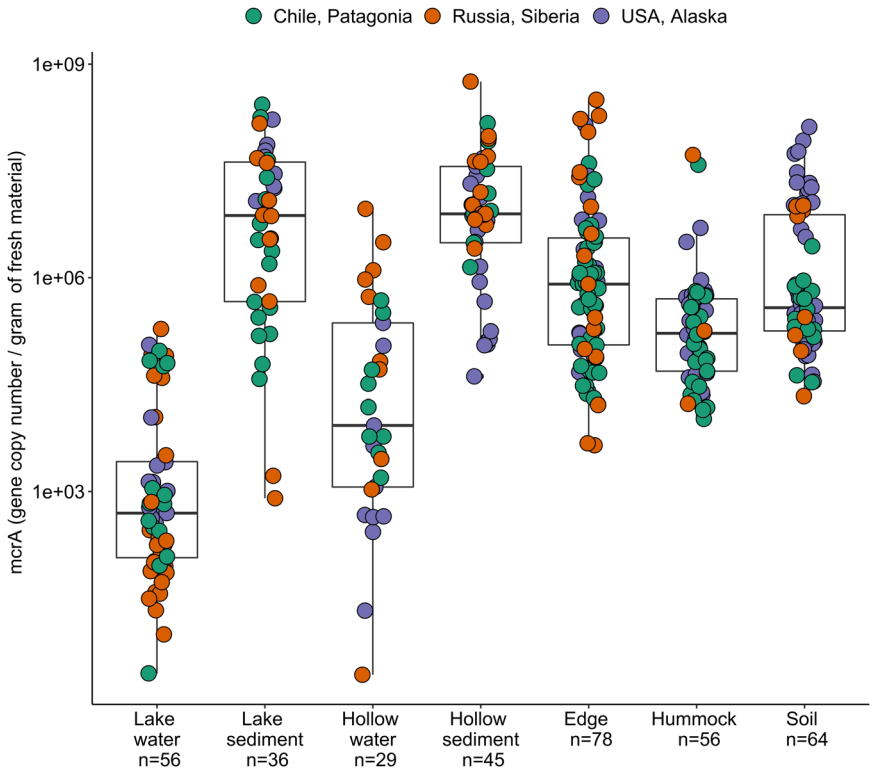



**Fig. 6** Abundance of methanogens in samples collected in Patagonia, Siberia and Alaska. Methanogen abundances were derived from qPCR assays targeting mcrA gene and were pooled according to environmental feature, environmental package and microtopography descriptors.

*Filtration.* After pre-filtration at 80 µm (nylon net filters, Merck Millipore, Cork Ireland), water samples were filtered at 0.22 µm (nitrocellulose GSWP membrane filters, Merck Millipore, Cork Ireland) up to filter clogging (corresponding to 636 ± 521 mL on average, ranging from 70 to 2930 mL depending on the highly variable suspended matter content of the samples). The filter was frozen at −20 °C prior to DNA extraction. The filtrate was recovered and used to prepare four vials for further analysis of dissolved organic carbon (DOC), the isotopic composition ($\delta^{13}C$) of DOC, optical properties of dissolved organic matter, cations, anions, and trace elements.

*Pore water extraction.* The water extraction was carried out on soil and sediment samples to assess the mobile fraction of DOC, major anions and cations, trace elements, and the optical properties of dissolved organic matter (DOM). Following the procedure recommended in Jones & Willet[52], 40 g of sample were placed in 200 mL of deionized water, and gently agitated with a magnetic stirrer at room temperature for 1 hr. The liquid phase was then recovered using a microRhizon sampler (Rhizosphere, Netherlands). The same procedure as for water samples was used to prepare and analyse these extracts.

*Total and dissolved organic carbon.* In water samples collected in lakes and hollows, TOC and DOC concentrations were analysed in using a TOC-V CSH analyser (Shimadzu, Japan). For DOC concentrations, samples were acidified to pH 2 using HCl 6 N and stored in 10 mL baked clear glass vials. The limit of quantification (LoQ) was 1 mg $L^{-1}$.

*Anions and cations.* Major ions were quantified in water samples collected in lakes and hollows and in pore water using a HPLC (Dionex, USA), a Dionex DX-120 analyser for cations (Thermo Fisher Scientific, France) and a Dionex ICS-5000 + analyser for anions (Thermo Fisher Scientific, France), according to recommandations[53]. The LoQ was 0.5 mg $L^{-1}$ for calcium, chloride, sulphate, and magnesium; 0.25 mg $L^{-1}$ for bromide, sodium, and potassium; 0.025 mg $L^{-1}$ for ammonium and phosphate; 0.01 mg $L^{-1}$ for fluoride, nitrate, and nitrite.

