environmental microbiology

Environmental Microbiology (2018) **20**(12), 4328–4342 doi:10.1111/1462-2920.14348

Disentangling the complexity of permafrost soil by using high resolution profiling of microbial community composition, key functions and respiration rates

Oliver Müller,[1*] Toke Bang-Andreasen,[2,3] Richard Allen White III,[4] Bo Elberling,[5] Neslihan Taş,[6] Timothy Kneafsey,[6] Janet K. Jansson[4] and Lise Øvreås[1,7]

[1]*Department of Biological Sciences, University of Bergen, N-5020, Bergen, Norway.*

[2]*Department of Environmental Science, Aarhus University, DK-4000, Roskilde, Denmark.*

[3]*Department of Biology, University of Copenhagen, DK-2100, Copenhagen, Denmark.*

[4]*Pacific Northwest National Laboratory, Richland, WA 99354, USA.*

[5]*Department of Geosciences and Natural Resource Management, Center for Permafrost (CENPERM), University of Copenhagen, DK-1350, Copenhagen, Denmark.*

[6]*Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.*

[7]*University Center in Svalbard, UNIS, N-9171, Longyearbyen, Norway.*

Summary

Thawing permafrost can stimulate microbial activity, leading to faster decomposition of formerly preserved organic matter and CO_2 release. Detailed knowledge about the vertical distribution of the responsible microbial community that is changing with increasing soil depth is limited. In this study, we determined the microbial community composition from cores sampled in a high Arctic heath at Svalbard, Norway; spanning from the active layer (AL) into the permafrost layer (PL). A special aim has been on identifying a layer of recently thawed soil, the transition zone (TZ), which might provide new insights into the fate of thawing permafrost. A unique sampling strategy allowed us to observe a diverse and gradually shifting microbial community in the AL, a Bacteroidetes dominated community in the TZ and throughout the PL, a community strongly dominated by a single Actinobacteria family (*Intrasporangiaceae*). The contrasting abundances of these two taxa caused a community difference of about 60%, just within 3 cm from TZ to PL. We incubated subsamples at about 5°C and measured highest CO_2 production rates under aerobic incubations, yet contrasting for five different layers and correlating to the microbial community composition. This high resolution strategy provides new insights on how microbial communities are structured in permafrost and a better understanding of how they respond to thaw.

Received 5 February, 2018; revised 27 June, 2018; accepted 28 June, 2018. *For correspondence. E-mail oliver.muller@uib.no; Tel. +47 55 58 44 00; Fax: +47 55 58 96 43.

Introduction

Permafrost constitutes 25% of Earth's terrestrial surface and stores a vast amount of buried, ancient carbon (C) equaling about half of the global belowground soil organic matter (SOM) pool (Hugelius *et al.,* 2014). The current warming in the Arctic is therefore of special concern, as this may trigger increased microbial activity, leading to faster decomposition of formerly preserved organic matter and release of greenhouse gases such as carbon dioxide (CO_2) and methane (CH_4) (Schuur *et al.,* 2009; Grosse *et al.,* 2011; Mackelprang *et al.,* 2011; Xue *et al.,* 2016).

Permafrost thaw has expanded dramatically across the Arctic, as measured by the increasing extent of active layer thickness in the soil (Jorgenson *et al.,* 2001; Åkerman and Johansson, 2008; Romanovsky *et al.,* 2010; Hayes *et al.,* 2014). The active layer (AL) is the upper part of the permafrost, undergoing seasonal freezing–thawing cycles, while the permafrost layer (PL) remains constantly frozen throughout the year. Together, this is the underlying reason for the microbial differences between those layers (Jansson and Taş, 2014).

Several studies of Arctic permafrost have shown a higher microbial diversity, biomass and activity in the AL, which is decreasing towards the PL (Yergeau *et al.,* 2010; Mackelprang *et al.,* 2011; Frank-Fahle *et al.,* 2014; Taş *et al.,* 2014; Gittel *et al.,* 2014a,b; Deng *et al.,* 2015).

However, the impact of microbes on SOM degradation processes and greenhouse gas production remains unclear. Only a few studies have combined microbial community composition and activity measurements. Interestingly, experiments showed that microbial communities in the PL changed rapidly in structure and function upon thaw, indicating that the newly available SOM can be processed instantly (Mackelprang *et al.,* 2011; Deng *et al.,* 2015). Whether this degradation will result in increased CO_2 or CH_4 fluxes depends strongly on conditions at the permafrost site, including soil type, water content and if aerobic or anaerobic conditions are dominating (Lee *et al.,* 2012; Elberling *et al.,* 2013). Additionally, the microbial contribution to greenhouse gas emissions is not fully understood (Elberling *et al.,* 2004; McCalley *et al.,* 2014).

One important challenge in all soil studies is the heterogeneous character of a soil profile (Elkateb *et al.,* 2003). Nevertheless, most permafrost studies compare microbial communities based on broad scales, comparing surface (AL) with deep (PL) samples and knowledge regarding fine-scale shifts throughout the soil core is still scarce. In this study, we therefore investigated the changes in microbial community structure along a fine scaled depth profile, following, in a high resolution (every 3–4 cm), the transition from AL into the PL in a 2-m soil core from Svalbard. For activity measurements, five different depths were chosen based on the community differences identified in the high resolution analysis. Subsamples from the five zones from this core and another core from the same sampling site were thawed, incubated at 4°C (±1) and gas fluxes were monitored (CO_2, CH_4 and N_2O) over intervals of hours to months.

Our study aims to precisely capture the changes in microbial community composition throughout a 2-m soil profile, spanning from the AL into the permafrost and to identify potential connections between community structure and greenhouse gas fluxes. This information is crucial to improve computational predictions of climate change effects that include microbial processes (Schwalm *et al.,* 2010). To better understand these processes, we need detailed knowledge on how permafrost communities are structured and on their metabolic response to thaw.

Results

Soil characteristics

Several chemical parameters were analysed for core 1 as this core was used for the high resolution profile of 16S rRNA gene sequencing data. Overall, the soil structure throughout the core was found to be homogenous (Fig. 1A), as confirmed using X-ray Computer Tomography (Fig. 1E), where estimated bulk densities ranged between 0.3 and 0.4 g cm^{-3}. The soil sampling site is covered with Late Holocene loess deposits and is characterized by a high silt/sand content of up to ~47% (Bang-Andreasen *et al.,* 2017). Due to the high silt/sand content the site is relatively well-drained. In this location, the combination of continuous sedimentation and freezing conditions resulted in thickening of the PL, where organic C surface layers were buried with time and can be observed throughout the core in thin layers (Fig. 1A). This is reflected in the homogenous character of soil properties such as total carbon (TC), total nitrogen (TN) and pH (Fig. 1F).

Microbial community composition

The high resolution analysis of 16S rRNA genes from core 1 showed characteristic differences in relative abundance of the predominant microbial groups originating from the AL and PL (Fig. 1B) and these differences were also observed in corresponding depths of core 2 (Fig. 4C). In the AL, the most abundant phyla were Acidobacteria (14% on average), Actinobacteria (9%), Proteobacteria (24%) and Verrucomicrobia (16%) (Figs 1B and 4C). The community composition in the PL was entirely different and mainly dominated by Actinobacteria with an average relative abundance of 70% (±13%) in core 1 and 41% (±1%) in core 2 (Figs 1B and 4C). There were only four operational taxonomic units (OTUs) belonging to *Intrasporangiaceae* (family) that corresponded to the Actinobacteria sequences. In both cores, Archaea were generally underrepresented with a maximum relative abundance of 0.5% and OTUs were predominantly assigned to Crenarchaeota (class MBGA) and Euryarchaeota (class Thermoplasmata).

While broad community differences between AL and PL could be observed based on the subsampled segments from both cores, gradual fine-scale changes were only evident through the high resolution analysis of 16S rRNA genes from core 1 (Fig. 1B). As community composition was analysed every 3–4 cm, both gradual changes over depth and abrupt shifts could be observed. Especially the relative abundance of different classes of the phylum Proteobacteria showed fine-scale changes within the AL (Fig. 1B and Supporting Information Fig. S1). The abundance of Alphaproteobacteria and Gammaproteobacteria decreased gradually from 24% to 7% and 11% to 1.5% respectively (Fig. 1B). The decrease in relative abundance in the AL of those two classes correlated with depth (Alphaproteobacteria: $r = -0.85$, $p < 0.002$ and Gammaproteobacteria: $r = -0.89$, $p < 0.001$), while no significant correlations to C, pH or water content could be identified (Table 1). An opposite trend was observed for the class Betaproteobacteria, which was underrepresented down to 36 cm before their relative abundance increased from 4% to 22% down to 75 cm. Other phyla, like the Gemmatimonadetes and candidate phylum AD3 also increased in

