

Metaproteomics reveals functional partitioning and vegetational variation among permafrost-affected Arctic soil bacterial communities

Samuel E. Miller,¹ Albert S. Colman,¹ Jacob R. Waldbauer¹

AUTHOR AFFILIATION See affiliation list on p. 15.

ABSTRACT Microbial activity in Arctic soils controls the cycling of significant stores of organic carbon and nutrients. We studied *in situ* processes in Alaskan soils using original metaproteomic methods in order to relate important heterotrophic functions to microbial taxa and to understand the microbial response to Arctic greening. Major bacterial groups show strong metabolic specialization in organic topsoils. α - β - γ -Proteobacteria specialized in the acquisition of small, soluble compounds, whereas Acidobacteria, Actinobacteria, and other detritosphere groups specialized in the degradation of plant-derived polymers. α - β - γ -Proteobacteria dominated the expression of transporters for common root exudates and limiting nitrogenous compounds, supporting an ecological model of dependence upon plants for carbon and competition with plants for nitrogen. Detritosphere groups specialized in distinct substrates, with Acidobacteria producing the most enzymes for hemicellulose depolymerization. Acidobacteria was the most active group across the three plant ecotypes sampled—the largely nonvascular, lower biomass intertussock and the largely vascular, higher biomass tussock and shrub. Functional partitioning among bacterial groups was stable between plant ecotypes, but certain functions associated with α - β - γ -Proteobacteria were more strongly expressed in higher biomass ecotypes. We show that refined metaproteomic approaches can elucidate soil microbial ecology as well as biogeochemical trajectories of major carbon stocks.

IMPORTANCE The Arctic is warming twice as fast as the rest of the planet, and Arctic soils currently store twice as much carbon as the entire atmosphere—two facts that make understanding how Arctic soil microbial communities are responding to climate change particularly urgent. Greening of vegetation cover across the Arctic landscape is one of the most prominent climate-driven shifts in Arctic terrestrial ecology, with potentially profound effects on biogeochemical cycling by the soil microbiome. Here we use metaproteomics to document microbial metabolic functions that drive soil carbon and nutrient cycling processes in an Arctic tundra landscape. We identify functional roles among bacterial taxonomic groups that are largely stable across vegetation types, with certain functions strongly expressed by rhizosphere groups reflecting a community metabolic response to greening.

KEYWORDS soil, arctic, metaproteomics, permafrost, metabolism, microbial ecology, greening, warming

Arctic soils contain large quantities of organic carbon susceptible to respiration in a warming climate. Though Arctic soils comprise only 16% of soil surface area (1), they hold more than one-third of the global stock of soil organic carbon and roughly twice as much carbon as the entire atmosphere (2, 3). Low temperatures, high soil moisture, and

Editor Ileana M. Cristea, Princeton University, Princeton, New Jersey, USA

Address correspondence to Jacob R. Waldbauer, jwal@uchicago.edu.

The authors declare no conflict of interest.

Received 9 December 2022

Accepted 11 April 2023

Published 5 June 2023

Copyright © 2023 Miller et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

high acidity currently limit the microbial decomposition of abundant, labile soil organic matter (SOM) (4, 5), but with the current trajectory of warming, Arctic soils may release the equivalent of multiple years of anthropogenic C emissions by 2100 (6, 7). Currently, no consensus exists as to whether climate warming will result in net loss of stored soil carbon. Models and meta-analyses indicate either declines (8–12) or no net change (13, 14) in soil C stocks, with wide uncertainties (15, 16) especially in the Arctic (17). Arctic warming has widespread ecosystem-level consequences, which affect both the soil and the above-ground plant community in complex, interacting ways (18, 19). For Arctic tundra landscapes, these ecosystem consequences include lengthened growing seasons and increased plant biomass with concomitant changes in rooting depth, plant-microbial interactions, and nutrient demands (14, 20, 21).

Greening of vegetation cover across the Arctic landscape—notably the shift from tussock tundra to taller, higher-biomass woody shrubs—is one of the most prominent climate-driven shifts in Arctic terrestrial ecology, and its consequences for the soil microbial community remain unclear (22–24). Long-term greenhouse warming experiments at Toolik Field Station, Arctic Alaska, USA, have documented progressive reduction in topsoil trophic complexity and increases in plant stature, rooting depth, leaf litter accumulation, the evenness of topsoil microbial diversity, and subsoil microbial respiration (25, 26). Greening itself might substantially offset soil C losses (27), since woody plant biomass can store more carbon for a given amount of nitrogen, generally the limiting nutrient during the Arctic growing season (28). Interactions between microbes in the seasonally thawed active layer of the soil and the overlying, progressively greening vegetation are biogeochemically critical, since roughly one-third of Arctic soil carbon is within 1 m of the land surface (2).

Uncertainty in predictions of changes in Arctic soil biogeochemistry is due in part to an incomplete understanding of interactions between microbes, plants, and the physical environment (29–31). The structural complexity and occlusion of SOM precludes the *in situ* measurement of most reaction intermediates linking plant inputs to efflux gases, such as CO₂, via microbial metabolisms (32). One approach to analyzing SOM transformations is the identification of characteristic relationships between microbial taxonomic groups and processes they control (33). Surveys of 16S rRNA diversity in soils have revealed that the relative abundances of high-level bacterial clades strongly correlate with edaphic properties such as acidity and plant productivity (33–37). Many genes encoding key heterotrophic processes are distributed among phylogenetically diverse genomes, often varying in their occurrence between closely related strains (38–40). Taxa sharing the potential to express a metabolic pathway but exhibiting different growth rates, resuscitation rates from dormancy, or substrate use efficiencies can respond divergently to a given condition, such as substrate concentration (41, 42). The increasing use of methods that more directly interrogate phenotype has begun to clarify functional partitioning among taxa (43). Exometabolomic experiments have shown that rhizosphere bacterial isolates grown on natural mixtures of exudate metabolites preferentially assimilate classes of compounds exuded at different stages of plant development (43, 44). Population-resolved metatranscriptomics has revealed the emergence and succession of carbohydrate-degrading taxa associated with plant growth and detritus addition (45). Stable isotope probing has identified patterns of substrate utilization among both broader and narrower taxonomic groups (46), including transference of the isotopic label between groups over the course of days, suggesting interactions across trophic levels (47).

Some of the aforementioned approaches have the drawback of extrapolating experimental manipulations to *in situ* activity in soil. Metaproteomic analyses can characterize *in situ* protein expression of microbial communities and thereby functional partitioning realized in the environment, with recent studies linking proteins to metagenome-assembled genomes from deeply sequenced soils (48–50). The application of metaproteomics to complex ecosystems has been hindered by low identification rates of peptide mass spectra and the challenge of relating peptides to annotated proteins

(51, 52). We developed a novel computational pipeline called ProteinExpress to increase the number and quality of protein identifications from complex samples and applied ProteinExpress to Arctic soil samples from Toolik Field Station. We hypothesized that greening will restructure microbial community protein expression patterns, promoting metabolic activities involved in interactions with plant roots and altering soil nutrient cycling. To understand changes in microbial activity associated with greening, soils were analyzed from three plant ecotypes: the lower biomass, largely nonvascular intertussock ecotype and the higher biomass, and largely vascular tussock and woody shrub ecotypes. These metaproteomic data reveal functional traits of major bacterial groups and how they change with vegetation.

MATERIALS AND METHODS

Sample collection and proteomic mass spectrometry

Cores of permafrost-affected soil were collected at the end of the lower Arctic growing season (10–12 August 2014) in the vicinity of Toolik Field Station, Alaska, USA (TFS) (Fig. S1; Table S1). Cores were taken from three of the common plant ecotypes of Arctic North American moist acidic tundra: tussock, intertussock, and shrub (53, 54). The sedge, *Eriophorum vaginatum*, forms hemispherical tussocks ~0.5 m in radius, with a dense structure of slender roots extending through the organic soil to the base of the active layer (55, 56). Tussock sedges are spaced ~0.5 m apart and surrounded by diverse, low-lying intertussock vegetation, including mosses, lichens, and prostrate herbaceous plants (57). Woody shrubs ~1 m in height (*Betula nana* and *Salix pulchra*) are abundant around near-surface runoff channels, with nonvascular and herbaceous plants growing in the understory. Shrubs have a higher density of fine roots distributed at shallower depths in the organic layer than tussocks (55, 58). Sampling sites were located in the north–south facing hillslopes just south of TFS and the west facing hillslope of Imnavait Creek valley 12 km east of TFS. The TFS and Imnavait soils (ruptic-histic aquiturbels) developed from glacial outwash deposits of ~55 ka and ~125 ka and have average soil pH values of 4.34 and 4.80, respectively (59, 60). Average TFS soil temperatures at 5 cm and 50 cm depth over the sampling period were 12°C and 5°C, respectively (61). Cores to the permafrost or water table were extracted using a serrated push-corer (6.35 cm inner diameter), returned to TFS on blue ice, and frozen at –80°C within 3 h of sampling. Cores were transported from TFS to the University of Chicago on dry ice and stored at –80°C.

Metaproteomic datasets were collected from 15 organic topsoil samples ~5 cm below the surface and from three mineral subsoil samples ~5 cm below the organic-mineral boundary. Core sectioning, protein extraction, protein digestion, peptide recovery, and LC-MS analysis are detailed in Supplemental Text (doi.org/10.6084/m9.figshare.c.6485806). Protein extraction was based on the method of Chourey et al. (62). Protein digestion and peptide recovery followed a modified eFASP (enhanced Filter-Aided Sample Preparation) procedure (63). LC-MS analysis was performed as in Miller et al. (64). Proteomic mass spectral data is archived in the MassIVE repository under accession MSV000084386.

Data analysis

Overview of ProteinExpress software

We developed a Python 3 application, ProteinExpress (Fig. S7), for protein identification and annotation. ProteinExpress is designed to address the problem of database incompleteness in metaproteomic peptide-spectrum matching and protein inference, which can be a significant issue when the search database consists of population bins of contigs from a paired metagenomic dataset. Genomic microdiversity prevents complete sampling of every sequence variant, even in cold Arctic soils with relatively low diversity compared to warmer soils (65). Additionally, many reads are not assigned to contigs, let alone bins, especially in metagenomes from complex environments (66). Therefore,

a match between a peptide mass spectrum and a protein-coding sequence does not guarantee that the mass spectrum originated from the taxon associated with the protein, as opposed to another taxon producing the same peptide in an unsampled coding sequence.

ProteinExpress addresses database incompleteness by first searching mass spectra against both reads and contigs from metagenomes/metatranscriptomes and then aligning matching sequences to the full set of bins representing the major taxonomic groups in the environment. The alignment scores to each bin from sequences assigned the same function (e.g., all sequences with spectral matches annotated as ATP synthase) are used to measure the relative expression levels of the function by the suite of taxa. These scores and the overall relative abundances of proteins, as determined from spectral counts, factor into the metric, Φ , introduced below, which is used to compare the protein expression profiles of taxonomic groups and how a single taxonomic group's expression profile differs between samples. ProteinExpress uses both new algorithms and existing bioinformatic tools, which are implemented by the software in a pipeline (Fig. S1 and S2). ProteinExpress is publicly available at <https://github.com/semiller10/protein-express>.