*Trace elements.* For trace element analysis, samples were acidified with ultrapure $HNO_3$ prior to ICP-MS (7500ce, Agilent Technologies) analysis, and kept in 15 mL polypropylene vials. LoQ were $<0.5\,\mu g\,g^{-1}$ for aluminium, iron, manganese, $<0.05\,\mu g\,g^{-1}$ for vanadium, chromium, cobalt, nickel, copper, zinc, and $<0.005\,\mu g\,g^{-1}$ for arsenic, strontium, cadmium, antimony, lead, uranium.

*Optical properties.* Subsamples were collected in 30-mL polypropylene vial for optical properties of DOM. The UV absorption spectra of pore water were measured with a spectrophotometer (Secoman UVi-lightXT5) from 190 to 700 nm in a 1 cm quartz cell. The specific UV absorbance at 254 nm (SUVA, L mg $C^{-1}$ $m^{-1}$) was calculated

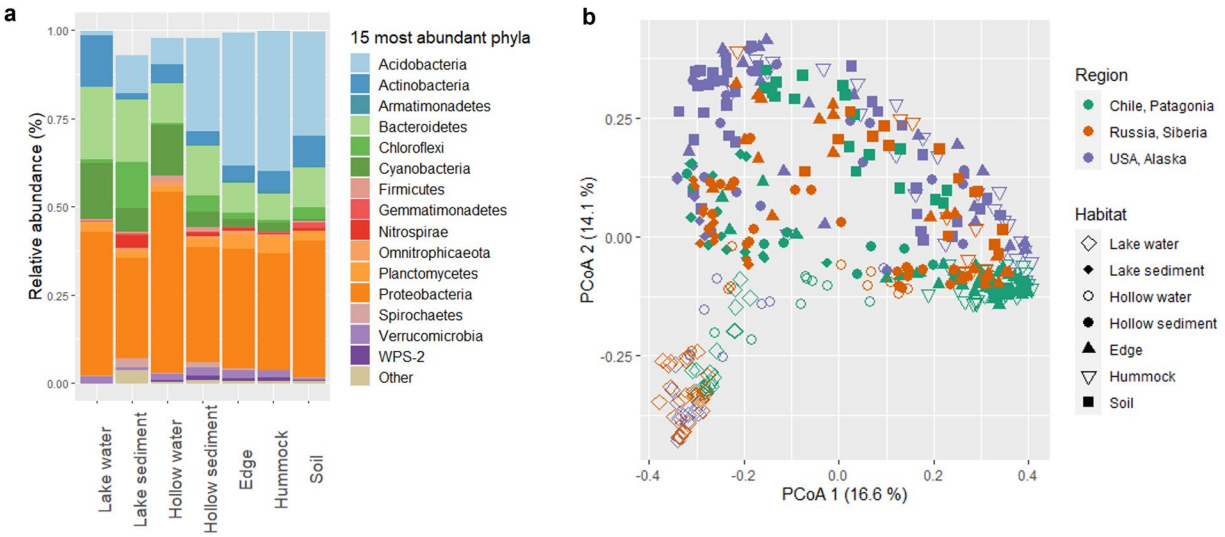



**Fig. 7** Taxonomic composition and similarities between the 387 microbial communities. The taxonomic composition of bacteria is presented at the phylum level, representing only the 15 most abundant phyla (panel **a**). Relative abundances of the phyla were calculated for seven habitats, i.e. a combination of environmental feature, environmental package and microtopography descriptors. Only the 15 more abundant phyla are displayed. The principal coordinate analysis (PCoA) of the filtered and standardized OTU abundance table was computed with Bray-Curtis distance to visualize similarities between microbial communities of the different habitats (indicated by the symbol shape) across the three studied regions (indicated by the symbol color; panel **b**). The percentage of total variance explained by each component is indicated along the axis, showing high microbial community variability mainly according to the different habitats.

as follows: SUVA = $A_{254}$/b*DOC[54], where $A_{254}$ is the sample absorbance at 254 nm (non-dimensional), b is the optical path length (m), and DOC is in mg $L^{-1}$. Fluorescence measurements were performed using a spectrofluorometer (Synergy MX, Biotek). The emission spectrum was recorded for a 370 nm excitation wavelength. The fluorescence Index (FI) was determined for a 370 nm excitation wavelength, as the ratio of the 470 nm emission to 520 nm emission[55,56].