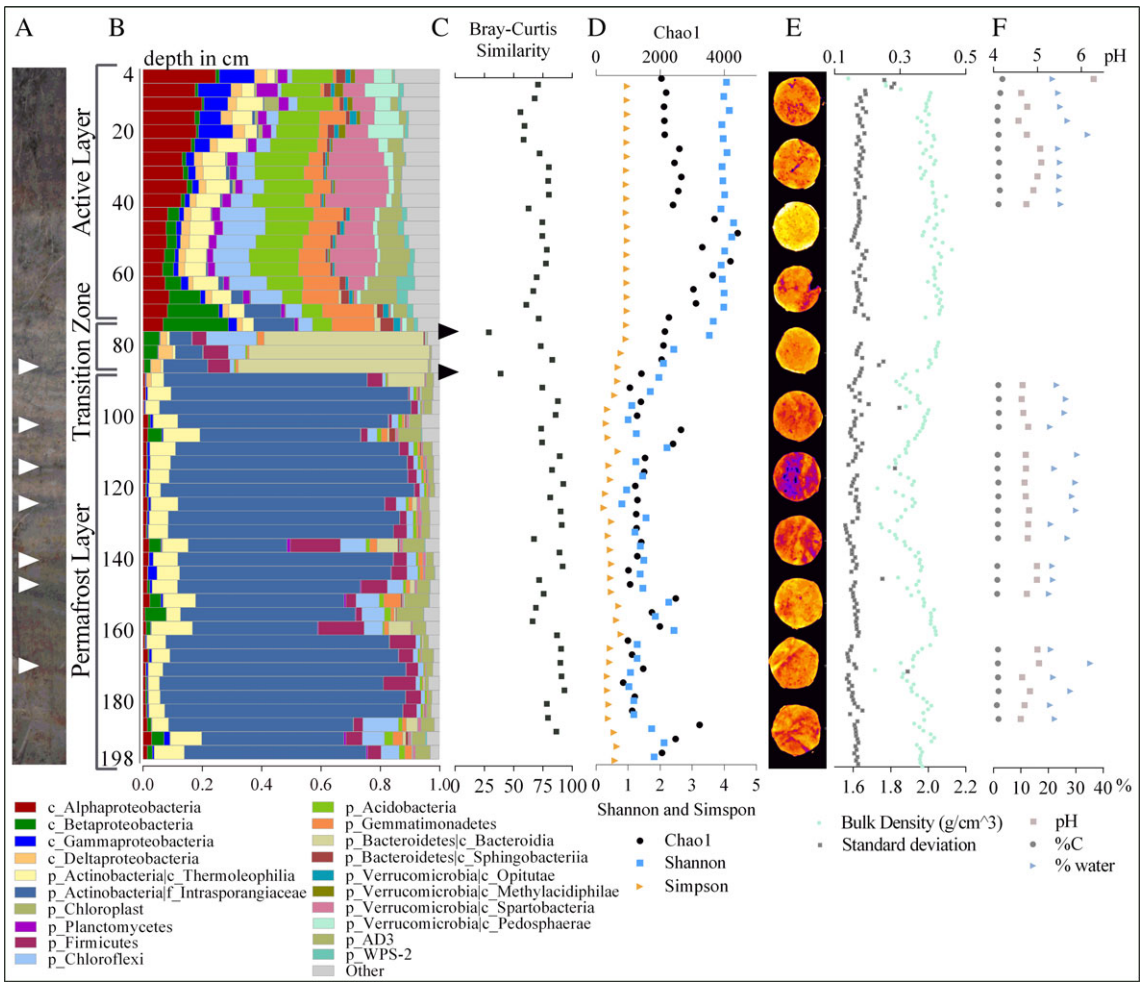

Fig. 1. Permafrost soil profile illustrating:
A. The soil structure in Adventdalen. Photograph shows the sedimentation process with periods of low sedimentation rates (small darker organic rich layers, marked with white triangles) and periods with higher sedimentation rates (thicker more pale layers).
B. Relative abundance of the 20 most abundant taxa at different taxonomical levels based on 16S rRNA gene sequence data showing the taxonomic prokaryotic community composition of the permafrost core profile. When taxonomical classes within a phylum showed very contrasting trends, the most abundant classes were illustrated instead of the entire phyla. Taxonomical levels are indicated by a one-letter code (p = phylum; c = class; f = family). Taxa comprising <1% of the total number of sequences within a sample were summarized as 'Other'. Relative abundances for each sampling point are average values calculated from 2 to 6 replicates (Supporting Information Table S1).
C. Bray–Curtis dissimilarity values calculated for 16S rRNA gene sequence data at OTU level. Black triangles mark the transition from AL to TZ and from TZ to PL.
D. Alpha diversity indices (Chao1, Shannon and Simpson).
E. Bulk density measurement using X-ray CT scanning with measuring points for every centimetre (11 example images illustrated on the left).
F. Soil chemistry (n = 1) showing in blue the water content, in grey the carbon content and in red the pH throughout the core. Note that data was not available between 30 and 80 cm.

abundance from 42 cm down to 75 and 68 cm respectively. A general decrease down to 75 cm could be observed for the phyla Acidobacteria, Verrucomicrobia, Chloroflexi and Planctomycetes. Throughout the entire core, an increase or decrease in relative abundance of certain taxa significantly correlated with depth, while correlations to C concentrations, pH and water content were predominantly not significant (Table 1).

We detected a major shift in microbial community composition at 75–78 cm depth in the AL. This was mainly driven by changes in relative abundance of Bacteroidetes (class Bacteroidia), which increased from 2% to 54% and Proteobacteria, which decreased from 34% to 8%. Bacteroidetes (class Bacteroidia) stayed abundant until 85 cm before decreasing to <0.5% within the next 7 cm of the soil core. We identified mayor differences in Bray–Curtis dissimilarity values, with 29% similarity at the transition from 75 to 78 cm and with 39% similarity between 85 and 88 cm (Fig. 1C). As community similarity between consecutive samples is on average about 78%, we interpreted those two shifts as the transition from AL to TZ and TZ to PL respectively. Consequently, this change at

Table 1. Statistical analysis using the Pearson correlation coefficient to identify linear correlations throughout the entire core and the first 30 cm of active layer by comparing depth, and the soil properties carbon concentration, pH and water content with the relative abundance of the 20 most abundant taxa at class level.

		c_Alpha proteobacteria	c_Beta proteobacteria	c_Gamma proteobacteria	c_Delta proteobacteria	p_Actinobacteria c_Thermoleophilia	p_Actinobacteria f_Intrasporangiaceae	p_Chloroplast	p_Planctomycetes	p_Firmicutes	p_Chloroflexi	p_Acidobacteria	p_Gemmatimonadetes	p_Bacteroidetes c_Bacteroidia	p_Bacteroidetes c_Sphingobacteriia	p_Verrucomicrobia c_Opitutae	p_Verrucomicrobia c_Methylacidiphilae	p_Verrucomicrobia c_Spartobacteria	p_Verrucomicrobia c_pedosphaerae	p_AD3	p_WpS-2
Depth	Entire core																				
	r	-0.82	-0.14	-0.64	-0.84	0.13	0.81	-0.43	-0.79	0.62	-0.21	-0.83	-0.46	-0.04	-0.52	-0.69	-0.69	-0.7	-0.66	-0.04	-0.41
	p (two-tailed)	< 0.0001	0.321	< 0.0001	< 0.0001	0.3554	< 0.0001	0.0017	< 0.0001	< 0.0001	0.141	< 0.0001	0.001	0.797	0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.759	0.0035
	p summary	****	ns	****	****	ns	****	**	****	****	ns	****	***	ns	***	****	****	****	****	ns	**
	First 30 cm																				
	r	-0.85	0.52	-0.89	-0.82	0.57	0.64	-0.42	-0.51	0.4	0.93	0.85	0.57	0.13	-0.84	-0.74	-0.76	0.9	-0.7	0.52	-0.69
	p (two-tailed)	0.002	0.123	0.001	0.004	0.089	0.046	0.224	0.132	0.258	0	0.002	0.089	0.719	0.003	0.015	0.011	0	0.025	0.124	0.028
	p summary	**	ns	***	**	ns	*	ns	ns	ns	***	**	ns	ns	**	*	*	***	*	ns	*
% Carbon	Entire core																				
	r	0.54	0.35	0.54	0.54	-0.41	-0.44	0.09	0.53	-0.4	-0.04	0.44	0.05	-0.02	0.69	0.42	0.51	0.31	0.44	-0.38	0.33
	p (two-tailed)	0.0021	0.0559	0.0023	0.0023	0.0256	0.0143	0.6369	0.0026	0.0288	0.8444	0.015	0.7857	0.9044	< 0.0001	0.0201	0.0042	0.0999	0.0159	0.039	0.0713
	p summary	**	ns	**	**	*	*	ns	**	*	ns	*	ns	ns	****	*	**	ns	*	*	ns
	First 30 cm																				
	r	0.52	0	0.42	0.39	-0.73	-0.47	-0.24	0.36	0.19	-0.37	-0.1	-0.54	-0.39	0.65	0.11	0.3	-0.2	0.17	-0.6	-0.01
	p (two-tailed)	0.126	0.99	0.228	0.262	0.016	0.175	0.51	0.302	0.599	0.295	0.793	0.108	0.265	0.043	0.759	0.403	0.583	0.644	0.07	0.977
	p summary	ns	ns	ns	ns	*	ns	ns	ns	ns	ns	ns	ns	ns	*	ns	ns	ns	ns	ns	ns
pH	Entire core																				
	r	0.44	0.29	0.4	0.34	-0.15	-0.33	0.05	0.22	-0.12	0.1	0.31	0.02	-0.13	0.5	0.27	0.18	0.23	0.13	-0.1	0.05
	p (two-tailed)	0.0148	0.116	0.0296	0.0631	0.4423	0.0786	0.7758	0.2474	0.5246	0.6011	0.0985	0.9279	0.4913	0.0046	0.1428	0.3343	0.2212	0.4874	0.586	0.799
	p summary	*	ns	*	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	**	ns	ns	ns	ns	ns	ns
	First 30 cm																				
	r	0.58	0.07	0.25	0.27	-0.42	-0.37	-0.15	-0.16	-0.22	-0.21	0.01	-0.42	-0.16	0.48	0.05	-0.12	-0.08	-0.18	-0.19	-0.32
	p (two-tailed)	0.079	0.849	0.495	0.452	0.232	0.299	0.682	0.66	0.545	0.556	0.989	0.229	0.652	0.157	0.888	0.75	0.836	0.61	0.606	0.371
	p summary	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
% water	Entire core																				
	r	0.02	-0.07	0.1	0.06	-0.07	-0.01	0.06	-0.01	-0.1	-0.17	0.02	0.04	-0.09	0.1	0.1	0.13	-0.05	0.2	-0.07	0.15
	p (two-tailed)	0.9051	0.7094	0.5964	0.7431	0.7063	0.9723	0.767	0.9672	0.6098	0.3756	0.8979	0.8245	0.6326	0.6015	0.5833	0.5047	0.8011	0.2802	0.715	0.4297
	p summary	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
	First 30 cm																				
	r	-0.06	-0.31	0.3	0.38	-0.23	0.13	0.1	-0.16	-0.18	-0.24	-0.11	0.12	0.02	0.15	0.31	0.35	-0.27	0.57	-0.29	0.38
	p (two-tailed)	0.863	0.384	0.405	0.274	0.515	0.727	0.778	0.661	0.617	0.502	0.753	0.742	0.954	0.688	0.381	0.319	0.453	0.083	0.424	0.285
	p summary	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns

Table 2. ANOSIM analysis of Bray–Curtis dissimilarities for the three different permafrost zones AL (Active layer; n = 19), TZ (Transition zone; n = 3) and PL (Permafrost layer; n = 27). R indicates the grade of dissimilarity (1 = most dissimilar) and p the statistical probability.