Construction of search databases

Peptide sequences were assigned to mass spectra by searching a sequence database generated from publicly available Alaskan soil metagenomic and metatranscriptomic datasets (Table S2), combining sequence data from multiple sites and sequencing projects to achieve representation of the protein sequence space of the sampled Toolik-area soils. We used 12 metagenomes and 6 metatranscriptomes from Imnavait, the area of one of our metaproteomic sampling sites (67), and 10 metagenomes from the CiPEHR site 200 km south of TFS (65). The ProteinExpress pipeline trimmed reads from these 28 2×150 bp datasets with SolexaQA++ dynamictrim at a 1% nucleotide error probability cutoff (68). Full and partial prokaryotic genes were called and translated with Prodigal (69). Predicted proteins from each dataset were substring dereplicated with CD-Hit (70), forming 28 search databases of unassembled sequences from each reference dataset. Contigs were assembled by MEGAHIT from each dataset (minimum *k*-mer length 27, maximum length 87, and step size 10) (71). Using the Graph2Pro pipeline (66), another 28 databases of predicted proteins were produced from the assembled datasets.

We binned contigs from the 12 Imnavait metagenomes sharing a common taxonomic rank, e.g., bins of *Pseudomonas* and Rhizobiales (Table S3; Fig. S7). The method used to construct the bins was designed to maximize both the scope and quality of taxonomic bins by starting with single assemblies of the 12 Toolik-area metagenomes and then rebinning co-assembled reads mapping to the set of preliminary bins with the same taxonomy. Initially, assemblies from each dataset were kept separate. Contigs $\geq 2,500$ bp were binned by MetaBAT 2 on the basis of read coverage, as determined by Bowtie 2 (72), and tetranucleotide frequency (73). Binned contigs were aligned to RefSeq (release 83) using DIAMOND (74), retaining 500 hits per query. The lowest common ancestor was assigned from subject sequences with alignment bit scores within 10% of the top bit score and its taxonomic hierarchy retrieved from the NCBI Taxonomy Database (75). Taxonomic ranks of binned contigs were compared in *anvi'o* (76). We excluded bins representing taxa, such as Verrucomicrobia, that did not have at least 90% CheckM single-copy gene completeness. The final bins captured ~80% of the prokaryotic diversity in the active layer found in Toolik-area 16S rRNA libraries (Fig. S2). Taxonomically consistent contigs were exported and pooled with contigs sharing the same taxonomic classification from the other datasets. Reads mapping to these contigs were coassembled with MEGAHIT as before, generating a final set of contigs for each taxonomic bin.

Proteomic data analysis

The ProteinExpress pipeline used MS-GF+ to search each metaproteomic dataset against the 28 databases of unassembled reads and 28 databases of assembled contigs

(Table S1) (77). Each spectrum could return a peptide-spectrum match (PSM) from multiple reads and/or contigs. Database search against assembled nucleotide datasets was refined with the Graph2Pro pipeline (see Supplemental Text at doi.org/10.6084/m9.figshare.c.6485806) (66). PSMs meeting the 1% false discovery rate cutoff from each target-decoy search were selected (78). The one or more full- or partial-length proteins containing a PSM are henceforth called PSM-protein candidates.

PSM-protein candidates were aligned by BLASTp against RefSeq (release 83) to find homologous proteins (79, 80). Of the 500 subject sequences returned per query, those within the top 10% by bit score were retained. If more than five subject sequences were selected, then five were evenly subsampled in descending order of bit score to make the following steps computationally tractable. In the case of a PSM-protein candidate with 500 hits to identical or near-identical RefSeq sequences, hits #1, 125, 250, 375, and 500 would be retained. Selected subject sequences were sorted by NCBI Taxonomy ID into fasta files for Bacteria and Archaea and searched by eggNOG-mapper (via DIAMOND) against superkingdom-level databases of protein Orthologous Groups (see Supplemental Text at doi.org/10.6084/m9.figshare.c.6485806) (81, 82). At this point in the ProteinExpress pipeline, each PSM is associated with a set of PSM-protein candidates, which in turn are associated with up to five functionally annotated RefSeq hits. Only PSMs with identical eggNOG functional description strings (e.g., “Citrate synthase”) among all of the PSM-protein candidate hits were retained. A set of PSMs sharing a functional annotation is called a metaprotein (83).

For purposes of interpretation, we defined 141 Functional Groups of significance for cellular biology and biogeochemical cycles from 2,678 unique pairs of eggNOG gene family and description annotations (Table S4 at doi.org/10.6084/m9.figshare.c.6485806). For example, “CITA” + “Citrate synthase” was assigned to the “TCA Cycle” Functional Group. Some Functional Groups, such as “TCA Cycle,” are equivalent to biochemical pathways. Others, such as “Ribosome,” “Transposase,” and “Ribose Transport,” are comprised of functions not necessarily related through a common pathway. When needed, eggNOG-mapper gene family and functional descriptions were cross-referenced to databases (UniProtKB, InterPro, KEGG Orthology, and CAZy) to assign Functional Groups. For example, sequences with a description of “Hydrolase, family 38” lacking an eggNOG gene family assignment were assigned to “Mannose Cleavage,” as CAZy records that enzymes assigned to Family 38 on the basis of sequence homology are only known to be active on α -linked mannose residues (84).

The relative expression levels of peptides, metaproteins, and Functional Groups were estimated by NSAF (normalized spectral abundance factor) (85):

$$\text{NSAF}_N = \frac{D_N/L_N}{\sum_{i=1}^n (D_i/L_i)} \quad (1)$$

where N is the peptide index; D_N is the number of spectra matched to peptide N ; L_N is the average length of protein subject sequences returned from BLASTp searches of PSM-protein candidates containing peptide N (see above); and n is the total number of peptides identified in the dataset. Spectrum count is normalized to protein length, as longer proteins generate more tryptic peptides and are thereby more likely to be observed. The average standard deviation of relative subject sequence lengths from sets of PSM-protein candidates was 2.0%. Metaprotein and Functional Group NSAF values were calculated by summing the NSAF values of peptides associated with the given metaprotein or Functional Group.

We developed a metric, Φ , to estimate relative metaprotein expression levels within and between taxa. Φ takes into account both the relatedness of a metaprotein to a particular bin of contigs (serving as a measure of the likeliness of expression by organisms in the bin) and its NSAF, which represents the level of metaprotein expression irrespective of the bin. To determine Φ , each peptide’s PSM-protein candidates were aligned to each bin by BLASTp. The single candidate that produced the alignment with the lowest E-value was selected, only considering alignments with E-values

<0.01. Selection of this candidate ensures the strongest possible nonrandom association between a peptide and each taxonomic bin. The strength of each association was measured by the alignment's bit score, S' . S' was normalized to the range, [0, 1], by subtracting the minimum S' among the bins (S'_{\min}) and dividing by $S'_{\max} - S'_{\min}$. Normalization allows comparison of relative levels of production of a peptide among the taxonomic groups in the sample—which must be well-represented by the bins—with minimum and maximum production by groups corresponding to values of 0 and 1, respectively. The normalized value of S' was multiplied by the NSAF value of the peptide in the dataset, producing φ_N , a metric representing both peptide abundance and taxonomic similarity.

$$\begin{cases} \varphi_N = \text{NSAF}_N \times \frac{S' - S'_{\min}}{S'_{\max} - S'_{\min}} & \text{if E-value of peptide-bin alignment} < 0.01, \text{ else} \\ \varphi_N = 0 & \end{cases} \quad (2)$$

φ_N was extended to metaproteins and Functional Groups (Table S4 at doi.org/10.6084/m9.figshare.c.6485806) by summing φ_N values of PSMs categorized as a certain metaprotein or Functional Group, P .

$$\varphi_P = \sum_i^n \varphi_{N,i} \quad (3)$$

We used Φ_{bin} , which scales φ_P to the maximum value across bins ($\varphi_{P,\text{max bin}}$), to compare patterns of functional partitioning between taxonomic groups. We used Φ_P , which scales φ_P to the maximum value across metaproteins or Functional Groups ($\varphi_{P,\text{max function}}$), to compare a taxonomic group's protein expression profile between samples. In either case, if the maximum value of φ is 0, Φ is assigned a value of 0. Φ_{bin} and Φ_P lie in the range [0, 1], with at least one value across the set of bins or functions, respectively, equal to 1.

$$\Phi_{\text{bin}} = \frac{\varphi_P}{\varphi_{P,\text{max bin}}} \quad (4)$$

$$\Phi_P = \frac{\varphi_P}{\varphi_{P,\text{max function}}} \quad (5)$$

To understand differences in the functional profiles of taxonomic groups between plant ecotypes, linear discriminant analysis and principal component analysis were conducted with Φ_{bin} and Φ_P values. For analysis of between-ecotype variance of a given taxon's functional profile, each data point in the multivariate analysis is a taxonomic group with a vector of Φ_{bin} values over all Functional Groups. For analysis of between-ecotype variance of a given function's distribution between taxa, each data point is a Functional Group with a vector of Φ_P values over all taxonomic groups.

RESULTS

Identification of proteins and activity of taxonomic groups

We identified peptide sequences for 3,142–16,235 MS2 spectra (PSMs) from each of the 15 organic and 3 mineral soil samples, of which 1,079–5,003 (27.1–36.9%) were functionally annotated (Fig. S1; Table S1). Metaproteins, or groups of PSM-containing proteins with the same functional annotation, were compared to 12 bacterial taxonomic bins identified from Toolik-area topsoil metagenomes (Tables S2 and S3); prior work indicates that similar sets of taxa are present in the organic and mineral layers of tundra soils (86) and over the course of the growing season (87), albeit at differing relative abundances. The taxonomic specificity of the bins ranged from genus to phylum level. Bins represented the same phylum or class as 81.5% of the reads on average from nine Toolik-area 16S rRNA datasets (Fig. S2; 87–91) with Verrucomicrobia, Firmicutes, other lower abundance phyla, and unclassified reads not represented by bins.

Our calculated metric, Φ_{bin} (see Equation 4), was used as a measure of the relative expression of a given metaprotein by different taxonomic groups. We identified overarching patterns in functional partitioning by the 12 major bacterial groups using *k*-means clustering of all Functional Group (Table S4 at doi.org/10.6084/m9.figshare.c.6485806) Φ_{bin} values averaged across all organic soil samples, irrespective of plant ecotype. A breakpoint in the reduction of the sum of squared errors with the addition of clusters exists at *k* = 3 clusters (Table S5 at doi.org/10.6084/m9.figshare.c.6485806), which showed consistent expression patterns between taxonomic bins. Protein expression of Functional Groups in Cluster 1 was largely attributed to α - β - γ -Proteobacteria; proteins in Cluster 2 were expressed most strongly by Acidobacteria, but also at high levels by α - β - γ -Proteobacteria; and functions in Cluster 3 were dominated by Acidobacteria. Other taxa, including Actinobacteria, Bacteroidetes, and Myxococcales (δ -Proteobacteria), produced fewer proteins for most functions.

Many core cellular functions, including DNA replication and repair, transcription, translation, and cell division, fall into Cluster 2—being most highly expressed by Acidobacteria and secondarily by α - β - γ -Proteobacteria (Fig. 1)—suggesting that these two groups are highly active in the organic soils. Ribosomal proteins and chaperones, along with nucleoid structuring proteins, are among the most abundant proteins in the dataset (NSAF values >0.01; Fig. 1), reflecting the high metabolic and nutrient cost of central housekeeping functions. Cell envelope-related functions (Fig. S3) showed broadly similar distributions to core functions, being most highly expressed by Acidobacteria and α - β - γ -Proteobacteria. As expected, gliding motility proteins associated with Bacteroidetes and Myxococcales were found to be expressed essentially solely by those groups, validating the taxonomic specificity of protein assignment by the ProteinExpress pipeline and suggesting that taxon-specific functions were sufficiently represented in our sequence bins.