*Isotopes.* The stable isotopic signature of methane ($\delta^{13}C$-$CH_4$, shown in Fig. 5b, and $\delta^{2}H$-$CH_4$) was analyzed at the Stable Isotope Facility of UC-Davis (https://stableisotopefacility.ucdavis.edu/methane-ch4-gas), using a ThermoScientific Precon concentration unit interfaced to a ThermoScientific Delta V Plus isotope ratio mass spectrometer (ThermoScientific, Germany). Methane was extracted for IRMS analysis following the method of Yarnes *et al.*[57]. The LoQ was 5 ppm of $CH_4$ for $\delta^{2}H$ and 1.7 ppm of $CH_4$ for $\delta^{13}C$, and standard deviation was typically 2‰ for $\delta^{2}H$ and 0.2‰ for $\delta^{13}C$. The $\delta^{13}C$-$CO_2$ was analyzed using a mass spectrometer (Isoprime 100, Elementar, UK) coupled with an equilibration system (MultiFlow-Geo, Elementar, UK). Samples were acidified using phosphoric acid and flushed with helium. The $\delta^{13}C$-DOC was analysed at the UC Davis Stable Isotope Facility, following the described procedure (http://stableisotopefacility.ucdavis.edu/doc.html). A TOC Analyzer (OI Analytical, College Station, TX) was interfaced to a PDZ Europa 20–20 isotope ratio mass spectrometer (Sercon Ltd., UK) utilizing a GD-100 Gas Trap Interface (Graden Instruments).

*DNA extraction.* Soil and sediments were subsampled and frozen at −20 °C. DNA was extracted from 0.5 g of the soil or sediment subsamples and from the previously frozen 0.22-µm filters using the PowerSoil and PowerWater DNA isolation kits, respectively (Qiagen, Hilden, Germany), following manufacturer instructions. The DNA extracts were stored at −20 °C.

*qPCR assay.* The abundances of four genes were measured by quantitative PCR (qPCR): bacterial 16S rRNA gene, archaeal 16S rRNA gene, *pmoA* gene (marker gene for aerobic methane oxidizing bacteria through the particulate methane monooxygenase), and *mcrA* gene (marker gene for methanogens and ANMEs through the methyl coenzyme M reductase). Duplicate measurements were run in 20 µL, using the Takyon SYBR master mix (Eurogentec, Belgium) with a CFX96 thermocycler (Bio-Rad Laboratories, Hercules, CA, US) and AriaMX thermocycler (Agilent, CA, US). Primer sequences and concentrations, thermocycling conditions, and standard curve preparation were detailed in Thalasso *et al.*[18]. As an illustration, the abundance of *mcrA* gene according to habitat (i.e. category combining the environmental material and the microtopography) is displayed in Fig. 6.

*High-throughput amplicon sequencing.* Archaeal and bacterial diversity was assessed using metabarcoding and targeting the V4-V5 region of 16S rRNA gene. Amplicons were obtained from DNA extracts using 515 F (GTGYCAGCMGCCGCGGTA) and 928 R (CCCCGYCAATTCMTTTRAGT) primers[58]. MTP Taq DNA polymerase was acquired from Sigma (France). The thermocycling procedure was the following: 2 min at 94 °C; 30

cycles of 60 s at 94 °C, 40 s at 65 °C, and 30 s at 72 °C; and finally, 10 min at 72 °C. PCR products were used for pair-end sequencing using Illumina Miseq (2 × 250-bp). After pre-processing of raw reads through the FROGS pipeline[59], a total of 18 369 310 sequences were obtained from the 387 samples, and clustered into 121 971 OTUs using Swarm[60]. The OTUs were further filtered at 0.005% of relative abundance, as previously recommended[61], and taxonomically annotated against SILVA 132 rRNA database. Community analysis was carried out in R software, version 4.1.1, with '*phyloseq*' package[62]. The taxonomic composition of bacteria according to habitat was represented by a barplot at the phylum level (Fig. 7a). As an illustration of the microbial diversity outcomes from this dataset and the community variability according to the different habitats, the dissimilarity among the 387 community structures was visualized by a principal coordinate analysis PCoA, a.k.a. Multidimensional scaling (MDS) with the *ordinate* function using Bray Curtis distance ('phyloseq' package[62]) computed on the filtered and standardized (percentage) OTUs relative abundances (Fig. 7b).