Sample grouping	Global *R*	*p*
Active layer – Transition zone	0.997	0.002
Active layer – Permafrost layer	0.993	0.001
Transition zone – Permafrost layer	1	0.001

88 cm, where the microbial community composition shifted towards a dominance of Actinobacteria, can be interpreted as the beginning of the PL and the depths between 78 and 85 cm as the TZ. An ANOSIM analysis on Bray–Curtis dissimilarities between the three layers observed in core 1 showed that they are significantly different to each other and justifies the introduction of the TZ as a biological independent layer in the permafrost core (Table 2). The three layers showed also significant differences in richness (Chao1), evenness (Simpson) and diversity (Shannon), illustrated in Fig. 2 (ANOVA, $p < 0.0001$). All three alpha diversity indices were higher in the AL and decreased significantly with depth down to the PL (Pearson's r: Shannon [$r = -0.8242$] and Simpson [$r = -0.7666$] with $p < 0.0001$ and Chao1 [$r = -0.5121$] with $p = 0.0002$) (Fig. 1D). Beta-diversity analysis confirmed these differences and is illustrated by a multidimensional scaling analysis (MDS) plot (Fig. 3). All samples from the PL shared 60% similarity, while AL samples only shared 20% similarity. Furthermore, the AL, TZ and PL samples grouped separately, with AL samples showing the widest distribution with a clear, gradual clustering from surface to deeper AL samples.

The subsampled segments (AL-1, AL-2, TZ-3, PL-4 and PL-5) from core 1 and core 2 used in the incubation experiments were chosen based on community composition results from the high resolution analysis of core 1. All five segments from core 1 revealed a characteristic community composition and grouped according to the depth they were taken from (Fig. 3 and Supporting Information

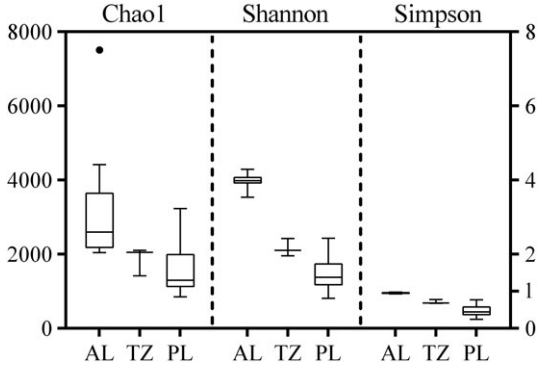

Fig. 2. Richness and alpha diversity indices (Chao1, Shannon and Simpson). Pooled samples from the AL (n = 19), TZ (n = 3) and PL (n = 27) were used to calculate alpha diversity indices.

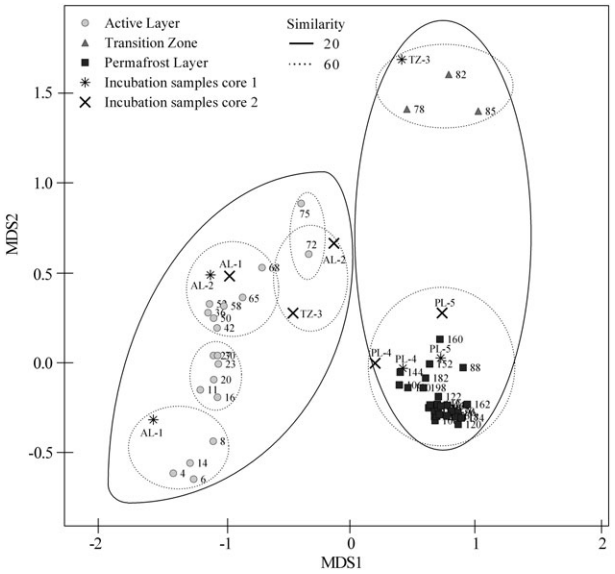

Fig. 3. MDS plot based on Bray–Curtis dissimilarity values illustrating the variation of bacterial community composition based on sequenced 16S rRNA gene fragments at OTU level. The analysis included the high resolution core samples and samples from the five layers of each core used for the incubation experiment. The soil type of each core sample is given by symbols and the sample depth and incubation segment is added as caption. Numbers correspond to sampling depths in core 1. The stress value for the MDS plot is 0.06. The black circles around clusters indicate similarities of 20% and the dotted circles 60% similarity.

Table S3). The five segments from core 2 were taken from similar depths and community composition was comparable to the corresponding layers in core 1 (Figs 3 and 4C). One exception was the absence of a Bacteroidetes dominated TZ in core 2, instead, community composition in the TZ-3 sample from this core was more similar to the AL-2 sample from core 1 (Fig. 3 and Supporting Information Table S3). The characteristic differences between AL and PL segments observed for core 1 were also present in core 2 and included the decrease of Alphaproteobacteria and increase of Betaproteobacteria with increasing AL depth and the dominance of *Intrasporangiaceae* in PL samples (Fig. 4C).

Degradation potential of soil organic matter

The five segments (AL-1, AL-2, TZ-3, PL-4 and PL-5) were chosen based on community composition differences and used as models to simulate thawing conditions in different layers and to document CO_2 fluxes. Subsamples from the five segments from both cores were incubated for up to 122 days and CO_2 was measured regularly at eight time points for the first 19 days (Fig. 4A). Overall CO_2 emissions were higher in core 1 than core 2, but showed similar trends for the different segments (Fig. 4A). In core 1, highest CO_2 values were measured for PL-5 and lowest for AL-2 with 65 μg C-CO_2 g soil^{-1} and 19 μg C-CO_2 g soil^{-1} respectively.

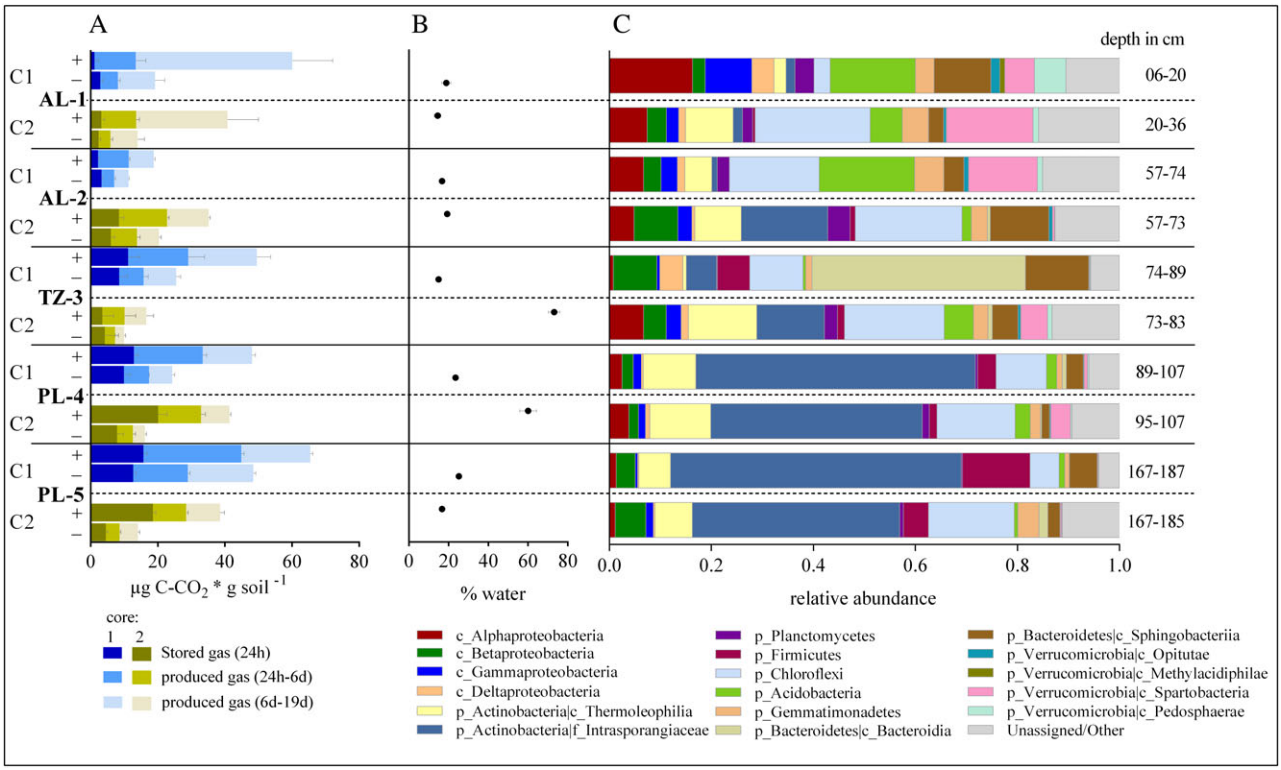

Fig. 4. CO_2 emissions, water content and microbial community structure of five characteristic layers from two replicate permafrost cores. C1: Core 1 from 2011 and C2: Core 2 from 2014.
A. CO_2 emissions of permafrost samples measured over 19 days and categorized in 3 phases which are indicated by the different shades of blue and green as stacked bars. Samples from five segments of the cores were incubated at 4 (±1)°C under aerobic (+) or anaerobic (–) conditions (n = 3).
B. Water content in % of weight for the five different layers (n = 3).
C. Microbial community composition of the five segments from the two permafrost cores used for the incubation experiment covering the AL, TZ and PL (n = 1). Illustrated are the most abundant taxa at different taxonomical levels. Taxonomical levels are indicated by a one-letter code (p = phylum; c = class; f = family). Taxa comprising <1% of the total number of sequences within a sample were summarized as 'Other'.