We also examined the expression of Functional Groups when normalized to that of ribosomal proteins (by subtracting $\Phi_{\text{bin,Ribosome}}$ values, as a proxy for overall cellular activity; Fig. S4). Taxonomic expression profiles with marked differences in ribosomal proteins suggest that the function was not expressed in proportion to an overall activity of the different taxonomic groups. This comparison reinforces the result from clustering that many functions were skewed strongly either toward (Cluster 1) or away (Cluster 3) from α - β - γ -Proteobacteria.

Carbon cycling and energy conservation

Acidobacteria, class Actinobacteria, Bacteroidetes, and Myxococcales dominated the expression of extracellular enzymes responsible for depolymerization of plant cell wall polysaccharides (Fig. 2). Acidobacteria had the highest Φ_{bin} values for most core C metabolism and energy conservation functions, including the TCA cycle and ATP synthase. The group strongly expressed enzymes acting on diverse heteropolysaccharide bonds, including β -1,4-linked D-xylose, β -1,4-linked D-glucose, and β -1,4-linked D-mannose, and likewise expressed enzymes required for the metabolism of diverse pentoses, such as xylose and arabinose. Acidobacteria also produced the highest levels of enzymes involved in the breakdown of starch and glycogen, including α -amylase, 4- α -glucanotransferase, amylo- α -1,6-glucosidase, and α -glucosidase. Overall, these protein expression patterns point toward catabolism of hemicelluloses, starch, and glycogen as the main carbon and energy sources for the highly active Acidobacteria in these soils.

C metabolism protein expression patterns of other bacterial groups indicate specialization for breakdown of other plant-derived substrates. Actinobacteria dominated the expression of endoglucanase and cellobiosidase, enzymes required for the debranching and cleavage of oligosaccharides from cellulose. Bacteroidetes and Myxococcales were distinguished from Acidobacteria and Actinobacteria by strong expression of pectinases. Pathways such as the protocatechuate-4,5 cleavage pathway for the aerobic degradation of heterogeneous aromatic compounds derived from lignin were most heavily expressed by Burkholderiaceae. Anaerobic aromatic degradation pathways had significantly lower

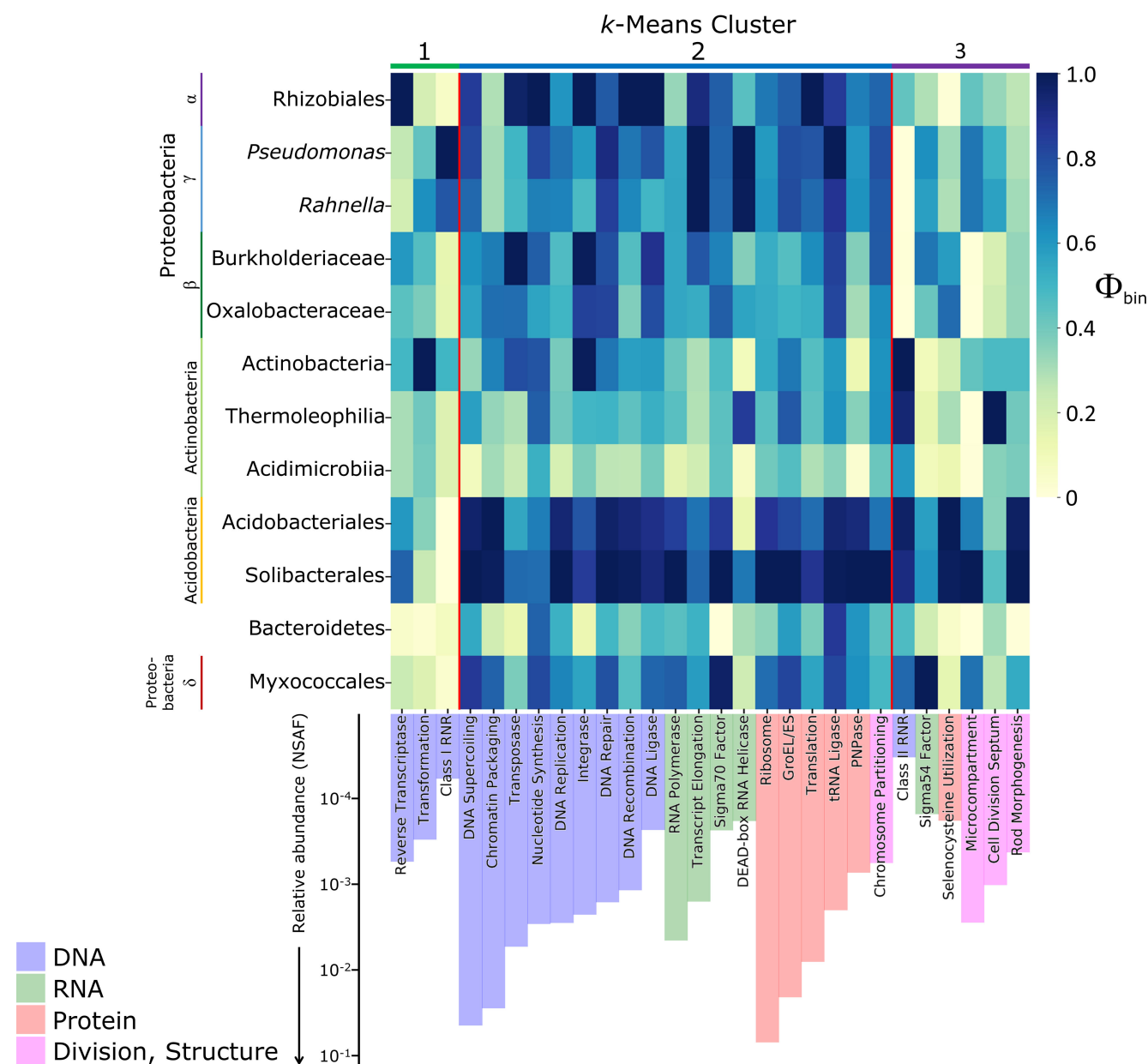


FIG 1 Cell growth-related Functional Group Φ_{bin} values (heatmap) and average overall relative abundances (bars) across organic soil samples. Columns are ordered by cluster assignment, then by Functional Group category (bar color), then by overall relative abundance (NSAF). Φ_{bin} values are column-normalized to between 0 and 1 based on $\Phi_{bin,max}$ for that column.

levels overall, which was generally true of anaerobic versus aerobic pathways in our samples. Lactate and ethanol fermentation enzymes were about two orders of magnitude less abundant than TCA cycle enzymes and ATP synthase, and terminal electron acceptors besides O_2 are relatively scarce in organic-rich tundra soils (92).

Different taxonomic groups dominated the expression of pathways required for utilization of small soluble molecules and gases. Carboxidotrophy and methylotrophy had relatively high overall abundances, with Rhizobiales dominating the expression of CO dehydrogenase, and β - and γ -Proteobacteria most strongly expressing methanol dehydrogenase (MxaF) and formaldehyde oxidation pathways. The nondetection of methane monooxygenase, low representation of Archaea in nucleotide datasets, and identification of only three metaproteins with strong alignments to Archaea in RefSeq indicate that biogenic methane was not a significant substrate. Rhizobiales produced the

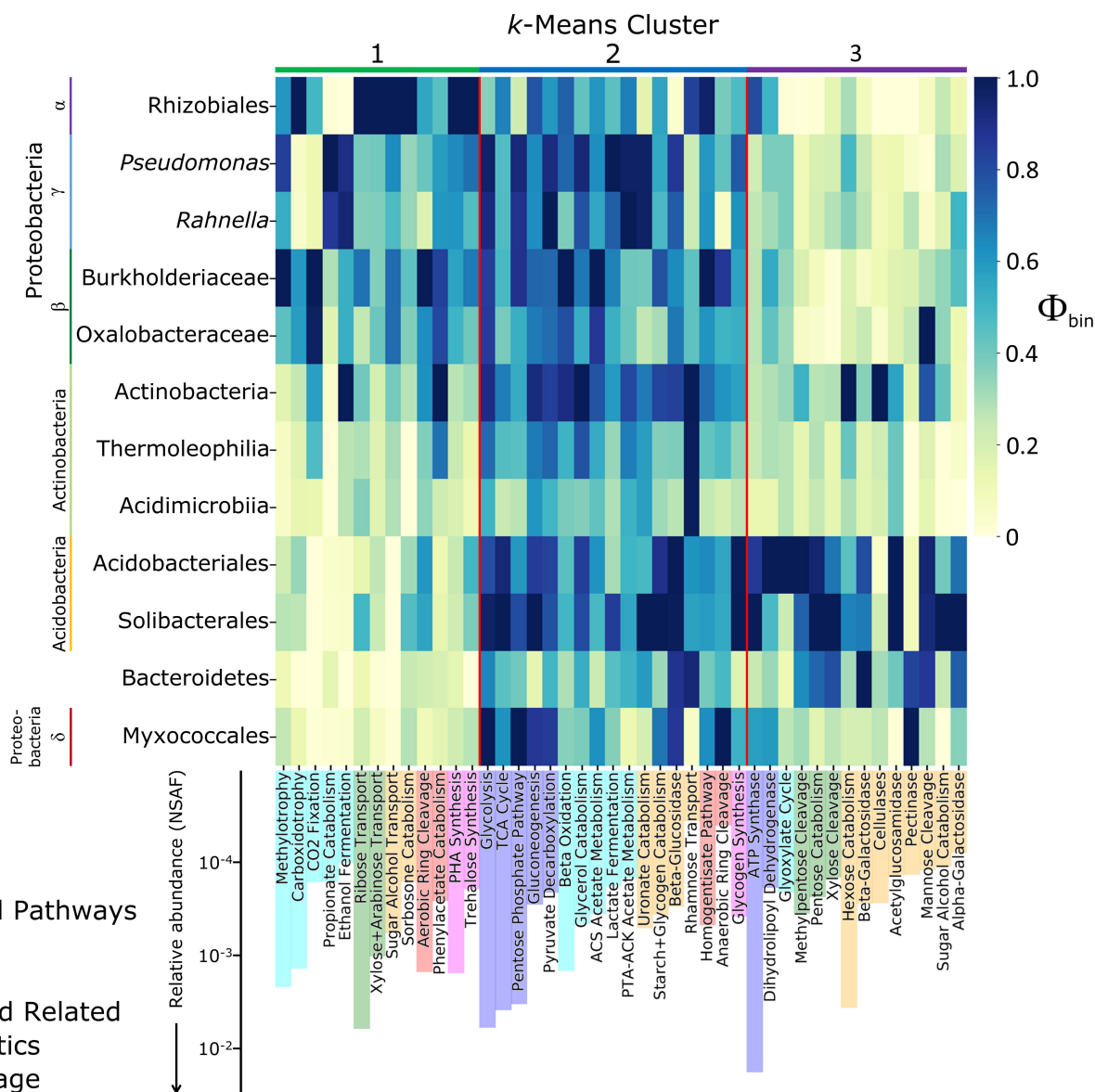


FIG 2 Carbon-related Functional Group Φ_{bin} values (heatmap) and average overall relative abundances (bars), across organic soil samples. Columns are ordered by cluster assignment, then by Functional Group category (bar color), then by overall relative abundance (NSAF). Φ_{bin} values are column-normalized to between 0 and 1 based on $\Phi_{bin,max}$ for that column.

highest levels of monosaccharide transporters but low levels of enzymes for polysaccharide depolymerization, glycolysis, and the catabolism of nonglucose monosaccharides. Rhizobiales dominated the expression of enzymes for polyhydroxyalkanoate synthesis, a C storage mechanism with higher overall abundance than the other identified C storage pathways of trehalose and glycogen synthesis.