## Data Records

This paper presents a combination of sample metadata, environmental data (gas flux and biogeochemical measurements), and high-throughput microbiome sequencing data co-located in time and space. Linking these data of different nature is crucial for their effective interpretation and reuse. The geo-referenced dataset was documented in the DRYAD platform[63] and is fully available under a CC0 1.0 Universal (CC0 1.0) Public Domain Dedication license. The dataset is publicly accessible with the following doi: 10.5061/dryad.rfj6q57dp. The dataset in DRYAD also includes a 'readme' file intended to provide key information for understanding and reuse of the dataset. The data is organized in a standard datasheet table easily downloadable (csv format) with 387 samples (in rows) and 120 parameters (in columns). The first row of the table is the parameter name and the second row of the table is the unit of each parameter. The parameters are organized as follows: Sampling context; Ecosystem characteristics; Sequencing method details; Basic physicochemical parameters; Organic matter characteristics; Nutrients, anions, cations; Greenhouse gases; qPCR quantifications; Micro elements; Bioclimatic variables. The data is easily downloadable in csv format, and the clear and unique sample ID enables to link the data to the sequence set. The raw sequence data in FASTQ format without preprocess, were archived in the European Nucleotide Archive (ENA) with accession codes PRJEB36731 (Siberia)[64], PRJEB36732 (Alaska)[65], and PRJEB36733 (Patagonia)[66]. These microbial datasets along with sample metadata have been published in the Global Biodiversity Information Facility (GBIF)[67–69] separately for Patagonia[67], Siberia[68], and Alaska[69]. Standardized information about sequence data[70] were reported together with environmental data, and formatted as defined by the Genomic Standards Consortium[71], based on MIMARKS sheet for miscellaneous natural environment.

## Technical Validation

Operator training and strategic harmonization for meta(data) collection occurred at the beginning of the first field campaign to ensure all operators used identical and replicable methods in terms of data acquisition in field, sampling, sample processing in the field and in the laboratory, and data recording. All data were checked and accurately transferred to MIMARKS database. The database was eventually manually curated by a dedicated data manager.

During $CH_4$ and $CO_2$ flux measurement, two criteria were tested before emission trends were validated[72]: (i) that the initial concentration was nearly equal to ambient atmospheric concentration; and (ii) that the linear correlation coefficient ($R^2$) from the regression analysis reached 0.90. When a measurement did not meet these criteria, additional replicates were done, which occurred in only a few occasions.

Reference material ION-915 and ION 96.4, both acquired from Environment and Climate Change Canada (Canada), were included in the analytical loop of TOC and major anions and cations determination. Recovery was >95% of the certified value. The trace element certified river water[53] SLRS6 (National Research Council – Conseil National de Recherches Canada) was used as a reference material on every run for ICPMS analysis, with indium as an internal standard, and accuracy (i.e. recovery >95%) was checked. The analytical routine included the analysis of blanks, calibration standards, and a multi-element quality control solution (EPOND) every 12 samples.

For isotopic analysis of $\delta^{13}C$-$CO_2$ analysis, standards included $Na_2CO_3$ and $NaHCO_3$ as well as internal water standards, that were analyzed every 8 samples to check for instrument stability. All samples were analyzed in replicates. Standard deviation was typically around or below 0.2‰.

Blanks (sterile pure water) were included in the DNA extraction process, PCR, and qPCR protocols. The absence of amplification on negative controls (contamination) was checked by gel electrophoresis. The correct size of 16S rRNA amplicons and the PCR specificity (unique band) were also checked by gel electrophoresis.

For qPCR of bacterial and archaeal 16S rRNA gene and *mcrA* gene, standard curves were prepared from 10-fold serial dilutions of each target gene, amplified from the following pure strains: *Pseudomonas stutzeri* SLG510A3-8 (KT153610 accession number), Arch_21F_10-Berre_sed clone (KT351355 accession number), and *Methanosarcina barkeri* CM1 (AKJ39604 accession number), respectively, and cloned in pGEM-T plasmid (Promega). For *pmoA*, the standard was synthetized by Eurofins from *Methylobacter* sp. BB5.1 *pmoA* gene sequence (AF016982 accession number), inserted in TOPO-TA pCR2.1 plasmid. qPCR efficiencies were always >90% and amplicon size and specificity were confirmed by melting curve analysis and agarose gels.

Two samples with less than 1325 sequences retrieved from high-throughput sequencing were discarded from the sequence datasets deposited in the European Nucleotide Archive.

## Code availability

No custom code was used to process the dataset presented in this article. To visualize microbial structure and composition, the diversity data provided in this dataset have been processed through the FROGS pipeline[59] version 3.2.3 (http://frogs.toulouse.inra.fr/), available on the Galaxy server (https://vm-galaxy-prod.toulouse.inrae.fr/galaxy_main/) of the Genotoul bioinformatic platform (http://bioinfo.genotoul.fr/) under GNU GPLv3 licence.