Differences could be observed when the 19 days were divided into three phases. The first 24 h cover the release of trapped CO_2 (phase 1), the next 5 days mark the first CO_2 production phase (phase 2) and the last 12 days represent the later CO_2 production phase (phase 3). PL-5, PL-4 and TZ-3 had all high amounts of stored CO_2 which were released during the first 24 h. AL-1 and AL-2 released only small amounts during this first phase. The production during phase 2 was highest in PL-5 (29 µg C-CO_2 g soil^{-1}) and decreased towards AL-1 (12 µg C-CO_2 g soil^{-1}). The most contrasting rates between the five segments were observed during production phase 3, when AL-1 showed the highest CO_2 emissions with 47 µg C-CO_2 g soil^{-1} making up for 78% of the total emissions within 19 days.

Overall similar trends, as described above, could be observed for the representative segments from core 2. By comparing both treatments, we found that CO_2 emissions were always higher under aerobic incubations, with up to 4.2 times more CO_2 produced for AL-1 in phase 3 (Fig. 4A). The last sampling point was after 122 days and during this 103 day period, AL-1 samples produced the most CO_2 with 1,207 µg C-CO_2 g soil^{-1} (Supporting Information Fig. S4), while the segments TZ-3, PL-4 and PL-5 released around 800 µg C-CO_2 g soil^{-1} and AL-2 produced only 330 µg C-CO_2 g soil^{-1}.

All layers released significant amounts of CO_2, thereby showing spatial (ANOVA, core 1: p = 0.0369; core 2: p = 0.007) and temporal (ANOVA, core 1: p = 0.0032; core 2: p = 0.0041) differences. These differences were compared with the community composition in the different layers and visualized in a redundancy analysis biplot (Fig. 5). The redundancy analysis (RDA) model included the OTU abundance matrix as response variables and CO_2 emission values for the three categories (24 h; 24 h to 6 days; and 6–19 days) as predictor variables, which explained 46.2% (p = 0.016) of the total variance. Bacterial abundance indicated as 16S rRNA gene copies determined by qPCR (Supporting Information Fig. S5) showed no statistical significant correlation with CO_2 emission values. For core 1, chemical parameters such as DOC, DTN, NH_4-N, pH and water content, before and after 19 days of incubation, were measured. The results showed no major differences between the two different treatments and incubation time (Supporting Information Fig. S3).

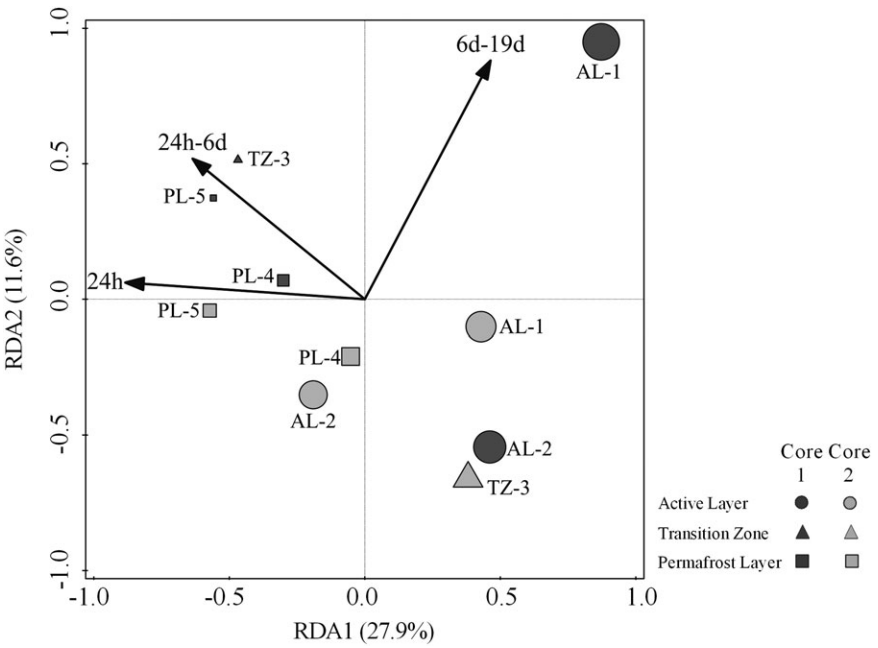

Fig. 5. Redundancy analysis biplot summarizing the variation of bacterial community composition in the five segments from both cores used for the incubation experiment plotted in relation to CO_2 emission under aerobic conditions used as explanatory variables (black arrows). CO_2 emission was grouped into three periods covering the first 24 h, 24 h to 6 days and 6–19 days. The size of the symbols relates to the sample diversity (Shannon-Wiener index) at species level.

Metagenomics analysis

We were interested in identifying functional differences in the different layers of the core as well as differences due to the induced thawing conditions of the incubation experiment (Fig. 6). First, differences between the four analysed segments were greater than differences caused by incubation for up to 16 days (Supporting Information Fig. S4). Changes in gene abundance over the course of the 16 days of incubation were more pronounced under

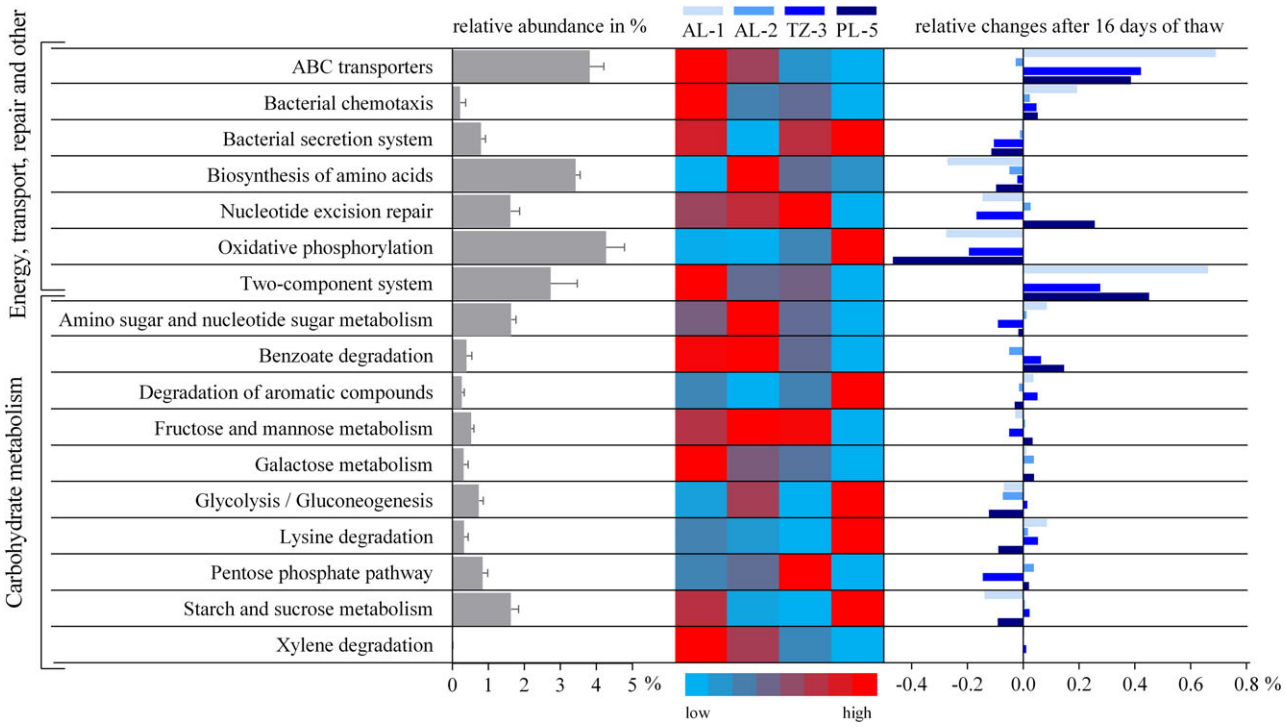

Fig. 6. Heat map displaying the relative abundance of KEGG annotated metagenomics reads assigned to a selection of different metabolic pathways in four different layers of the permafrost core. The relative abundance of reads assigned to each pathway is displayed by the grey bar graph to the left and the changes in relative abundance for each layer after 16 days of incubation is indicated by the bar graph to the right.

aerobic conditions than in anaerobic treatments, where microbial communities remained relatively unchanged (data not shown). Overall, most reads (>60%) were assigned to genes involved in various metabolisms, of which the carbohydrate metabolism was most represented with 10%–15% of all reads, but varying for each layer. Glycolysis, starch and sucrose metabolism and degradation of aromatic compounds were overrepresented in the PL, while the fructose metabolism and the pentose phosphate pathway were pronounced in the TZ (Fig. 6). In the AL, benzoate and xylene degradation and the galactose metabolism were overrepresented. After 16 days of thaw, the relative abundance of these KEGG pathways changed considerably in three of the four layers, only AL-2 samples showed no or minimal changes (Fig. 6). Common for all layers was an increase in reads assigned to DNA repair mechanisms and decarboxylases. The most pronounced change after 16 days of thaw was an increase in relative abundance of genes connected to ABC transporters and the two-component system. Changes in the relative abundance of carbohydrate pathways were considerably less pronounced after 16 days of thaw. Additionally, the layers where a certain carbohydrate pathway was overrepresented compared with the other layers, also had the strongest reduction of the respective pathway after 16 days of thaw, for example, the decrease in relative abundance of the starch and sucrose metabolism in AL-1 and PL-5 (Fig. 6).