Nutrient cycling

Proteobacteria—particularly Rhizobiales—dominated the production of N transporters (Fig. 3). Proteins involved in amino acid transport had a higher overall abundance than those involved in oligopeptide and polyamine transport and urea assimilation, which in turn were more abundant than proteins involved in ammonium transport, nitrate/nitrite reduction, and N_2 fixation. Intracellular N cycling functions, such as amino acid synthesis and catabolism pathways, were much more evenly expressed between taxonomic groups. The abundant ammonia metabolism Functional Group, which largely

consists of the glutamine synthetase metaprotein, was more strongly expressed by Acidobacteria.

Phosphate acquisition proteins, including phosphate transporters and phosphatases, were higher in overall abundance than polyphosphate and phosphonate acquisition proteins, and were relatively evenly expressed across groups. Iron is known to be an important substrate in microbial iron mats that float on submerged soils around Toolik (93). In our soils, none of which were submerged, Fe-related functions were relatively abundant but not clearly related to dissimilatory Fe metabolism, which was not identified. Ferritins were especially abundant and overwhelmingly produced by Acidobacteria. Heme synthesis pathways were most heavily expressed by γ -Proteobacteria, and proteins involved in Fe import, largely through siderophore synthesis, were most strongly expressed by β -Proteobacteria. These functions had a more even taxonomic distribution than ferritin.

Differences between plant ecotypes

Multivariate analyses of Φ were used to understand differences in the functional profiles of taxonomic groups between plant ecotypes. Complementary linear discriminant analyses of Φ_{bin} and $\Phi_{\text{Functional Group}}$ values reveal that organic soil samples cluster by plant ecotype when considering the functional profiles of taxonomic groups (Fig. 4A) but not the taxonomic profiles of Functional Groups (Fig. 4B)—taxonomic groups maintained functional specialization across plant ecotypes, but the expression levels of functions varied across ecotypes. This result was supported by principal component analysis of Φ_{bin} values (Fig. 4C), which shows that taxonomic groups generally maintained their relative positions in functional space while the whole community of all taxonomic groups shifted systematically between environments.

Four Functional Groups in tussock and intertussock organic soil samples were found to have statistically significant differences (Welch's t-test) in overall relative abundance between environments (shrub is ignored due to the fewer number of metaproteomic datasets than tussock and intertussock): ribose transport ($P = 0.00021$), xylose+arabinose transport ($P = 0.019$), sugar alcohol transport ($P = 0.014$), and succinoglycan synthesis ($P = 0.014$) (Fig. S5; Supplemental Text at doi.org/10.6084/m9.figshare.c.6485806). These functions were more highly represented in tussock than intertussock samples and were most strongly expressed by Rhizobiales. Linear discriminant analysis of the NSAF data (Supplemental Text at doi.org/10.6084/m9.figshare.c.6485806) shows that organic soil samples distinctly cluster by plant ecotype, whereas the three mineral soil samples, one from each plant ecotype, form a separate cluster (Fig. S6).

DISCUSSION

We profiled microbial protein expression in Arctic Alaskan soils using a metaproteomic approach that accounts for sequence microdiversity, finding that major heterotrophic functions are consistently partitioned among broad taxonomic groups (Fig. 5). While this study was limited to a snapshot view of the late Arctic growing season, these data elucidate functional specialization among microbial taxa, clarifying how each group contributes to the breakdown of plant-derived organic carbon and to soil nutrient cycling. Acidobacteria, the most active group (as indicated by expression of ribosomal proteins, chaperones, and other abundant core proteins; Fig. 1), specialized in the degradation of relatively labile polysaccharides (Fig. 2). In accordance with findings from soils along a permafrost thaw gradient (48) and in *Sphagnum* peat (94), Acidobacteria produced a full complement of enzymes for xylan depolymerization and xylose assimilation: endo-/exo-xylanases, acetyl xylan esterase, xylose isomerase, and xylulokinase. We additionally found that Acidobacteria strongly expressed numerous glycosidases for the breakdown of diverse hemicelluloses and pectins, although Myxococcales and Bacteroidetes dominated the production of pectinases required for depolymerization of pectin's galacturonan backbone.

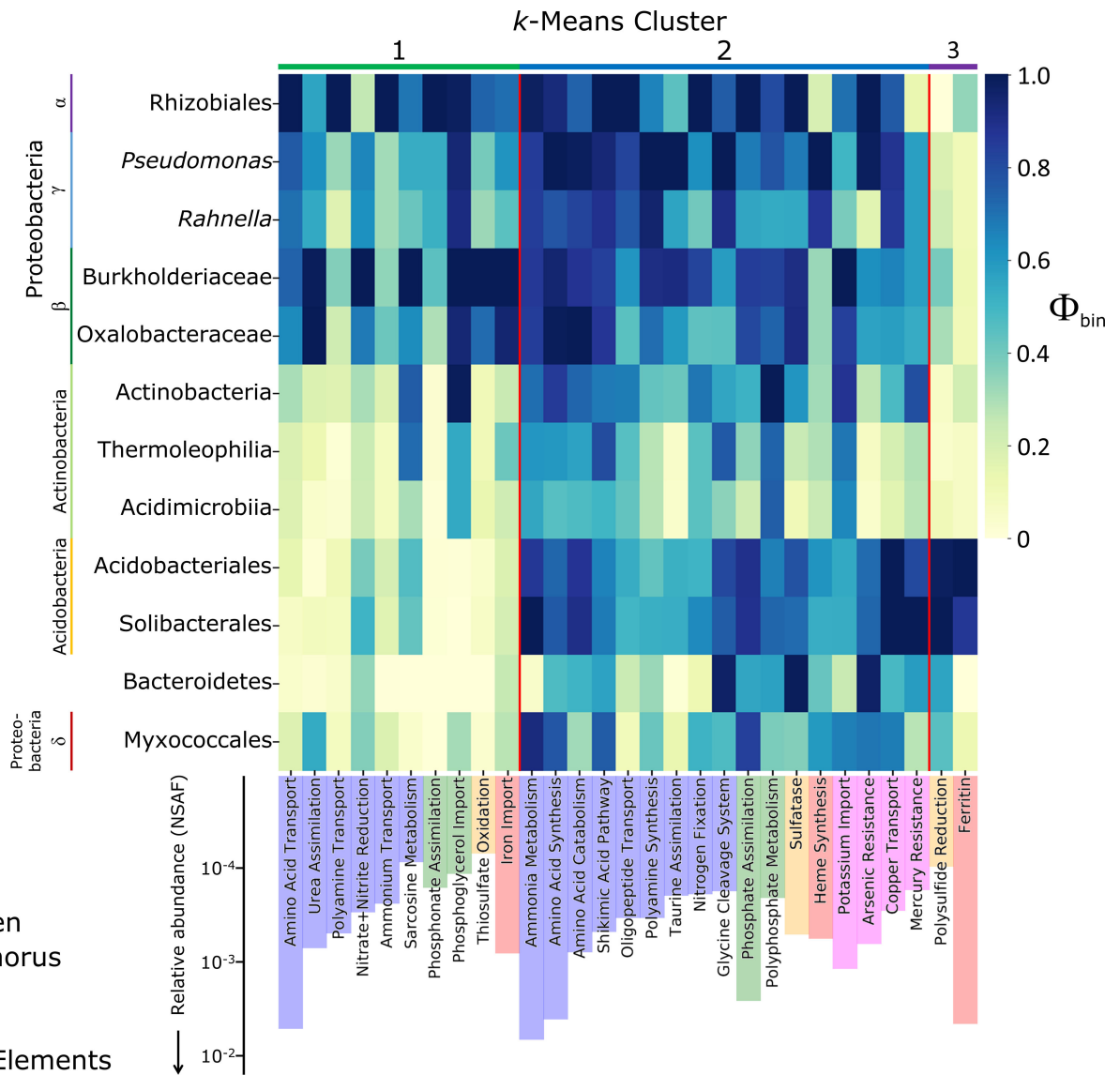


FIG 3 Nutrient-related Functional Group Φ_{bin} values (heatmap) and average overall relative abundances (bars), across organic soil samples. Columns are ordered by cluster assignment, then by Functional Group category (bar color), then by overall relative abundance (NSAF). Φ_{bin} values are column-normalized to between 0 and 1 based on $\Phi_{bin,max}$ for that column.

In these Toolik-area soils, Actinobacteria dominated cellulase production, in contrast to the Swedish soils in which Acidobacteria predominated, while Actinobacteria had high genomic potential for cellulases but negligible production of transcripts and proteins (48). Our results are consistent with the ability of the majority of Acidobacteria in pure culture to use xylan or starch but not crystalline cellulose or carboxymethylcellulose as a carbon source (95). A metatranscriptomic investigation of Svalbard peat soils also shares certain patterns of functional partitioning with those we observe at Toolik, including the importance of Actinobacteria in the expression of cellulases, Bacteroidetes in branched polysaccharide glycosidases, and Proteobacteria in phenolic degradation pathways (96).

α - β - γ -Proteobacteria specialized in the acquisition of small, soluble compounds, including monosaccharides, despite expressing very low levels of glycosidases. These taxa appear to be adapted primarily to the rhizosphere or other soil hot spots receiving pulses of labile carbon substrates or nutrients (97). At the time when we sampled for this

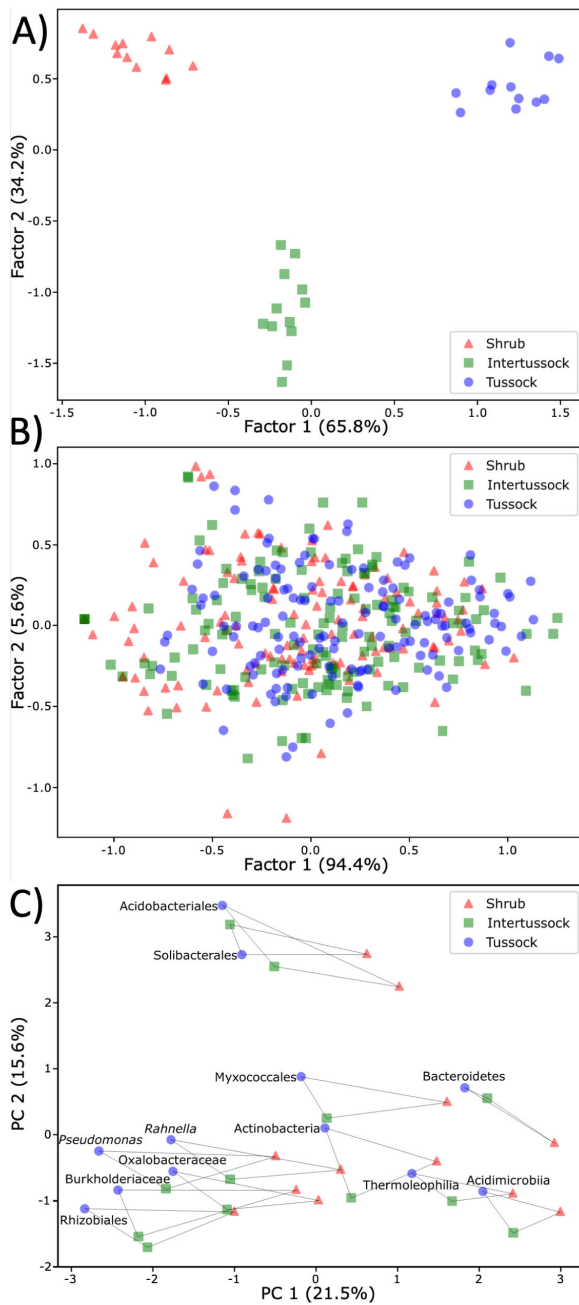


FIG 4 Linear discriminant analyses for organic soil samples between plant ecotypes of (A) Φ_{bin} values, with each data point representing one of 12 taxonomic bins in a given ecotype and (B) $\Phi_{Functional\ Group}$ values, with each data point representing one of 141 Functional Groups in a given ecotype. The percentage is the proportion of separation due to the discriminant function. (C) Principal component analyses of Functional Group Φ_{bin} values, showing principal components 1 and 2 (percentage indicates the proportion of variance explained by each PC), with lines connecting points representing the same taxonomic group in different plant ecotypes.