Received: 21 April 2022; Accepted: 4 October 2022;
Published online: 04 November 2022

## References

1. Saunois, M. *et al*. The Global Methane Budget 2000–2017. *Earth Syst. Sci. Data* **12**, 1561–1623 (2020).
2. Dodds, W. K. *et al*. The freshwater biome gradient framework: predicting macroscale properties based on latitude, altitude, and precipitation. *Ecosphere* **10** (2019).
3. Lee, J. Y. *et al*. Future Global Climate: Scenario-Based Projections and Near-Term Information. in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press. IPCC, 2021, 1–195).
4. Cohen, J. *et al*. Recent Arctic amplification and extreme mid-latitude weather. *Nature Geosci* **7**, 627–637 (2014).
5. Myhre, G. *et al*. Anthropogenic and Natural Radiative Forcing. in *Climate Change 2013: The Physical Science Basis* vol. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley, 2013).
6. Schuur, Ea. G. *et al*. Climate change and the permafrost carbon feedback. *Nature* **520**, 171–179 (2015).
7. Rozzi, R. *et al*. Integrating Ecology and Environmental Ethics: Earth Stewardship in the Southern End of the Americas. *BioScience* **62**, 226–236 (2012).
8. Rozzi, R. *et al*. Changing lenses to assess biodiversity: patterns of species richness in sub-Antarctic plants and implications for global conservation. *Frontiers in Ecology and the Environment* **6**, 131–137 (2008).
9. Gerardo-Nieto, O., Astorga-españa, M. S., Mansilla, A. & Thalasso, F. Initial report on methane and carbon dioxide emission dynamics from sub-Antarctic freshwater ecosystems: A seasonal study of a lake and a reservoir. *Science of the Total Environment* **593–594**, 144–154 (2017).
10. Münchberger, W., Knorr, K.-H., Blodau, C., Pancotto, V. A. & Kleinebecker, T. Zero to moderate methane emissions in a densely rooted, pristine Patagonian bog - biogeochemical controls as revealed from isotopic evidence. *Biogeosciences Discuss*. 1–35, https://doi.org/10.5194/bg-2018-301 (2018).
11. Conrad, R. Importance of hydrogenotrophic, aceticlastic and methylotrophic methanogenesis for methane production in terrestrial, aquatic and other anoxic environments: A mini review. *Pedosphere* **30**, 25–39 (2020).
12. Günthel, M. *et al*. Reply to 'Oxic methanogenesis is only a minor source of lake-wide diffusive CH4 emissions from lakes'. *Nat Commun* **12**, 1205 (2021).
13. Grossart, H.-P., Frindte, K., Dziallas, C., Eckert, W. & Tang, K. W. Microbial methane production in oxygenated water column of an oligotrophic lake. *Proceedings of the National Academy of Sciences* **108**, 19657–19661 (2011).
14. Peeters, F. & Hofmann, H. Oxic methanogenesis is only a minor source of lake-wide diffusive CH4 emissions from lakes. *Nat Commun* **12**, 1206 (2021).
15. DelSontro, T., del Giorgio, P. A. & Prairie, Y. T. No Longer a Paradox: The Interaction Between Physical Transport and Biological Processes Explains the Spatial Distribution of Surface Water Methane Within and Across Lakes. *Ecosystems* **21**, 1073–1087 (2018).
16. Bižić-Ionescu, M., Ionescu, D., Günthel, M., Tang, K. W. & Grossart, H.-P. Oxic Methane Cycling: New Evidence for Methane Formation in Oxic Lake Water. in *Biogenesis of Hydrocarbons* (eds. Stams, A. J. M. & Sousa, D.) 1–22, https://doi.org/10.1007/978-3-319-53114-4_10-1 (Springer International Publishing, 2018).
17. Bastviken, D., Cole, J. J., Pace, M. L. & Van de Bogert, M. C. Fates of methane from different lake habitats: Connecting whole-lake budgets and $CH_4$ emissions: FATES OF LAKE METHANE. *J. Geophys. Res.* **113**, n/a–n/a (2008).
18. Thalasso, F. *et al*. Sub-oxycline methane oxidation can fully uptake CH4 produced in sediments: case study of a lake in Siberia. *Sci Rep* **10**, 3423 (2020).
19. Knief, C. Diversity and Habitat Preferences of Cultivated and Uncultivated Aerobic Methanotrophic Bacteria Evaluated Based on pmoA as Molecular Marker. *Front. Microbiol*. **6** (2015).
20. Cabrol, L. *et al*. Anaerobic oxidation of methane and associated microbiome in anoxic water of Northwestern Siberian lakes. *Science of The Total Environment* **736**, 139588 (2020).
21. Martinez-Cruz, K. *et al*. Ubiquitous and significant anaerobic oxidation of methane in freshwater lake sediments. *Water Research* **144**, 332–340 (2018).
22. Miller, K., Lai, C.-T., Dahlgren, R. & Lipson, D. Anaerobic Methane Oxidation in High-Arctic Alaskan Peatlands as a Significant Control on Net CH4 Fluxes. *Soil Syst.* **3**, 7 (2019).
23. Gupta, V. *et al*. Stable Isotopes Reveal Widespread Anaerobic Methane Oxidation Across Latitude and Peatland Type. *Environ. Sci. Technol*. 130717064455005, https://doi.org/10.1021/es400484t (2013).
24. Joye, S. B. A piece of the methane puzzle. *Nature* **491**, 538–539 (2012).
25. Ettwig, K. F. *et al*. Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature* **464**, 543–548 (2010).
26. Oswald, K. *et al*. Light-Dependent Aerobic Methane Oxidation Reduces Methane Emissions from Seasonally Stratified Lakes. *PLoS ONE* **10**, e0132574 (2015).
27. Juncher Jørgensen, C., Lund Johansen, K. M., Westergaard-Nielsen, A. & Elberling, B. Net regional methane sink in High Arctic soils of northeast Greenland. *Nature Geosci* **8**, 20–23 (2015).
28. Oh, Y. *et al*. Reduced net methane emissions due to microbial methane oxidation in a warmer Arctic. *Nat. Clim. Chang.* **10**, 317–321 (2020).
29. Rosentreter, J. A. Half of global methane emissions come from highly variable aquatic ecosystem sources. *Nature Geoscience* **14**, 22 (2021).
30. Hugelius, G. *et al*. Large stocks of peatland carbon and nitrogen are vulnerable to permafrost thaw. *Proc Natl Acad Sci USA* **117**, 20438–20446 (2020).
31. Frolking, S. *et al*. Peatlands in the Earth's 21st century climate system. *Environ. Rev.* **19**, 371–396 (2011).
32. Treat, C. C. *et al*. The role of wetland expansion and successional processes in methane emissions from northern wetlands during the Holocene. *Quaternary Science Reviews* **257**, 106864 (2021).
33. Turetsky, M. R. *et al*. Short-term response of methane fluxes and methanogen activity to water table and soil warming manipulations in an Alaskan peatland. *J. Geophys. Res.* **113**, G00A10 (2008).
34. McLaughlin, J. & Webster, K. Effects of Climate Change on Peatlands in the Far North of Ontario, Canada: A Synthesis. *Arctic, Antarctic, and Alpine Research* **46**, 84–102 (2014).