Discussion

Fine-scale resolution

Numerous studies have characterized the microbial community structure of permafrost-affected soils, but mostly on broader scales (Yergeau *et al.,* 2010; Wilhelm *et al.,* 2011; Frank-Fahle *et al.,* 2014; Koyama *et al.,* 2014; Taş *et al.,* 2014; Gittel *et al.,* 2014a,b; Deng *et al.,* 2015; Schostag *et al.,* 2015). So far, only two studies characterized shifting microbial communities at a finer scale (Kim *et al.,* 2016; Tripathi *et al.,* 2018). They analysed the AL of Alaskan organic rich tundra soils in steps of 5 cm and showed that a high resolution analysis of the soil profile can reveal unique changes in relative abundance of certain bacterial groups with depth.

Here, a permafrost core from Svalbard was analysed at 3–4 cm intervals to a depth of 2 m. Due to the fine resolution, we were able to identify microbial community shifts in the AL and the PL, and further to distinguish a structurally unique and different TZ between these layers. Throughout the entire core, bacteria were dominating over archaea, which could be detected at various depths, but did not exceed abundances of more than 0.5% of the total community. Low archaeal abundances have been reported before and especially the absence of methanogenic archaea has been associated with low water content, similar to the soil investigated in this study (Høj *et al.,* 2006; Yergeau *et al.,* 2010; Rivkina *et al.,* 2016). We identified similar trends as seen in Kim *et al*. (2016) for the bacterial community in the AL, where the relative abundance of Alphaproteobacteria, Gammaproteobacteria and Acidobacteria decreased, whereas Betaproteobacteria, Gemmatimonadetes and candidate phylum AD3 increased with soil depth. The community shifts caused by the different abundances of Proteobacteria and the other phyla have to some extent been described before (Koyama *et al.,* 2014; Deng *et al.,* 2015; Kim *et al.,* 2016), and correlated with different soil horizons based on C turnover. Although those studies were based on organic soils, we could identify similar community shifts in our mineral soil core with <5% C content. The higher abundance of Alphaproteobacteria and Gammaproteobacteria in the upper AL has been speculated to correlate with higher C and nutrient availability (Koyama *et al.,* 2014; Kim *et al.,* 2016). However, Alphaproteobacteria and Gammaproteobacteria were also more abundant in the upper AL of our core, despite the overall low C content in this mineral soil. This indicates that C availability alone might not be suitable to explain changes in the relative abundance of Alphaproteobacteria and Gammaproteobacteria in Arctic soils with both low and high carbon concentrations. The current study can, however, not discriminate whether the observed decline in abundance is related to carbon related factors like C quality or a combination of several factors, including oxygen availability and redox effects.

Betaproteobacteria and especially taxa belonging to the uncultured and uncharacterized order SBla14 and the order *Gallionellales* increased in abundance in the deeper AL. As *Gallionellales* are chemolithotrophic iron-oxidizers, their increase in relative abundance might be connected to an increase in available Fe(II), which they use as primary energy source (Emerson *et al.,* 2015). In mineral soils, Fe(II) is less likely to be bound to organic ligands and therefore available for oxidation by iron-oxidizing bacteria (Liang *et al.,* 1993). Other taxa that increased in relative abundance in the deeper AL belong to the phylum Gemmatimonadetes and the candidate phylum AD3 and have been detected in deeper soils before (Costello, 2007; Taş *et al.,* 2014; Deng *et al.,* 2015). It is unclear which environmental factors are causing their increase, as members of these phyla are largely uncharacterized. While dry soils and neutral pH have been associated with a higher abundance of Gemmatimonadetes (DeBruyn *et al.,* 2011), higher abundances of candidate phylum AD3 were associated with low carbon concentrations (Jansson and Taş, 2014). Interestingly, all phyla, independent of increasing or decreasing

abundance towards the deeper AL, abruptly changed from 75 to 78 cm, due to the tremendous increase in relative abundance of Bacteroidetes (from 2% to 54%; Fig. 1). The Bacteroidetes-dominated zone likely coincides with a geocryological defined transition zone, which, according to the three-layer conceptual model, describes a zone between the seasonally thawed AL and the stable PL, with subdecadal freeze–thaw transitions (Shur *et al.,* 2005). The thickness of the AL is variable in Adventdalen and can range between 63 and 90 cm within a 20 m radius of our study site (field measurements in 2016 and Cable *et al.,* 2018). These AL maximum depths however suggest that a TZ above the permafrost table in core 1 might have been located where we also identified the high Bacteroidetes abundance. It might also explain why we did not identify this zone of high Bacteroidetes abundance at similar depths in core 2. As the community composition of the TZ-3 segment from core 2 was most similar to the AL-2 samples from core 1, the AL might have extended deeper in core 2 than in core 1 and hence might explain why we did not identify the actual TZ with high Bacteroidetes abundance in core 2. The current study can, however, not provide evidence whether the high Bacteroidetes abundance is associated to the effects of recent thawing in the transition zone. High Bacteroidetes abundances in general have been associated with their metabolic flexibility and their ability to quickly respond to the easy available C and nutrients (Padmanabhan *et al.,* 2003; Fierer *et al.,* 2007). Chemical analysis of the TZ-3 samples from core 1 however, showed neither elevated carbon nor nitrogen levels compared with the neighbouring layers AL-2 and PL-4. It remains uncertain what factors caused this clear separation of the TZ or whether this layer might be a record of a particular condition in the past when it was formed and buried due to sedimentation processes. Interestingly, the dominance by one taxonomical group (unknown family; within the class Bacteroidia) in the TZ is followed by the dominance of another family (*Intrasporangiaceae*; within the phylum Actinobacteria) in the PL. The sharp community changes between the different layers, which we here demonstrate for the first time, are an indicator that community structures are highly flexible and might adapt quickly to distinct conditions.

Studies of permafrost from other Arctic regions have reported a high variability of bacteria belonging to the phyla of Actinobacteria, Proteobacteria, Verrucomicrobia, Chloroflexi, Bacteroidetes and Firmicutes (Hansen *et al.,* 2007; Yergeau *et al.,* 2010; Mackelprang *et al.,* 2011; Yang *et al.,* 2012; Jansson and Taş, 2014). In our study, we observed a dominance of only four OTUs, belonging to the family of *Intrasporangiaceae*, which accounted for up to 80% of the entire community in the PL in core 1 (Fig. 1B) and up to 42% in core 2 (Fig. 4C). High dominance of this Actinobacteria family has been seen in permafrost soils before [(Gittel *et al.,* 2014b): Greenland up to 8%; (Gittel *et al.,* 2014a): Siberian Arctic up to 47%], but to our knowledge never been reported in such consistent high numbers throughout the entire PL. It has been discussed that these Actinobacteria are globally successful in permafrost environments due to their adaptions to low C availability and capabilities to degrade complex C compounds, like cellulose, cellobiose and lignin (DeAngelis *et al.,* 2011; Giongo *et al.,* 2013; Gittel *et al.,* 2014b).

Greenhouse gas fluxes and functional genes

Although respiration takes place at sub-zero temperature, several studies have shown a rapid increase in respiration upon thaw, probably due to the availability of liquid water (Clein and Schimel, 1995; Larsen *et al.,* 2002; Elberling and Brandt, 2003; Mackelprang *et al.,* 2011). Nikrad *et al*. (2016) therefore suggested that several parameters, including temperature, C content and changes in the physical environment should be included when comparing respiration rates. In our study, we included the microbial community composition as an additional factor regarding respiration rates.

We were able to show, that both incubation conditions and soil depth with associated differences in microbial community structure influence CO_2 fluxes during permafrost thaw (Fig. 4A). Incubations with a different community structure showed different respiration rates. The highest difference was observed between AL-1 and AL-2 incubations from core 1. Within 19 days, AL-1 samples released 60 μg C-CO_2 per gram soil, while samples from AL-2 released three times less (Fig. 4A). Similar low CO_2 production rates were observed for samples from the TZ-3 of core 2 (Fig. 4A). As the community composition of TZ-3 from core 2 was highly similar to the AL-2 segment from core 1 (Fig. 3 and Supporting Information Table S3), suggests that depth-dependent differences in community composition might be one indicator for the CO_2 production potential (Fig. 5). If these findings can be applied to longer term and field conditions requires further investigations, as well as the identification of the active drivers of CO_2 production and their abundance in the different communities. Further, the correlation between community composition and CO_2 fluxes that we observed, might also be driven by substrate specific factors which we did not measure. Nevertheless, our results highlight the complexity of microbial driven respiration throughout a permafrost core and the necessity for further detailed gas flux studies considering the different responses in the different layers.

Independent of soil layer and community structure, respiration rates were higher under aerobic than anaerobic

conditions, with up to four times more produced CO_2 (Fig. 4A). Similar results have been documented in a study with permafrost from Greenland, incubated for 12 years (Elberling *et al.,* 2013) and a comparative study investigating aerobic and anaerobic permafrost incubations from different locations in Alaska and Siberia (Lee *et al.,* 2012).