study, relatively late in the growing season, labile C compounds from root exudates may have been largely depleted, contributing to the observed lower apparent activity of Proteobacteria compared to Acidobacteria. Rhizobiales (α -Proteobacteria) dominated the expression of transporters for low molecular weight compounds that are among the most abundant types of root exudates (98) as well as proteins for the production of succinoglycan, an exopolysaccharide known to be essential for interactions between

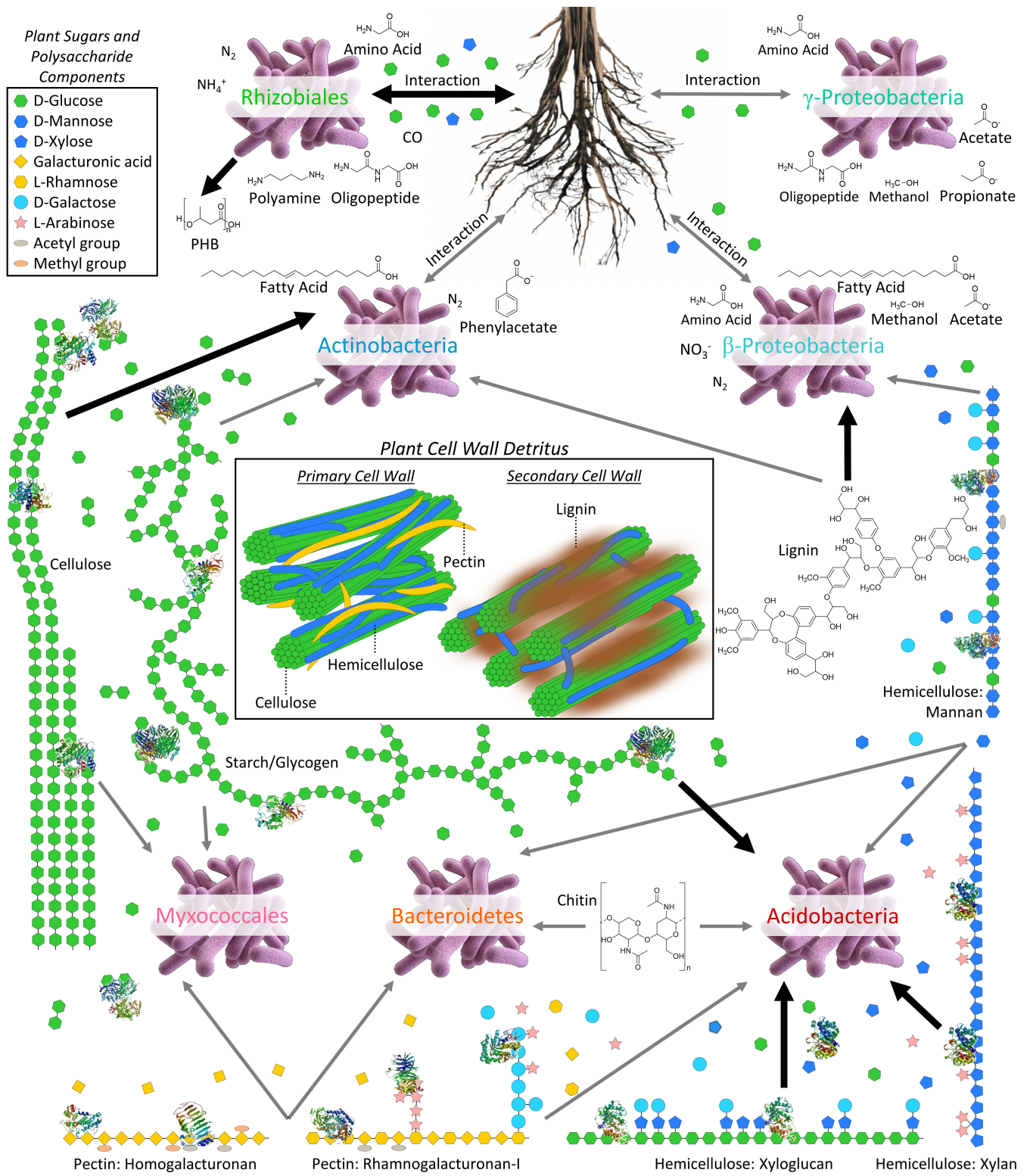


FIG 5 Summary of resource partitioning by major taxonomic groups in organic soils, as determined from Φ_{bin} data.

nodulating rhizobia and their leguminous hosts (99, 100). Monosaccharide and sugar alcohol transporters and succinoglycan production were the functions that increased most significantly in relative abundance between our soils with low and high root biomass. Furthermore, Rhizobiales dominated the production of CO dehydrogenase for

CO oxidation, a process that removes most CO produced by roots, a major source in soils (101, 102). Trace atmospheric CO may also provide a carbon source for subsistence in the absence of root exudates (103), which could be a widespread condition among Rhizobiales in Arctic soils in light of their high expression of the synthesis pathway for polyhydroxyalkanoates, polymers used for C and energy storage and metabolized during later periods of low plant productivity (104).

α - β - γ -Proteobacteria appeared to compete strongly with plants for scarce N in Toolik-area soils. By the time we sampled at the end of the growing season, bioavailable N—which is mainly in organic forms—has been found to fall to nearly undetectable levels in soils around Toolik (105). Two mechanisms have been proposed to explain pervasive N limitation of plant productivity in Arctic soils: slow microbial decomposition generating little bioavailable N and high N use efficiency by rhizosphere microbes, with the latter hypothesis supported by ^{15}N dilution experiments (106) and meta-analysis of N turnover rates (107). In our datasets, α - β - γ -Proteobacteria dominated the production of transporters for nitrogenous compounds, including amino acids, oligopeptides, polyamines, urea, and ammonium, whereas proteins involved in intracellular N cycling—glutamine synthetase being the most abundant—had relatively even expression between Proteobacteria and detritosphere groups such as Acidobacteria (Fig. 3). High N uptake efficiency among α - β - γ -Proteobacteria may be driven by depletion of N around growing root tips and demand for nutrients to generate biomass from influxes of root exudates or other soluble C sources (97, 98, 108). Our observation of cells poised for N assimilation agrees with the rapidly responding copiotrophic behavior of Proteobacteria in N fertilization experiments (109). We found that P transporters, in contrast, were more evenly expressed across taxonomic groups, which may be attributed to lower demand relative to available supplies, as Arctic soils are less often P limited.

Bacterial protein relative abundances differed between plant ecotypes in organic soils (Fig. 4A), yet functional partitioning across bacterial groups remained relatively invariant (Fig. 4B)—implying that taxa have consistent metabolic roles in soils supporting different types of vegetation, but the relative prominence of those various roles shifts. A 16S rRNA study of nearby soils documented a shift toward Proteobacteria and away from Acidobacteria in shrub versus tussock soils (87). It was inferred that copiotrophic Proteobacteria were better adapted to use the labile C pool in shrub soils (110). Given that we found higher relative abundances of α - β - γ -Proteobacterial functions in highly rooted tussock than sparsely rooted intertussock soils, it may be that greater root density and plant activity generally leads to the formation of rhizosphere hot spots favorable for Proteobacteria. Interestingly, the protein expression profiles of bacterial groups shifted with plant ecotype in the same directions along principal component axes (Fig. 4C), which may signify that different plant stimuli and inputs affected the expression of functions independent of whether those functions were expressed at high or low levels by particular bacterial groups. Lastly, the relatively uniform profiles of overall functional abundance in deeper mineral soils (Fig. S6) may be explained by higher representation of core cellular functions versus more specialized metabolism in relatively inactive cells or the advective homogenization of plant inputs from nearby ecotypes (111).

Patterns of functional partitioning that we found among major bacterial groups seem likely to persist as plant ecotypes continue to shift in response to Arctic warming. These patterns may also apply over large areas of the Arctic, as suggested by the occurrence of nearly identical metagenome-assembled genomes hundreds of kilometers apart (65). Building from the snapshot of the late Arctic summer season presented here, assaying potential shifts in functional partitioning over the full course of the growing season (and during wintertime, increasingly recognized as crucial to Arctic soil C cycling (112)) will test the broader applicability of these observations. Our study particularly highlights the metabolic interplay between rhizosphere α - β - γ -Proteobacteria and detritosphere groups such as Acidobacteria in determining the fate and rate of transformation of plant-derived organic carbon in Arctic soils. Understanding the biogeography of

functional roles of the full microbial community across seasons and in diverse soils will be crucial to robust predictions of the fate of Arctic soil carbon stocks.

ACKNOWLEDGMENTS

We are grateful to Anne Giblin for fieldwork preparation and coordination; to James Laundre for assistance at Toolik Field Station; to Maureen Coleman, Michael Foote, and David Archer for advice on data interpretation and methods; to Lichun Zhang for maintenance of the biogeochemical proteomics facility at UChicago; and to Gerard Olack for organizational aid and feedback. We also acknowledge the support of the Arctic Long-Term Ecological Research Program by the U.S. National Science Foundation (Award # NSF-1637459).

The authors declare they have no competing financial or other interests.

AUTHOR AFFILIATION

¹Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois, USA

PRESENT ADDRESS

Samuel E. Miller, Josephine Bay Paul Center, Marine Biological Laboratory, Woods Hole, Massachusetts, USA

Albert S. Colman, Department of Earth, Environmental, and Planetary Sciences, Rice University, Houston, Texas, USA

AUTHOR CONTRIBUTIONS

Samuel E. Miller, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review and editing | Albert S. Colman, Conceptualization, Funding acquisition, Supervision, Writing – review and editing | Jacob R. Waldbauer, Conceptualization, Formal analysis, Methodology, Project administration, Resources, Supervision, Visualization, Writing – original draft, Writing – review and editing

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Fig. S1 (mSystems01238-22-s0001.pdf). Map of sampling locations in the vicinity of Toolik Lake, Alaska. See Table S1 for details of individual sample numbers.

Fig. S2 (mSystems01238-22-s0002.pdf). 16S rRNA datasets from organic soils in the area of Toolik Field Station, collected during the growing season from June–August (55–59). Taxa above the red line in the legend are represented by metagenomic bins, whereas those below the line, as well as any other low abundance taxa not listed, are not represented by bins.