35. Sepulveda-Jauregui, A., Walter Anthony, K. M., Martinez-Cruz, K., Greene, S. & Thalasso, F. Methane and carbon dioxide emissions from 40 lakes along a north–south latitudinal transect in Alaska. *Biogeosciences* **12**, 3197–3223 (2015).
36. Rask, H., Schoenau, J. & Anderson, D. Factors in¯uencing methane ¯ux from a boreal forest wetland in Saskatchewan, Canada. *Soil Biology* 9 (2002).
37. Ye, R. *et al.* pH controls over anaerobic carbon mineralization, the efficiency of methane production, and methanogenic pathways in peatlands across an ombrotrophic–minerotrophic gradient. *Soil Biology and Biochemistry* **54**, 36–47 (2012).
38. Boetius, A. Global change microbiology — big questions about small life for our future. *Nat Rev Microbiol* **17**, 331–332 (2019).
39. Cavicchioli, R. *et al.* Scientists' warning to humanity: microorganisms and climate change. *Nat Rev Microbiol* **17**, 569–586 (2019).
40. Buttigieg, P. *et al.* The environment ontology: contextualising biological and biomedical entities. *J Biomed Sem* **4**, 43 (2013).
41. Brown, J., Ferrians, O., Heginbottom, J. A. & Melnikov, E. Circum-Arctic Map of Permafrost and Ground-Ice Conditions, *NSIDC* https://doi.org/10.7265/skbg-kf16 (2002).
42. Olson, D. M. *et al.* Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience* **51**, 933 (2001).
43. Streletskiy, D. A. *et al.* Permafrost hydrology in changing climatic conditions: seasonal variability of stable isotope composition in rivers in discontinuous permafrost. *Environ. Res. Lett.* **10**, 095003 (2015).
44. Gandois, L. *et al.* Contribution of Peatland Permafrost to Dissolved Organic Matter along a Thaw Gradient in North Siberia. *Environ. Sci. Technol.* **53**, 14165–14174 (2019).
45. Pfeiffer, M. *et al.* CHLSOC: the Chilean Soil Organic Carbon database, a multi-institutional collaborative effort. *Earth System Science Data* **12**, 457–468 (2020).
46. Lavergne, C. *et al.* Temperature differently affected methanogenic pathways and microbial communities in sub-Antarctic freshwater ecosystems. *Environment International* **154**, 106575 (2021).
47. Fick, S. & Hijmans, R. Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, https://doi.org/10.1002/joc.5086 (2017).
48. Laiho, R. Decomposition in peatlands: Reconciling seemingly contrasting results on the impacts of lowered water levels. *Soil Biology and Biochemistry* **38**, 2011–2024 (2006).
49. Liebner, S. *et al.* Shifts in methanogenic community composition and methane fluxes along the degradation of discontinuous permafrost. *Front. Microbiol.* **6** (2015).
50. Singleton, C. M. *et al.* Methanotrophy across a natural permafrost thaw environment. *ISME J* **12**, 2544–2558 (2018).
51. Gonzalez-valencia, R. *et al.* *In Situ* Measurement of Dissolved Methane and Carbon Dioxide in Freshwater Ecosystems by Off -Axis Integrated Cavity Output Spectroscopy. *Environmental Science and Technology* **48**, 11421–11428 (2014).
52. Jones, D. L. & Willett, V. B. Experimental evaluation of methods to quantify dissolved organic nitrogen (DON) and dissolved organic carbon (DOC) in soil. *Soil Biology and Biochemistry* **38**, 991–999 (2006).