Under anaerobic conditions, the release of CO_2 was with 50 μg C-CO_2 per gram soil twice as high in PL-5 as in PL-4 and TZ-3 and almost three times as high as in AL-1 and AL-2. This indicates that the fermentative potential is highest in the deepest PL layer, which also represents the only layer which has not been thawed at any point since the beginning of the Holocene. Even though TZ-3 has been thawed during the last 15 years, similar amounts of stored CO_2 were released during the first 24 h compared with PL-5. This indicates that CO_2 is produced under freezing conditions and that CO_2 can be stored until it is released upon thawing (as previous documented by Elberling and Brandt, 2003).

In contrast to other permafrost incubations, we could not detect CH_4 production in any of our treatments. The absence of methane production has been shown to be connected to iron and sulphate reduction processes, which energetically outcompete methanogenesis under anaerobic conditions (Patrick and Jugsujinda, 1992; Megonigal *et al.,* 2004). The metagenomics analysis confirmed that methanogenesis might not have been activated, as genes involved in methane production were at a very low level or absent throughout the entire soil core, while genes encoding iron transporter proteins and sulfoxide reductases, were widely expressed and abundant throughout the core.

This analysis also revealed that of all metabolisms, most genes were assigned to several carbohydrate pathways, varying for the different layers. The microbial community in the AL was involved in various carbohydrate degradation pathways, whereas communities in deeper layers depended on smaller organic substrates for energy conservation and degradation of complex aromatic compounds. Especially, transport and signalling genes were overrepresented in the AL and increased upon incubation. The abundance of certain genes changed in nearly all soil layers over the course of the incubation, with the exception of the lower part of the AL (AL-2) (Supporting Information Fig. S4). This was also the layer that, despite the genomic potential to degrade various carbon sources, produced the least amount of CO_2 during the incubation. It remains unclear why CO_2 production rates differed so substantially within the AL. These spatial and temporal differences in CO_2 production within the different layers indicate important implications for future climate change models which incorporate permafrost greenhouse gas release.

Conclusion

In this study, high resolution profiling of the microbial community throughout an entire 2-m permafrost core from Svalbard allowed us to closely follow changes. It showed significant differences between the AL and the PL, with the latter being dominated by *Intrasporangiaceae* and revealed a ~8 cm spanning TZ with a clearly different community, dominated by Bacteroidetes. Further, we showed that CO_2 release from our permafrost samples varied within the different layers and during different incubation periods. Together, our results indicate that the spatial and temporal variability of CO_2 is reflected in the fine-scale variations of microbial communities throughout the permafrost core, that can change dramatically even within 3 cm. Due to observed structural variability of permafrost soils we cannot expect our identified microbial and activity patterns to be universal. However, the results from our study strengthen the necessity of detailed high resolution microbial community permafrost profiles to understand how the different microbes are distributed, how they interact and how they function in this globally important and changing Arctic environment.

Experimental procedures

Study site and sampling

Soil samples were obtained from a characteristic low-centred ice-wedge polygon site in the valley Adventdalen on Svalbard (78.186 N, 15.9248E). Overall, two separate cores were drilled; the first core was obtained in April 2011 and is referred to as 'core 1' and a replicate core was drilled at the same coordinates in July 2014 ('core 2'). The high resolution sequencing and chemical analyses are based on soil samples from core 1. Subsamples (n = 2–6) from each of the analysed 50 sections of the core were used as replicates for the 16S rRNA gene sequencing (Supporting Information Table S1). The incubation experiments included samples from five representative segments from core 1 and at similar depths five segments from core 2 (Supporting Information Table S2). Samples from these five layers used in the incubation experiments were also used for 16S rRNA gene sequencing. The incubation experiments and chemical analyses included three to four replicates.

Soil processing and chemical properties

To remove potential core surface contaminants, introduced during the drilling procedures, the outermost 2 cm were scraped off with sterile blades. The remaining inner part was kept frozen on dry ice and transferred to a sterile thick plastic bag and homogenized by hammering. Characteristic soil properties, such as pH, water content,

dissolved organic carbon, NO_3 and NH^{4+} were measured for most of the subsections of core 1 used for the high resolution profile and for all segments from both cores used for the incubation experiments.

X-ray CT scanning

The core was X-ray CT scanned using a modified Siemens Somatom HiQ medical CT scanner at 133 kV. To obtain bulk density, the CT scanner was calibrated by scanning a number of materials having known density. The data were analysed using imageJ (Schindelin *et al.,* 2015).

Soil incubations and gas-flux measurements

All preparations for soil incubations were performed in a –20°C cold room. All incubation experiments included three replicates and covered five different segments (2 from AL, 1 from TZ and 2 from PL) from each of the two permafrost cores. Two gram of soil were added to sterilized glass vials (20 ml) and sealed with butyl rubber stoppers and aluminium crimps. Samples incubated under anaerobic conditions were flushed with nitrogen for 1 min. Three millilitres of headspace gas were collected from the vials immediately after the incubation was started and at 9 time points (4 h, 12 h, 24 h, 2 days, 4 days, 6 days, 16 days, 19 days and 122 days) during incubation. After each gas collection, aerobic or anaerobic conditions were re-established. The gas samples were transferred into 3 ml Exetainer glass vials and analysed for CO_2, CH_4 and N_2O on a SRI 8610C gas chromatograph. Measurements for CH_4 and N_2O were below the detection limit in all the segments and data was therefore not included. All samples were incubated at 4($\pm$1)°C in the dark.

DNA extraction, 16S rRNA gene amplification and amplicon sequencing

Frozen soil from homogenized segments of the core was used for DNA extractions, following manufacturer's instructions of the PowerSoil® DNA Isolation Kit (Mobio, Carlsbad, USA) and performed in triplicates with minor modifications. Instead of 0.25 g, 0.3–0.4 g soil was used for the DNA extraction and an additional incubation step of 70°C for 5 min was included after solution C1 was added. Information regarding the subsample depth, the number of replicates, DNA concentration and the number sequences before and after processing is listed in Supporting Information Table S1. DNA amplification included a two-step nested PCR approach with primers 519F [CAGCMGCCGCGGTAA; (Øvreås *et al.,* 1997)] and 806R [GGACTACHVGGGTWTCTAAT; (Caporaso *et al.,* 2011)] targeting the bacterial and archaeal 16S rRNA gene V4 hypervariable region. Details can be found in (Wilson *et al.,* 2017). Libraries were send to the Yale Center for Genome analysis (Yale University, the W.M. Keck Biotechnology Resource Laboratory, CT, USA) and the Norwegian Sequencing Centre (Oslo, Norway) for high-throughput sequencing on a MiSeq platform (Illumina). Sequencing data is available at 'The European Bioinformatics Institute' under study accession number PRJEB21759 (http://www.ebi.ac.uk).

Bioinformatic sequence analysis

Sequencing data was processed using different bioinformatic tools incorporated in the Qiime-processing platform using version 1.9.1. (Caporaso *et al.,* 2010). A total of 13,669,151 sequences were retrieved from 151 samples. Prokaryotic OTUs were selected at a sequence similarity threshold of 97% using a de novo uclust (Edgar, 2010) OTU clustering method with default parameters and taxonomy assigned, using the Greengenes reference database (DeSantis, 2006). OTUs with a taxonomic identification were assembled to an OTU table providing abundances for each sample excluding singletons and unassigned OTUs. After removal of singletons and unassigned OTUs, a total of 37,016 unique OTUs at 97% sequence identity were retrieved. Alpha diversity, including Shannon and Simpson indices (Shannon, 1948; Simpson, 1949) and Chao1 richness (Chao, 1984), was calculated using a rarefied sample set, standardized to the smallest read number of 3,500 sequences per sample. To test for multivariate environmental correlation with the prokaryotic community structure, the programs primer-e version 6 (Plymouth, UK) and Canoco 5 (ter Braak and Šmilauer, 2012) were used. Among others, Bray-Curtis resemblance, ANOSIM and principal component analyses (PCA) were calculated using these programs. MDS plot was used to illustrate the variation of bacterial diversity based on Bray–Curtis dissimilarity values of sequenced 16S rRNA gene data. For the RDA, the OTU abundance matrix was log transformed and used as response variables, while CO_2 emission values for the three categories (24 h; 24 h to 6 days; and 6–19 days) were used as predictor variables.

Metagenomics analysis

Samples for metagenomics Illumina High-Seq sequencing were collected from four of the five different segments (2 from AL, 1 from TZ and 1 from PL) from core 1 used in the incubation experiment. Triplicate soil samples of 0.3–0.4 g were collected before the incubation started and after 16 days of thaw under aerobic conditions, frozen immediately in liquid nitrogen and stored in at –80°C.

DNA was extracted using the PowerSoil® DNA Isolation Kit (Mobio, Carlsbad, USA) as described above. DNA was sent to WSU Genomics Core (Washington, Spokane) where DNA was quality checked and libraries prepared using 100 ng DNA per sample and the Illumina® TruSeq® Nano DNA Library Prep kit (Illumina), according to manufacturer`s instructions. Metagenomics shot sequencing data was analysed using the ATLAS software package and standard settings (White *et al.,* 2017). Raw paired-end Illumina reads (.fastq format) were extended for overlaps by using FLASH (Magoč and Salzberg, 2011), after which φX174 was removed using Bowtie2 (Langmead and Salzberg, 2012), the reads were trimmed with trimmomatic (Bolger *et al.,* 2014) and then quality control was performed with FastQC (Andrews, 2010). Overlapped paired-end reads (from FLASH) and unpaired reads were assembled using MEGAHIT (Li *et al.,* 2015). The resulting contigs were subsampled for lengths >1 kbp and translated to protein coding open reading frames (ORFs) using Prodigal (Hyatt *et al.,* 2010) in metagenome mode and annotated using DIAMOND (Buchfink *et al.,* 2015) blastp for protein–protein searching. DIAMOND blastp high-scoring pairs were filtered to user specified bitscore and *e*-value cut-offs (defaults >200 and $< 1 \times 10^7$ respectively). Functional annotation utilizes non-redundant RefSeq (O'Leary *et al.,* 2016) and obtains KEGG (Ogata *et al.,* 1999) (i.e., KO number) annotations from EggNOG reference database. KEGG reads were normalized to metagenome size and relative abundance differences between layers were normalized for each pathway. ATLAS uses RefSeq high-scoring pairs along with NCBI's taxonomy assignments reference tree via a modified majority voting-method that utilizes lowest common ancestor (Hanson *et al.,* 2016), to determine the lowest common ancestor represented across all ORFs present within a single contig. Functional and taxonomic count data was obtained by mapping quality controlled reads to assembled contig annotations using Bowtie2, then parsed using featureCounts of the Subread package (Liao *et al.,* 2014). Sequencing details can be found in Supporting Information Table S4. Metagenomics sequencing data is deposited in the NCBI Sequence Read Archive under the bioproject accession number PRJNA476617.