Fig. S3 (mSystems01238-22-s0003.pdf). Cell envelope-related Functional Group Φ_{bin} values (heatmap) and average overall relative abundances (bars) across organic soil samples. Columns are ordered by cluster assignment, then by Functional Group category (bar color), and then by overall relative abundance (NSAF). Φ_{bin} values are column-normalized to between 0 and 1 based on $\Phi_{bin,max}$ for that column.

Fig. S4 (mSystems01238-22-s0004.pdf). Functional Group Φ_{bin} values with $\Phi_{bin,Ribosome}$ values subtracted, and average overall relative abundances (bars) across organic soil samples. Columns are ordered by cluster assignment, then by Functional Group category (bar color), and then by overall relative abundance (NSAF). A) Cell growth-related Functional Groups; B) Carbon-related Functional Groups; C) Nutrient-related Functional Groups; D) Cell envelope-related Functional Groups.

Fig. S5 (mSystems01238-22-s0005.pdf). Overall expression levels (NSAF) of Functional Groups associated with A) DNA, B) RNA, C) translation, D) central carbon metabolism and energy conservation, E) polysaccharide degradation, F) carbon metabolism, and G) nitrogen, H) sulfur, I) phosphorus, J) trace elements, K) stress, L) membrane and wall synthesis, M) movement, N) cell division and structure, and O) other functions.

Fig. S6 (mSystems01238-22-s0006.pdf). Linear discriminant analyses of NSAF data from metaproteins defined by multiple functional annotation systems: A) GO terms, B) KEGG Orthology IDs, C) Gene Families + eggNOG descriptions, and D) Functional Groups of Gene Family + eggNOG description metaproteins. Each point is a different sample, with both organic (unlabeled points) and mineral soils (labeled) considered. The percentage is the proportion of separation due to the discriminant function.

Fig. S7 (mSystems01238-22-s0007.pdf). Flowcharts of metaproteomic data analysis in ProteinExpress and of taxonomic bin construction from metagenomes.

Table S1 (mSystems01238-22-s0008.pdf). Metaproteomic sample information. For map of sample locations, see Figure S1. "NA" in "Depth to water table (cm)" signifies that the water table was at the base of the active layer.

Table S2 (mSystems01238-22-s0009.pdf). Metagenomic and metatranscriptomic sample datasets used for metaproteomic search database construction. Asterisks indicate data sets that were used in binning.

Table S3 (mSystems01238-22-s0010.pdf). Taxonomic bins constructed from metagenomic data, taxonomic ranks, and number of contigs (all $\geq 2,500$ nt) per taxonomic bin.

REFERENCES

- Brown J, Ferrians OJ, Heginbottom JA, Melnikov ES. 1997. Circum-arctic map of permafrost and ground-ice conditions. Circum-Pacific Map. <http://pubs.er.usgs.gov/publication/cp45>.
- Hugelius G, Strauss J, Zubrzycki S, Harden JW, Schuur EAG, Ping C-L, Schirmermeister L, Grosse G, Michaelson GJ, Koven CD, O'Donnell JA, Elberling B, Mishra U, Camill P, Yu Z, Palmtag J, Kuhry P. 2014. Estimated stocks of circumpolar permafrost carbon with quantified uncertainty ranges and identified data gaps. *Biogeosciences* 11:6573–6593. <https://doi.org/10.5194/bg-11-6573-2014>
- Scharlemann JPW, Tanner EVJ, Hiederer R, Kapos V. 2014. Global soil carbon: understanding and managing the largest terrestrial carbon pool. *Carbon Manag* 5:81–91. <https://doi.org/10.4155/cmt.13.77>
- Zimov SA, Davydov SP, Zimova GM, Davydova AI, Schuur EAG, Dutta K, Chapin FS. 2006. Permafrost carbon: stock and decomposability of a globally significant carbon pool. *Geophys Res Lett* 33. <https://doi.org/10.1029/2006GL027484>
- Dutta K, Schuur EAG, Neff JC, Zimov SA. 2006. Potential carbon release from permafrost soils of Northeastern Siberia. *Glob Chang Biol* 12:2336–2351. <https://doi.org/10.1111/j.1365-2486.2006.01259.x>
- Schuur EAG, McGuire AD, Schädel C, Grosse G, Harden JW, Hayes DJ, Hugelius G, Koven CD, Kuhry P, Lawrence DM, Natali SM, Olefeldt D, Romanovsky VE, Schaefer K, Turetsky MR, Treat CC, Vonk JE. 2015. Climate change and the permafrost carbon feedback. *Nature* 520:171–179. <https://doi.org/10.1038/nature14338>
- Koven CD, Ringeval B, Friedlingstein P, Ciais P, Cadule P, Khvorostyanov D, Krinner G, Tarnocai C. 2011. Permafrost carbon-climate feedbacks accelerate global warming. *Proc Natl Acad Sci U S A* 108:14769–14774. <https://doi.org/10.1073/pnas.1103910108>
- Crowther TW, Todd-Brown KEO, Rowe CW, Wieder WR, Carey JC, Machmuller MB, Snoek BL, Fang S, Zhou G, Allison SD, Blair JM, Bridgman SD, Burton AJ, Carrillo Y, Reich PB, Clark JS, Classen AT, Dijkstra FA, Elberling B, Emmett BA, Estiarte M, Frey SD, Guo J, Harte J, Jiang L, Johnson BR, Kröel-Dulay G, Larsen KS, Laudon H, Lavallee JM, Luo Y, Lupascu M, Ma LN, Marhan S, Michelsen A, Mohan J, Niu S, Pendall E, Peñuelas J, Pfeifer-Meister L, Poll C, Reinsch S, Reynolds LL, Schmidt IK, Sistla S, Sokol NW, Templer PH, Treseder KK, Welker JM, Bradford MA. 2016. Quantifying global soil carbon losses in response to warming. *Nature* 540:104–108. <https://doi.org/10.1038/nature20150>
- Hicks Pries CE, Castanha C, Porras RC, Torn MS. 2017. The whole-soil carbon flux in response to warming. *Science* 355:1420–1423. <https://doi.org/10.1126/science.aal1319>
- Bond-Lamberty B, Bailey VL, Chen M, Gough CM, Vargas R. 2018. Globally rising soil heterotrophic respiration over recent decades. *Nature* 560:80–83. <https://doi.org/10.1038/s41586-018-0358-x>
- García-Palacios P, Crowther TW, Dacal M, Hartley IP, Reinsch S, Rinnan R, Rousk J, van den Hoogen J, Ye J-S, Bradford MA. 2021. Evidence for large microbial-mediated losses of soil carbon under anthropogenic warming. *Nat Rev Earth Environ* 2:507–517. <https://doi.org/10.1038/s43017-021-00178-4>
- Larson EJJ, Schiferl LD, Commann R, Munger JW, Trugman AT, Ise T, Euskirchen ES, Wofsy S, Moorcroft PM. 2022. The changing carbon balance of Tundra ecosystems: results from a vertically-resolved peatland biosphere model. *Environ Res Lett* 17:014019. <https://doi.org/10.1088/1748-9326/ac4070>
- Romero-Olivares AL, Allison SD, Treseder KK. 2017. Soil microbes and their response to experimental warming over time: a meta-analysis of field studies. *Soil Biol Biochem* 107:32–40. <https://doi.org/10.1016/j.soilbio.2016.12.026>
- van Gestel N, Shi Z, van Groenigen KJ, Osenberg CW, Andresen LC, Dukes JS, Hovenden MJ, Luo Y, Michelsen A, Pendall E, Reich PB, Schuur EAG, Hungate BA. 2018. Predicting soil carbon loss with warming. *Nature* 554:E4–E5. <https://doi.org/10.1038/nature25745>
- Bradford MA, Wieder WR, Bonan GB, Fierer N, Raymond PA, Crowther TW. 2016. Managing uncertainty in soil carbon feedbacks to climate change. *Nature Clim Change* 6:751–758. <https://doi.org/10.1038/nclimate3071>
- Jansson JK, Hofmockel KS. 2020. Soil microbiomes and climate change. *Nat Rev Microbiol* 18:35–46. <https://doi.org/10.1038/s41579-019-0265-7>
- Wieder WR, Sulman BN, Hartman MD, Koven CD, Bradford MA. 2019. Arctic soil governs whether climate change drives global losses or gains in soil carbon. *Geophys Res Lett* 46:14486–14495. <https://doi.org/10.1029/2019GL085543>
- Loranty MM, Abbott BW, Blok D, Douglas TA, Epstein HE, Forbes BC, Jones BM, Kholodov AL, Kropp H, Malhotra A, Mamet SD, Myers-Smith IH, Natali SM, O'Donnell JA, Phoenix GK, Rocha AV, Sonntag O, Tape KD, Walker DA. 2018. Reviews and syntheses: changing ecosystem influences on soil thermal regimes in northern high-latitude permafrost regions. *Biogeosciences* 15:5287–5313. <https://doi.org/10.5194/bg-15-5287-2018>
- DeAngelis KM, Roy Chowdhury P, Pold G, Romero-Olivares A, Frey S. 2019. Microbial responses to experimental soil warming: five testable