53. Yeghicheyan, D. *et al.* A New Interlaboratory Characterisation of Silicon, Rare Earth Elements and Twenty-Two Other Trace Element Concentrations in the Natural River Water Certified Reference Material SLRS-6 (NRC-CNRC). *Geostandards and Geoanalytical Research* **43**, 475–496 (2019).
54. Weishaar, J. L. *et al.* Evaluation of Specific Ultraviolet Absorbance as an Indicator of the Chemical Composition and Reactivity of Dissolved Organic Carbon. *Environ. Sci. Technol.* **37**, 4702–4708 (2003).
55. McKnight, D. M. *et al.* Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. *Limnology and Oceanography* **46**, 38–48 (2001).
56. Jaffé, R. *et al.* Spatial and temporal variations in DOM composition in ecosystems: The importance of long-term monitoring of optical properties. *Journal of Geophysical Research: Biogeosciences* **113** (2008).
57. Yarnes, C. $\delta^{13}C$ and $\delta^{2}H$ measurement of methane from ecological and geological sources by gas chromatography/combustion/pyrolysis isotope-ratio mass spectrometry. *Rapid Communications in Mass Spectrometry* **27**, 1036–1044 (2013).
58. Wang, Y. & Qian, P.-Y. Conservative Fragments in Bacterial 16S rRNA Genes and Primer Design for 16S Ribosomal DNA Amplicons in Metagenomic Studies. *PLOS ONE* **4**, e7401 (2009).
59. Escudié, F. *et al.* FROGS: Find, Rapidly, OTUs with Galaxy Solution. *Bioinformatics* **34**, 1287–1294 (2018).
60. Mahé, F., Rognes, T., Quince, C., de Vargas, C. & Dunthorn, M. Swarm: robust and fast clustering method for amplicon-based studies. *PeerJ* **2**, e593 (2014).
61. Bokulich, N. A. *et al.* Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods* **10**, 57–59 (2013).
62. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **8**, e61217 (2013).
63. Barret, M. *et al.* A combined microbial and biogeochemical dataset from high-latitude ecosystems with respect to methane cycle. *Dryad, Dataset* https://doi.org/10.5061/dryad.rfj6q57dp (2022).
64. *European Nucleotide Archive*, https://identifiers.org/ena.embl:PRJEB36731 (2020).
65. *European Nucleotide Archive*, https://identifiers.org/ena.embl:PRJEB36732 (2020).
66. *European Nucleotide Archive*, https://identifiers.org/ena.embl:PRJEB36733 (2020).
67. Barret, M. *et al.* Bacteria and Archaea biodiversity in terrestrial ecosystems of Chilean Patagonia. SCAR - Microbial Antarctic Resource System. *Metadata dataset. GBIF.org* https://doi.org/10.15468/wv6xcg (2019).
68. Barret, M. *et al.* Bacteria and Archaea biodiversity in Arctic terrestrial ecosystems affected by climate change in Northern Siberia. SCAR - Microbial Antarctic Resource System. *Metadata dataset GBIF.org* https://doi.org/10.15468/dooh47 (2019).
69. Barret, M. *et al.* Bacteria and Archaea biodiversity in Arctic and Subarctic terrestrial ecosystems in Alaska. SCAR - Microbial Antarctic Resource System. *Metadata dataset GBIF.org* https://doi.org/10.15468/hhkhz2 (2019).
70. Yilmaz, P. *et al.* Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nat Biotechnol* **29**, 415–420 (2011).
71. Field, D. *et al.* The Genomic Standards Consortium. *PLoS Biol* **9**, e1001088 (2011).
72. Duchemin, E., Lucotte, M. & Canuel, R. Comparison of Static Chamber and Thin Boundary Layer Equation Methods for Measuring Greenhouse Gas Emissions from Large Water Bodies§. *Environ. Sci. Technol.* **33**, 350–357 (1999).