Statistical analysis

ANOSIM analysis comparing Bray–Curtis dissimilarities between samples for the three different permafrost zones was performed using the program primer-e version 6 (Plymouth, UK). Calculations for ANOVA and the Pearson correlation coefficient (Pearson`s *r*) were performed using GraphPad Prism v 6.01 for Windows (GraphPad Software, CA, USA). Differences in CO_2 emission in the different layers and for the three time periods (24 h; 24 h to 6 days; 6–9 days) during the incubation experiment were analysed by one-way ANOVA with $p < 0.05$ as threshold for statistical significance. To compare the differences of the three alpha diversity indices for the different layers, one-way ANOVA was performed and Pearson`s *r* to indicate whether those indices decreased significantly with depth. Further, Pearson`s *r* linear correlations were carried out to investigate the significance of correlation between the 16S rRNA relative abundance data and depth and the soil properties C concentration, pH and water content.

Acknowledgements

We would like to thank Graham Gilbert for obtaining the soil core from Adventdalen, Svalbard and Hanne Christiansen for relevant site information and UNIS for logistical support. We are grateful to be able to include data from the Circumpolar Active Layer Monitoring (CALM) program. We further like to thank Shi Wang and Yanhong Wang (Lawrence Berkeley National Laboratory), for laboratory assistance processing the soil core. We also like to thank CENPERM for excellent support regarding gas flux experiments and chemical analyses. Thanks also to Bryan Wilson for discussion and help regarding bioinformatics analyses. We further would like to thank Aviaja Lyberth Hauptmann, Jacob Bælum and Rachel Mackelprang for helpful discussions and ideas. This study is part of the project 'Microorganisms in the Arctic: Major drivers of biogeochemical cycles and climate change' (RCN 227062) funded by the Norwegian Research Council. Lise Øvreås was awarded Fulbright Arctic chair 2012/2013 funded by the Fulbright foundation and Bo Elberling and Toke Bang-Andreasen were funded by the Danish National Research Foundation (CENPERM DNRF100). This work was supported in part by the Office of Biological and Environmental Research, Office of Science, US Department of Energy (DOE) grants awarded to Lawrence Berkeley National Laboratory under contract number DE-AC02-05CH11231. The authors acknowledge additional financial support from the Microbiomes in Transition (MinT) Initiative at Pacific Northwest National Laboratory, under contract number DE-AC05-76LO1803 and the DOE Next Generation Ecosystem Experiment-Arctic (NGEE-Arctic) project.

Conflict of Interest

The authors declare no conflict of interest.

References

Åkerman, H. J., and Johansson, M. (2008) Thawing permafrost and thicker active layers in sub-arctic Sweden. *Permafr Periglac Process* **19**: 279–292.

Andrews, S. (2010) FastQC: A quality control tool for high throughput sequence data. [WWW document]. URL http://www.bioinformatics.babraham.ac.uk/projects/Fastqc/

Bang-Andreasen, T., Schostag, M., Priemé, A., Elberling, B., Jacobsen, C. S., and Douglas, T. (2017) Potential microbial contamination during sampling of permafrost soil assessed by tracers. *Sci Rep* **7**: 43338.

Bolger, A. M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Buchfink, B., Xie, C., and Huson, D. H. (2015) Fast and sensitive protein alignment using DIAMOND. *Nat Methods* **12**: 59–60.

Cable, S., Elberling, B., and Kroon, A. (2018) Holocene permafrost history and cryostratigraphy in the high-Arctic Adventdalen Valley, Central Svalbard. *Boreas* **47**: 423–442.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., *et al*. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., *et al*. (2011) NIH public access **7**: 335–336.

Chao, A. (1984) Nonparametric estimation of the number of classes in a population. *Source Scand J Stat* **11177125**: 265–270.

Clein, J. S., and Schimel, J. P. (1995) Microbial activity of tundra and taiga soils at sub-zero temperatures. *Soil Biol Biochem* **27**: 1231–1234.

Costello, E.K. (2007) Molecular phylogenetic characterization of high altitude soil microbial communities and novel, uncultivated bacterial lineages.

DeAngelis, K. M., Allgaier, M., Chavarria, Y., Fortney, J. L., Hugenholtz, P., Simmons, B., *et al*. (2011) Characterization of trapped lignin-degrading microbes in tropical Forest soil. *PLoS One* **6**: e19306.

DeBruyn, J. M., Nixon, L. T., Fawaz, M. N., Johnson, A. M., and Radosevich, M. (2011) Global biogeography and quantitative seasonal dynamics of Gemmatimonadetes in soil. *Appl Environ Microbiol* **77**: 6295–6300.

Deng, J., Gu, Y., Zhang, J., Xue, K., Qin, Y., Yuan, M., *et al*. (2015) Shifts of tundra bacterial and archaeal communities along a permafrost thaw gradient in Alaska. *Mol Ecol* **24**: 222–234.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., *et al*. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72**: 5069–5072.

Edgar, R. C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

Elberling, B., and Brandt, K. K. (2003) Uncoupling of microbial CO2 production and release in frozen soil and its implications for field studies of arctic C cycling. *Soil Biol Biochem* **35**: 263–272.

Elberling, B., Jakobsen, B. H., Berg, P., Søndergaard, J., and Sigsgaard, C. (2004) Influence of vegetation, temperature, and water content on soil carbon distribution and mineralization in four high Arctic soils. *Arctic, Antarct Alp Res* **36**: 528–538.

Elberling, B., Michelsen, A., Schädel, C., Schuur, E. A. G., Christiansen, H. H., Berg, L., *et al*. (2013) Long-term CO2 production following permafrost thaw. *Nat Clim Chang* **3**: 890–894.

Elkateb, T., Chalaturnyk, R., and Robertson, P. K. (2003) An overview of soil heterogeneity: quantification and implications on geotechnical field problems. *Can Geotech J* **40**: 1–15.

Emerson, D., Scott, J. J., Benes, J., and Bowden, W. B. (2015) Microbial iron oxidation in the arctic tundra and its implications for biogeochemical cycling. *Appl Environ Microbiol* **81**: 8066–8075.

Fierer, N., Bradford, M. A., and Jackson, R. B. (2007) Toward an ecological classification of soil bacteria. *Ecology* **88**: 1354–1364.

Frank-Fahle, B. A., Yergeau, É., Greer, C. W., Lantuit, H., Wagner, D., Kuhry, P., *et al*. (2014) Microbial functional potential and community composition in permafrost-affected soils of the NW Canadian Arctic. *PLoS One* **9**: e84761.

Giongo, A., Favet, J., Lapanje, A., Gano, K. A., Kennedy, S., Davis-Richardson, A. G., *et al*. (2013) Microbial hitchhikers on intercontinental dust: high-throughput sequencing to catalogue microbes in small sand samples. *Aerobiologia (Bologna)* **29**: 71–84.

Gittel, A., Bárta, J., Kohoutová, I., Mikutta, R., Owens, S., Gilbert, J., *et al*. (2014a) Distinct microbial communities associated with buried soils in the Siberian tundra. *ISME J* **8**: 841–853.

Gittel, A., Bárta, J., Kohoutová, I., Schnecker, J., Wild, B., Capek, P., *et al*. (2014b) Site- and horizon-specific patterns of microbial community structure and enzyme activities in permafrost-affected soils of Greenland. *Front Microbiol* **5**: 541.

Grosse, G., Romanovsky, V., Jorgenson, T., Anthony, K. W., Brown, J., and Overduin, P. P. (2011) Vulnerability and feedbacks of permafrost to climate change. *Eos, Trans Am Geophys Union* **92**: 73.

Hansen, A. A., Herbert, R. A., Mikkelsen, K., Jensen, L. L., Kristoffersen, T., Tiedje, J. M., *et al*. (2007) Viability, diversity and composition of the bacterial community in a high Arctic permafrost soil from Spitsbergen, Northern Norway. *Environ Microbiol* **9**: 2870–2884.

Hanson, N. W., Konwar, K. M., and Hallam, S. J. (2016) LCA: an entropy-based measure for taxonomic assignment within assembled metagenomes. *Bioinformatics* **32**: 3535–3542.

Hayes, D. J., Kicklighter, D. W., McGuire, A. D., Chen, M., Zhuang, Q., Yuan, F., *et al*. (2014) The impacts of recent permafrost thaw on land–atmosphere greenhouse gas exchange. *Environ Res Lett* **9**: 45005.

Høj, L., Rusten, M., Haugen, L. E., Olsen, R. A., and Torsvik, V. L. (2006) Effects of water regime on archaeal community composition in Arctic soils. *Environ Microbiol* **8**: 984–996.

Hugelius, G., Strauss, J., Zubrzycki, S., Harden, J. W., Schuur, E. A. G., Ping, C.-L., *et al*. (2014) Estimated stocks of circumpolar permafrost carbon with quantified uncertainty ranges and identified data gaps. *Biogeosciences* **11**: 6573–6593.

Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**: 119.

Jansson, J.K. and Taş, N. (2014) The microbial ecology of permafrost.

Jorgenson, M. T., Racine, C. H., Walters, J. C., and Osterkamp, T. E. (2001) Permafrost degradation and

Ecological changes associated with a WarmingClimate in Central Alaska. *Clim Change* **48**: 551–579.

Kim, H. M., Lee, M. J., Jung, J. Y., Hwang, C. Y., Kim, M., Ro, H.-M., *et al*. (2016) Vertical distribution of bacterial community is associated with the degree of soil organic matter decomposition in the active layer of moist acidic tundra. *J Microbiol* **54**: 713–723.

Koyama, A., Wallenstein, M. D., Simpson, R. T., and Moore, J. C. (2014) Soil bacterial community composition altered by increased nutrient availability in Arctic tundra soils. *Front Microbiol* **5**: 516.

Langmead, B., and Salzberg, S. L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.

Larsen, K. S., Jonasson, S., and Michelsen, A. (2002) Repeated freeze–thaw cycles and their effects on biological processes in two arctic ecosystem types. *Appl Soil Ecol* **21**: 187–195.

Lee, H., Schuur, E. A. G., Inglett, K. S., Lavoie, M., and Chanton, J. P. (2012) The rate of permafrost carbon release under aerobic and anaerobic conditions and its potential effects on climate. *Glob Chang Biol* **18**: 515–527.

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

Liang, L., McNabb, J. A., Paulk, J. M., Gu, B., and McCarthy, J. F. (1993) Kinetics of iron(II) oxygenation at low partial pressure of oxygen in the presence of natural organic matter. *Environ Sci Technol* **27**: 1864–1870.

Liao, Y., Smyth, G. K., and Shi, W. (2014) FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930.

Mackelprang, R., Waldrop, M. P., DeAngelis, K. M., David, M. M., Chavarria, K. L., Blazewicz, S. J., *et al*. (2011) Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. *Nature* **480**: 368–371.

Magoč, T., and Salzberg, S. L. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**: 2957–2963.

McCalley, C. K., Woodcroft, B. J., Hodgkins, S. B., Wehr, R. A., Kim, E.-H., Mondav, R., *et al*. (2014) Methane dynamics regulated by microbial community response to permafrost thaw. *Nature* **514**: 478–481.

Megonigal, J.P., Hines, M.E., and Visscher, P.T. (2004) Anaerobic metabolism: linkages to trace gases and aerobic processes.

Nikrad, M. P., Kerkhof, L. J., and Aggblom, M. M. (2016) The subzero microbiome: microbial activity in frozen and thawing soils. *FEMS Microbiol Ecol* **92**: 1–16.

Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa, M. (1999) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **27**: 29–34.

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., *et al*. (2016) Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* **44**: D733–D745.

Øvreås, L., Forney, L., and Daae, F. L. (1997) Distribution of Bacterioplankton in meromictic lake Sælenvannet, as determined by denaturing gradient gel electrophoresis of PCR-amplified gene fragments coding for 16S rRNA. *Appl Environmantal Microbiol* **63**: 3367–3373.

Padmanabhan, P., Padmanabhan, S., DeRito, C., Gray, A., Gannon, D., Snape, J. R., *et al*. (2003) Respiration of 13C-labeled substrates added to soil in the field and subsequent 16S rRNA gene analysis of 13C-labeled soil DNA. *Appl Environ Microbiol* **69**: 1614–1622.

Patrick, W. H., and Jugsujinda, A. (1992) Sequential reduction and oxidation of inorganic nitrogen, manganese, and iron in flooded soil. *Soil Sci Soc Am J* **56**: 1071.

Rivkina, E., Petrovskaya, L., Vishnivetskaya, T., Krivushin, K., Shmakova, L., Tutukina, M., *et al*. (2016) Metagenomic analyses of the late Pleistocene permafrost – additional tools for reconstruction of environmental conditions. *Biogeosciences* **13**: 2207–2219.

Romanovsky, V. E., Smith, S. L., and Christiansen, H. H. (2010) Permafrost thermal state in the polar Northern Hemisphere during the international polar year 2007-2009: a synthesis. *Permafr Periglac Process* **21**: 106–116.

Schindelin, J., Rueden, C. T., Hiner, M. C., and Eliceiri, K. W. (2015) The ImageJ ecosystem: an open platform for biomedical image analysis. *Mol Reprod Dev* **82**: 518–529.

Schostag, M., Stibal, M., Jacobsen, C. S., Bælum, J., Taş, N., Elberling, B., *et al*. (2015) Distinct summer and winter bacterial communities in the active layer of Svalbard permafrost revealed by DNA- and RNA-based analyses. *Front Microbiol* **6**: 399.

Schuur, E. A. G., Vogel, J. G., Crummer, K. G., Lee, H., Sickman, J. O., and Osterkamp, T. E. (2009) The effect of permafrost thaw on old carbon release and net carbon exchange from tundra. *Nature* **459**: 556–559.

Schwalm, C. R., Williams, C. A., Schaefer, K., Anderson, R., Arain, M. A., Baker, I., *et al*. (2010) A model-data intercomparison of CO_2 exchange across North America: results from the North American carbon program site synthesis. *J Geophys Res* **115**: G00H05.

Shannon, C. E. (1948) A mathematical theory of communication. *Bell Syst Tech J* **27**: 379–423.

Shur, Y., Hinkel, K. M., and Nelson, F. E. (2005) The transient layer: implications for geocryology and climate-change science. *Permafr Periglac Process* **16**: 5–17.

Simpson, E. H. (1949) Measurement of diversity. *Nature* **163**: 688.

Taş, N., Prestat, E., McFarland, J. W., Wickland, K. P., Knight, R., Berhe, A. A., *et al*. (2014) Impact of fire on active layer and permafrost microbial communities and metagenomes in an upland Alaskan boreal forest. *ISME J* **8**: 1904–1919.

ter Braak, C. and Šmilauer, P. (2012) Canoco Reference Manual and User's Guide: Software for Ordination. Version 5.0 t.

Tripathi, B. M., Kim, M., Kim, Y., Byun, E., Yang, J.-W., Ahn, J., and Lee, Y. K. (2018) Variations in bacterial and archaeal communities along depth profiles of Alaskan soil cores. *Sci Rep* **8**: 504.

White, R. A., III, Brown, J., Colby, S., Overall, C. C., Lee, J.-Y., Zucker, J., *et al*. (2017) ATLAS (automatic tool for local assembly structures) – a comprehensive infrastructure for assembly, annotation, and genomic binning of metagenomic and metatranscriptomic data. *PeerJ Prepr* **5**: e2843v1: 1–11.

Wilhelm, R. C., Niederberger, T. D., Greer, C., and Whyte, L. G. (2011) Microbial diversity of active layer and

permafrost in an acidic wetland from the Canadian high Arctic. *Can J Microbiol* **57**: 303–315.

Wilson, B., Müller, O., Nordmann, E.-L., Seuthe, L., Bratbak, G., and Øvreås, L. (2017) Changes in marine prokaryote composition with season and depth over an Arctic polar year. *Front Mar Sci* **4**: 95.

Xue, K., Yuan, M. M., Shi, Z. J., Qin, Y., Deng, Y., Cheng, L., *et al*. (2016) Tundra soil carbon is vulnerable to rapid microbial decomposition under climate warming. *Nat Clim Chang* **6**: 595–600.

Yang, S., Wen, X., Jin, H., Wu, Q., and Chavarria, K. (2012) Pyrosequencing investigation into the bacterial Community in Permafrost Soils along the China-Russia Crude Oil Pipeline (CRCOP). *PLoS One* **7**: e52730.

Yergeau, E., Hogues, H., Whyte, L. G., and Greer, C. W. (2010) The functional potential of high Arctic permafrost revealed by metagenomic sequencing, qPCR and microarray analyses. *ISME J* **4**: 1206–1214.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1: Fig. S1. Heat map displaying the highly represented (additive relative abundance >90%) bacterial groups at class level across the permafrost depth profile. Each of the bacterial groups is coloured according to its maximum abundance at any depth of the core. The darkest blue illustrates maximum, yellow medium and red lowest abundances.

Fig. S2. CO_2 emissions over 103 days incubation of permafrost samples, calculated for the period between days 19 and 122. Experiments were performed with 2 g soil samples from five segments of core 1 covering the AL, TZ and PL (Supporting Information Table S2). Samples were incubated at 4–6°C under aerobic (+) or anaerobic (–) conditions.

Fig. S3. Soil chemistry, including DOC, DTN, NH_4-N, C/N ratio, pH and water content of the permafrost incubations of the five segments from core 1, measured before (black) and after 19 days of incubation under aerobic (blue) and anaerobic (green) conditions.

Fig. S4. PCA of 16 metagenomes based on frequencies of KEGG annotated read frequency. Samples are coloured according to segments from the incubation experiment. Dark blue = Active Layer 1; Light blue = Active Layer 2; Green = Transition Zone 3; Red = Permafrost Layer 5. The time point of incubation is given by the sample label and only aerobic incubations are included in the analysis.

Table S1. Detailed 16S rRNA sequencing information regarding the number of replicates and sequences before and after processing in Qiime for the entire soil core profile.

Table S2. Overview of permafrost core segments used for the incubation experiments containing the depth and the characteristic layer represented by those segments.

Table S3. ANOSIM analysis of Euclidean distances between 16S rRNA gene sequencing data from the high resolution soil profile (core 1) and the sequencing data from the five chosen layers used in the incubation experiment from core 1 and core 2. The number indicates the grade of dissimilarity (1 = most dissimilar). The best similarity value for each incubation segment of the two cores is highlighted with a light green background.

Table S4. Summary of metagenomics sequencing results.