- hypotheses. *Ecosyst Consequences Soil Warm Microbes, Veg Fauna Soil Biogeochem*: 141–156
20. Hobbie JE, Kling GW. 2014. Alaska's changing Arctic: ecological consequences for tundra, streams, and lakes, p 303–324. Oxford University Press.
 21. Blume-Werry G, Milbau A, Teuber LM, Johansson M, Dorrepaal E. 2019. Dwelling in the deep—strongly increased root growth and rooting depth enhance plant interactions with thawing permafrost soil. *New Phytol* 223:1328–1339. <https://doi.org/10.1111/nph.15903>
 22. Shaver GR, Laundre JA, Bret-Harte MS, Chapin FS, Mercado-Díaz JA, Giblin AE, Gough L, Gould WA. 2014. Terrestrial ecosystems at Toolik Lake, Alaska. Alaska's changing arctic: ecological consequences for tundra, streams, and lakes, Oxford Univ. Press: New York, : 90–142. <https://doi.org/10.1021/pr060161n>
 23. Myers-Smith IH, Kerby JT, Phoenix GK, Bjerke JW, Epstein HE, Assmann JJ, John C, Andreu-Hayles L, Angers-Blondin S, Beck PSA, Berner LT, Bhatt US, Bjorkman AD, Blok D, Bryn A, Christiansen CT, Cornelissen JHC, Cunliffe AM, Elmendorf SC, Forbes BC, Goetz SJ, Hollister RD, de Jong R, Lorant MM, Macias-Fauria M, Maseyk K, Normand S, Olofsson J, Parker TC, Parmentier F-J, Post E, Schaepman-Strub G, Stordal F, Sullivan PF, Thomas HJD, Tømmervik H, Treharne R, Tweedie CE, Walker DA, Wilkening M, Wipf S. 2020. Complexity revealed in the greening of the Arctic. *Nat Clim Chang* 10:106–117. <https://doi.org/10.1038/s41558-019-0688-1>
 24. Doetterl S, Alexander J, Fior S, Frossard A, Magnabosco C, Broek M, Westergaard KB. 2022. Will accelerated soil development be a driver of Arctic greening in the late 21st century?. *J Plant Nutr Soil Sci* 185:19–23. <https://doi.org/10.1002/jpln.202100334>
 25. Sistla SA, Moore JC, Simpson RT, Gough L, Shaver GR, Schimel JP. 2013. Long-term warming restructures Arctic tundra without changing net soil carbon storage. *Nature* 497:615–618. <https://doi.org/10.1038/nature12129>
 26. Deslippe JR, Hartmann M, Simard SW, Mohn WW. 2012. Long-term warming alters the composition of Arctic soil microbial communities. *FEMS Microbiol Ecol* 82:303–315. <https://doi.org/10.1111/j.1574-6941.2012.01350.x>
 27. Liu J, Yuan F, Zuo Y, Zhou R, Zhu X, Li K, Wang N, Chen N, Guo Z, Zhang L, Sun Y, Guo Y, Song C, Xu X. 2022. Warming-induced vegetation growth cancels out soil carbon-climate feedback in the Northern Asian permafrost region in the 21st century. *Environ Res Lett* 17:084009. <https://doi.org/10.1088/1748-9326/ac7eda>
 28. Martin AC, Macias-Fauria M, Bonsall MB, Forbes BC, Zetterberg P, Jeffers ES. 2022. Common mechanisms explain nitrogen-dependent growth of Arctic shrubs over three decades despite heterogeneous trends and declines in soil nitrogen availability. *New Phytol* 233:670–686. <https://doi.org/10.1111/nph.17529>
 29. Graham EB, Knelman JE, Schindlbacher A, Siciliano S, Breulmann M, Yannarell A, Beman JM, Abell G, Philippot L, Prosser J, Foulquier A, Yuste JC, Granville HC, Jones DL, Angel R, Salminen J, Newton RJ, Bürgmann H, Ingram LJ, Hamer U, Siljanen HMP, Peltoniemi K, Pothast K, Bañeras L, Hartmann M, Banerjee S, Yu R-Q, Nogaro G, Richter A, Koranda M, Castle SC, Goberna M, Song B, Chatterjee A, Nunes OC, Lopes AR, Cao Y, Kaisermann A, Hallin S, Strickland MS, Garcia-Pausas J, Barba J, Kang H, Isobe K, Papaspyrou S, Pastorelli R, Lagomarsino A, Lindström ES, Basilikou N, Nemergut DR. 2016. Microbes as engines of ecosystem function: when does community structure enhance predictions of ecosystem processes? *Front Microbiol* 7:214. <https://doi.org/10.3389/fmicb.2016.00214>
 30. Bardgett RD, van der Putten WH. 2014. Belowground biodiversity and ecosystem functioning. *Nature* 515:505–511. <https://doi.org/10.1038/nature13855>
 31. Rijkers R, Rousk J, Aerts R, Sigurdsson BD, Weedon JT. 2022. Optimal growth temperature of Arctic soil bacterial communities increases under experimental warming. *Glob Chang Biol* 28:6050–6064. <https://doi.org/10.1111/gcb.16342>
 32. Lehmann J, Kleber M. 2015. The contentious nature of soil organic matter. *Nature* 528:60–68. <https://doi.org/10.1038/nature16069>
 33. Delgado-Baquerizo M, Oliverio AM, Brewer TE, Benavent-González A, Eldridge DJ, Bardgett RD, Maestre FT, Singh BK, Fierer N. 2018. A global atlas of the dominant bacteria found in soil. *Science* 359:320–325. <https://doi.org/10.1126/science.aap9516>
 34. Placella SA, Brodie EL, Firestone MK. 2012. Rainfall-induced carbon dioxide pulses result from sequential resuscitation of phylogenetically clustered microbial groups. *Proc Natl Acad Sci U S A* 109:10931–10936. <https://doi.org/10.1073/pnas.1204306109>
 35. Leff JW, Bardgett RD, Wilkinson A, Jackson BG, Pritchard WJ, De Long JR, Oakley S, Mason KE, Ostle NJ, Johnson D, Baggs EM, Fierer N. 2018. Predicting the structure of soil communities from plant community taxonomy, phylogeny, and traits. *ISME J* 12:1794–1805. <https://doi.org/10.1038/s41396-018-0089-x>
 36. Fierer N, Bradford MA, Jackson RB. 2007. Toward an ecological classification of soil bacteria. *Ecology* 88:1354–1364. <https://doi.org/10.1890/05-1839>
 37. Malard LA, Anwar MZ, Jacobsen CS, Pearce DA. 2019. Biogeographical patterns in soil bacterial communities across the Arctic region. *FEMS Microbiol Ecol* 95:fiz128. <https://doi.org/10.1093/femsec/fiz128>
 38. Berlemont R, Martiny AC. 2015. Genomic potential for polysaccharide deconstruction in bacteria. *Appl Environ Microbiol* 81:1513–1519. <https://doi.org/10.1128/AEM.03718-14>
 39. Berlemont R, Martiny AC. 2016. Glycoside hydrolases across environmental microbial communities. *PLOS Comput Biol* 12: e1005300. <https://doi.org/10.1371/journal.pcbi.1005300>
 40. Zimmerman AE, Martiny AC, Allison SD. 2013. Microdiversity of extracellular enzyme genes among sequenced prokaryotic genomes. *ISME J* 7:1187–1199. <https://doi.org/10.1038/ismej.2012.176>
 41. Bryson S, Li Z, Pett-Ridge J, Hettich RL, Mayali X, Pan C, Mueller RS, VerBerkmoes N. 2016. Proteomic stable isotope probing reveals taxonomically distinct patterns in amino acid assimilation by coastal marine bacterioplankton. *mSystems* 1. <https://doi.org/10.1128/mSystems.00027-15>
 42. Bryson S, Li Z, Chavez F, Weber PK, Pett-Ridge J, Hettich RL, Pan C, Mayali X, Mueller RS. 2017. Phylogenetically conserved resource partitioning in the coastal microbial loop. *ISME J* 11:2781–2792. <https://doi.org/10.1038/ismej.2017.128>
 43. Zhálnina K, Louie KB, Hao Z, Mansoori N, da Rocha UN, Shi S, Cho H, Karaoz U, Loqué D, Bowen BP, Firestone MK, Northen TR, Brodie EL. 2018. Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nat Microbiol* 3:470–480. <https://doi.org/10.1038/s41564-018-0129-3>
 44. Baran R, Brodie EL, Mayberry-Lewis J, Hummel E, Da Rocha UN, Chakraborty R, Bowen BP, Karaoz U, Cadillo-Quiroz H, Garcia-Pichel F, Northen TR. 2015. Exometabolite niche partitioning among sympatric soil bacteria. *Nat Commun* 6:8289. <https://doi.org/10.1038/ncomms9289>
 45. Nuccio EE, Starr E, Karaoz U, Brodie EL, Zhou J, Tringe S, Malmstrom RR, Woyke T, Banfield JF, Firestone MK, Pett-Ridge J. 2019. Niche differentiation is spatially and temporally regulated in the rhizosphere. *Ecology*. <https://doi.org/10.1101/611863>
 46. Wilhelm RC, Singh R, Eltis LD, Mohn WW. 2019. Bacterial contributions to delignification and lignocellulose degradation in forest soils with metagenomic and quantitative stable isotope probing. *ISME J* 13:413–429. <https://doi.org/10.1038/s41396-018-0279-6>
 47. Pepe-Ranney C, Campbell AN, Koechli CN, Berthrong S, Buckley DH. 2016. Unearthing the ecology of soil microorganisms using a high resolution DNA-SIP approach to explore cellulose and xylose metabolism in soil. *Front Microbiol* 7: 703. <https://doi.org/10.3389/fmicb.2016.00703>
 48. Woodcroft BJ, Singleton CM, Boyd JA, Evans PN, Emerson JB, Zayed AAF, Hoelzle RD, Lamberton TO, McCalley CK, Hodgkins SB, Wilson RM, Purvine SO, Nicora CD, Li C, Froliking S, Chanton JP, Crill PM, Saleska SR, Rich V, Tyson GW. 2018. Genome-centric view of carbon processing in thawing permafrost. *Nature* 560:49–54. <https://doi.org/10.1038/s41586-018-0338-1>
 49. Butterfield CN, Li Z, Andeer PF, Spaulding S, Thomas BC, Singh A, Hettich RL, Suttle KB, Probst AJ, Tringe SG, Northen T, Pan C, Banfield JF. 2016. Proteogenomic analyses indicate bacterial methylotrophy and archaeal heterotrophy are prevalent below the grass root zone. *PeerJ* 4: e2687. <https://doi.org/10.7717/peerj.2687>
 50. Diamond S, Andeer PF, Li Z, Crits-Christoph A, Burstein D, Anantharaman K, Lane KR, Thomas BC, Pan C, Northen TR, Banfield JF. 2019. Mediterranean grassland soil C-N compound turnover is dependent on