## Acknowledgements

We acknowledge MAEDI and MENESR French ministries and CONICYT (Chile) for financial support through the ERANet-LAC joint program METHANOBASE (ELAC2014_DCC-0092), and the University of Toulouse (France) and Toulouse INP for supporting fieldwork costs. We acknowledge the financial support of the Research Council of Norway project 256132/H30. We also gratefully acknowledge Conacyt (Mexico) for the financial support to Oscar Gerardo-Nieto (grant # 277238). We thank the ECOS Sud-CONICYT Project "MATCH" C16B03 for mobility and the ANR-20-CE01-0001 project for financial contribution. We thank the GeT-PlaGe Genotoul platform (Toulouse, France) for performing the MiSeq sequencing. We are grateful to Anatoly Pimov (Igarka) for field assistance, Frédéric Julien, Virginie Payre-Suc and Daniel Lambrigot for DOC and major elements analysis (PAPC platform, "Laboratoire Ecologie Fonctionnelle et Environnement" laboratory,

France), and Issam Moussa and Daniel Dalger for $\delta^{13}C$- DIC analysis (SHIVA platform, "Laboratoire Ecologie Fonctionnelle et Environnement", France). We thank Mary-Beth Leigh and Katey Walter Anthony for providing access to laboratories and facilities at, respectively, the Institute of Arctic Biology - Department of Biology and Wildlife, and the Water and Environmental Research Center (WERC), Institute of Northern Engineering, University of Alaska, Fairbanks, US (UAF). The authors also recognize the support and access to the NSF funded Toolik Field Station and the Poker Flat Research Range from the University of Alaska, Fairbanks (US). We are grateful to Karin Gerard and Claudio Gonzalez-Wevar for providing access to the "Laboratorio de Ecosistemas Marinos Antarticos y Subantarticos", Universidad de Magallanes, Punta Arenas (Chile). We thank Erwin Domínguez Díaz for access to the "Instituto de Investigaciones Agropecuarias" (INIA) field sites in Patagonia.

## Author contributions

Maialen Barret: Coordination, Funding and administrative management, Conceptualization, Field/lab work, Investigation, Methodology, Supervision, Data management, Writing - original draft, Draft revision. Laure Gandois: Field/lab work, Investigation, Methodology, Writing - original draft, Draft revision. Frederic Thalasso: Funding and administrative management, Conceptualization, Field/lab work, Investigation, Methodology, Draft revision. Karla Martinez Cruz: Field/lab work, Investigation, Methodology, Draft revision. Armando Sepulveda Jauregui: Field/lab work, Investigation, Methodology, Formal analysis. Céline Lavergne: Field/lab work, Investigation, Methodology, Supervision, Draft revision. Roman Teisserenc Field/lab work, Investigation. Polette Aguilar: Field/lab work, Investigation. Oscar Gerardo-Nieto: Field/lab work, Investigation. Claudia Etchebehere: Funding and administrative management, Conceptualization, Field/lab work, Investigation, Draft revision. Bruna Martins Dellagnezze: Field/lab work, Investigation. Patricia Bovio-Wrinkler: Field/lab work, Investigation. Gilberto J. Fochesatto: Field/lab work, logistic support. Nikita Tananaev: Field/lab work, logistic support. Mette M. Svenning: Funding and administrative management, Supervision, Investigation, Draft revision. Christophe Seppey: Investigation, Draft revision. Alexander Tveit: Field/lab work, Investigation, Methodology, Draft revision. Rolando Chamy: Funding and administrative management, Supervision. María Soledad Astorga-España: Field/lab work, logistic support, funding and administrative management. Andrés Mansilla: Logistic support, supervision. Anton Van de Putte: Funding and administrative management, Data management. Maxime Sweetlove: Data management. Alison E. Murray: Funding and administrative management, Conceptualization, Data management, Draft revision. Léa Cabrol: Coordination, Funding and administrative management, Conceptualization, Field/lab work, Investigation, Methodology, Supervision, Data management, Writing - original draft, Draft revision.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41597-022-01759-8.

**Correspondence** and requests for materials should be addressed to M.B. or L.C.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.