- rainfall and depth, and is mediated by genomically divergent microorganisms. *Nat Microbiol* 4:1356–1367. <https://doi.org/10.1038/s41564-019-0449-y>
51. Heyer R, Schallert K, Zoun R, Becher B, Saake G, Benndorf D. 2017. Challenges and perspectives of metaproteomic data analysis. *J Biotechnol* 261:24–36. <https://doi.org/10.1016/j.jbiotec.2017.06.1201>
 52. Starke R, Jehmlich N, Bastida F. 2019. Using proteins to study how microbes contribute to soil ecosystem services: the current state and future perspectives of soil metaproteomics. *J Proteomics* 198:50–58. <https://doi.org/10.1016/j.jprot.2018.11.011>
 53. Walker MD, Walker DA, Auerbach NA. 1994. Plant communities of a tussock tundra landscape in the Brooks range foothills, Alaska. *J Veg Sci* 5:843–866. <https://doi.org/10.2307/3236198>
 54. Walker DA, Reynolds MK, Daniëls FJA, Einarsson E, Elvebak A, Gould WA, Katenin AE, Kholod SS, Markon CJ, Melnikov ES, Moskalenko NG, Talbot SS, Yurtsev BA. 2005. The Circumpolar Arctic vegetation map. *J Veg Sci* 16:267–282. <https://doi.org/10.1111/j.1654-1103.2005.tb02365.x>
 55. Iversen CM, Sloan VL, Sullivan PF, Euskirchen ES, McGuire AD, Norby RJ, Walker AP, Warren JM, Wullschlegel SD. 2015. The unseen iceberg: plant roots in Arctic tundra. *New Phytol* 205:34–58. <https://doi.org/10.1111/nph.13003>
 56. Shaver GR, Cutler JC. 1979. The vertical distribution of live vascular phytomass in cottongrass tussock tundra. *AAAR* 11:335. <https://doi.org/10.2307/1550421>
 57. Shaver GR, Fetcher N, Chapin FS. 1986. Growth and flowering in *Eriophorum vaginatum*: annual and latitudinal variation. *Ecology* 67:1524–1535. <https://doi.org/10.2307/1939083>
 58. Wang P, Mommer L, van Ruijven J, Berendse F, Maximov TC, Heijmans MMPD. 2016. Seasonal changes and vertical distribution of root standing biomass of graminoids and shrubs at a Siberian tundra site. *Plant Soil* 407:55–65. <https://doi.org/10.1007/s11104-016-2858-5>
 59. Mercado-Díaz JA, Gould WA, González G. 2014. Soil nutrients, landscape age, and Sphagno-Eriophoretum vaginatum plant communities in Arctic moist-acidic tundra landscapes. *OJSS* 04:375–387. <https://doi.org/10.4236/ojss.2014.411038>
 60. Bockheim JG, Walker DA, Everett LR, Nelson FE, Shiklomanov NI, Bockheim JG. 1998. Soils and cryoturbation in moist nonacidic and acidic tundra in the Kuparuk River Basin, Arctic Alaska, U.S.A. *AAAR* 30:166. <https://doi.org/10.2307/1552131>
 61. Environmental Data Center Team, n.d. Meteorological monitoring program at Toolik, Alaska.
 62. Chourey K, Jansson J, VerBerkmoes N, Shah M, Chavarria KL, Tom LM, Brodie EL, Hettich RL. 2010. Direct cellular lysis/protein extraction protocol for soil metaproteomics. *J Proteome Res* 9:6615–6622. <https://doi.org/10.1021/pr100787q>
 63. Erde J, Loo RRO, Loo JA. 2014. Enhanced FASP (eFASP) to increase proteome coverage and sample recovery for quantitative proteomic experiments. *J Proteome Res* 13:1885–1895. <https://doi.org/10.1021/pr4010019>
 64. Miller SE, Rizzo AI, Waldbauer JR. 2018. Postnovo: postprocessing enables accurate and FDR-controlled de novo peptide sequencing. *J Proteome Res* 17:3671–3680. <https://doi.org/10.1021/acs.jproteome.8b00278>
 65. Johnston ER, Rodriguez-R LM, Luo C, Yuan MM, Wu L, He Z, Schuur EAG, Luo Y, Tiedje JM, Zhou J, Konstantinidis KT. 2016. Metagenomics reveals pervasive bacterial populations and reduced community diversity across the Alaska Tundra ecosystem. *Front Microbiol* 7:579. <https://doi.org/10.3389/fmicb.2016.00579>
 66. Tang H, Li S, Ye Y. 2016. A graph-centric approach for metagenome-guided peptide and protein identification in metaproteomics. *PLoS Comput Biol* 12:e1005224. <https://doi.org/10.1371/journal.pcbi.1005224>
 67. Ward CP, Nalven SG, Crump BC, Kling GW, Cory RM. 2017. Photochemical alteration of organic carbon draining permafrost soils shifts microbial metabolic pathways and stimulates respiration. *Nat Commun* 8:772. <https://doi.org/10.1038/s41467-017-00759-2>
 68. Cox MP, Peterson DA, Biggs PJ. 2010. SolexaQA: at-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics* 11:1–6. <https://doi.org/10.1186/1471-2105-11-485>
 69. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:1–11. <https://doi.org/10.1186/1471-2105-11-119>
 70. Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28:3150–3152. <https://doi.org/10.1093/bioinformatics/bts565>
 71. Li D, Liu CM, Luo R, Sadakane K, Lam TW. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31:1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>
 72. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>
 73. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7: e7359. <https://doi.org/10.7717/peerj.7359>
 74. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>
 75. Federhen S. 2012. The NCBI taxonomy database. *Nucleic Acids Res* 40:D136–D143. <https://doi.org/10.1093/nar/gkr1178>
 76. Eren AM, Kiehl E, Shaiber A, Veseli I, Miller SE, Schechter MS, Fink I, Pan JN, Yousef M, Fogarty EC, Trigodet F, Watson AR, Esen ÖC, Moore RM, Clayssen Q, Lee MD, Kivenson V, Graham ED, Merrill BD, Karkman A, Blankenberg D, Eppley JM, Sjödin A, Scott JJ, Vázquez-Campos X, McKay LJ, McDaniel EA, Stevens SLR, Anderson RE, Fuessel J, Fernandez-Guerra A, Maignien L, Delmont TO, Willis AD. 2021. Community-led, integrated, reproducible multi-Omics with AnviO. *Nat Microbiol* 6:3–6. <https://doi.org/10.1038/s41564-020-00834-3>
 77. Kim S, Pevzner PA. 2014. MS-GF+ makes progress towards a universal database search tool for proteomics. *Nat Commun* 5:1–10. <https://doi.org/10.1038/ncomms6277>
 78. Elias JE, Gygi SP. 2007. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* 4:207–214. <https://doi.org/10.1038/nmeth1019>
 79. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
 80. O’Leary NA, Wright MW, Brister JR, Ciuffo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, Astashyn A, Badredin A, Bao Y, Blinkova O, Brover V, Chetvernin V, Choi J, Cox E, Ermolaeva O, Farrell CM, Goldfarb T, Gupta T, Haft D, Hatcher E, Hlavina W, Joardar VS, Kodali VK, Li W, Maglott D, Masterson P, McGarvey KM, Murphy MR, O’Neill K, Pujar S, Rangwala SH, Rausch D, Riddick LD, Schoch C, Shkeda A, Storz SS, Sun H, Thibaud-Nissen F, Tolstoy I, Tully RE, Vatsan AR, Wallin C, Webb D, Wu W, Landrum MJ, Kimchi A, Tatusova T, DiCuccio M, Kitts P, Murphy TD, Pruitt KD. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 44:D733–D745. <https://doi.org/10.1093/nar/gkv1189>
 81. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, Bork P. 2017. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol Biol Evol* 34:2115–2122. <https://doi.org/10.1093/molbev/msx148>
 82. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, von Mering C, Bork P. 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* 47:D309–D314. <https://doi.org/10.1093/nar/gky1085>
 83. Muth T, Behne A, Heyer R, Kohrs F, Benndorf D, Hoffmann M, Lehtevä M, Reichl U, Martens L, Rapp E. 2015. The MetaProteomeAnalyzer: a powerful open-source software suite for metaproteomics data analysis and interpretation. *J Proteome Res* 14:1557–1565. <https://doi.org/10.1021/pr501246w>
 84. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42:D490–D495. <https://doi.org/10.1093/nar/gkt1178>
 85. Zybailov B, Mosley AL, Sardiou ME, Coleman MK, Florens L, Washburn MP. 2006. Statistical analysis of membrane proteome expression changes in

- Saccharomyces cerevisiae. *J Proteome Res* 5:2339–2347. <https://doi.org/10.1021/pr060161n>
86. Johnston ER, Hatt JK, He Z, Wu L, Guo X, Luo Y, Schuur EAG, Tiedje JM, Zhou J, Konstantinidis KT. 2019. Responses of tundra soil microbial communities to half a decade of experimental warming at two critical depths. *Proc Natl Acad Sci U S A* 116:15096–15105. <https://doi.org/10.1073/pnas.1901307116>
 87. Wallenstein MD, McMahon S, Schimel J. 2007. Bacterial and fungal community structure in Arctic tundra tussock and shrub soils. *FEMS Microbiol Ecol* 59:428–435. <https://doi.org/10.1111/j.1574-6941.2006.00260.x>
 88. Fierer N, Leff JW, Adams BJ, Nielsen UN, Bates ST, Lauber CL, Owens S, Gilbert JA, Wall DH, Caporaso JG. 2012. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc Natl Acad Sci U S A* 109:21390–21395. <https://doi.org/10.1073/pnas.1215210110>
 89. Koyama A, Wallenstein MD, Simpson RT, Moore JC. 2014. Soil bacterial community composition altered by increased nutrient availability in Arctic tundra soils. *Front Microbiol* 5: 516. <https://doi.org/10.3389/fmicb.2014.00516>
 90. Campbell BJ, Polson SW, Hanson TE, Mack MC, Schuur EAG. 2010. The effect of nutrient deposition on bacterial communities in Arctic tundra soil. *Environ Microbiol* 12:1842–1854. <https://doi.org/10.1111/j.1462-2920.2010.02189.x>
 91. Crump BC, Amaral-Zettler LA, Kling GW. 2012. Microbial diversity in Arctic freshwaters is structured by inoculation of microbes from soils. *ISME J* 6:1629–1639. <https://doi.org/10.1038/ismej.2012.9>
 92. Ye R, Jin Q, Bohannan B, Keller JK, McAllister SA, Bridgham SD. 2012. 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. <https://doi.org/10.1016/j.soilbio.2012.05.015>
 93. Emerson D, Scott JJ, Benes J, Bowden WB, Löffler FE. 2015. Microbial iron oxidation in the Arctic tundra and its implications for biogeochemical cycling. *Appl Environ Microbiol* 81:8066–8075. <https://doi.org/10.1128/AEM.02832-15>
 94. Reji L, Zhang X, Graham EB. 2022. Genome-resolved metagenomics informs the functional ecology of uncultured Acidobacteria in redox oscillated Sphagnum peat. *mSystems* 7. <https://doi.org/10.1128/msystems.00055-22>
 95. Kielak AM, Barreto CC, Kowalchuk GA, van Veen JA, Kuramae EE. 2016. The ecology of Acidobacteria: moving beyond genes and genomes. *Front Microbiol* 7: 744. <https://doi.org/10.3389/fmicb.2016.00744>
 96. Tveit AT, Urich T, Svenning MM. 2014. Metatranscriptomic analysis of Arctic peat soil microbiota. *Appl Environ Microbiol* 80:5761–5772. <https://doi.org/10.1128/AEM.01030-14>
 97. Kuzyakov Y, Blagodatskaya E. 2015. Microbial hotspots and hot moments in soil: concept & review. *Soil Biology and Biochemistry* 83:184–199. <https://doi.org/10.1016/j.soilbio.2015.01.025>
 98. Canarini A, Kaiser C, Merchant A, Richter A, Wanek W. 2019. Root exudation of primary metabolites: mechanisms and their roles in plant responses to environmental stimuli. *Front Plant Sci* 10:157. <https://doi.org/10.3389/fpls.2019.00157>
 99. Danhorn T, Fuqua C. 2007. Biofilm formation by plant-associated bacteria. *Annu Rev Microbiol* 61:401–422. <https://doi.org/10.1146/annurev.micro.61.080706.093316>
 100. Heavner ME, Qiu W-G, Cheng H-P, Mergaert P. 2015. Phylogenetic co-occurrence of ExoR, ExoS, and ChvI, components of the RSI bacterial invasion switch, suggests a key adaptive mechanism regulating the transition between free-living and host-invading phases in Rhizobiales. *PLoS ONE* 10:e0135655. <https://doi.org/10.1371/journal.pone.0135655>
 101. King GM, Weber CF. 2007. Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nat Rev Microbiol* 5:107–118. <https://doi.org/10.1038/nrmicro1595>
 102. King GM, Crosby H. 2002. Impacts of plant roots on soil CO cycling and soil-atmosphere CO exchange. *Glob Chang Biol* 8:1085–1093. <https://doi.org/10.1046/j.1365-2486.2002.00545.x>
 103. Cordero PRF, Bayly K, Man Leung P, Huang C, Islam ZF, Schittenhelm RB, King GM, Greening C. 2019. Atmospheric carbon monoxide oxidation is a widespread mechanism supporting microbial survival. *ISME J* 13:2868–2881. <https://doi.org/10.1038/s41396-019-0479-8>
 104. Ratcliff WC, Kadam SV, Denison RF. 2008. Poly-3-Hydroxybutyrate (PHB) supports survival and reproduction in starving rhizobia. *FEMS Microbiol Ecol* 65:391–399. <https://doi.org/10.1111/j.1574-6941.2008.00544.x>
 105. Weintraub MN, Schimel JP. 2005. The seasonal dynamics of amino acids and other nutrients in Alaskan Arctic tundra soils. *Biogeochemistry* 73:359–380. <https://doi.org/10.1007/s10533-004-0363-z>
 106. Wild B, Alves RJE, Bárta J, Čapek P, Gentsch N, Guggenberger G, Hugelius G, Knoltsch A, Kuhry P, Lashchinskiy N, Mikutta R, Palmtag J, Prommer J, Schnecker J, Shibistova O, Takriti M, Urich T, Richter A. 2018. Amino acid production exceeds plant nitrogen demand in Siberian tundra. *Environ Res Lett* 13:034002. <https://doi.org/10.1088/1748-9326/aaa4fa>
 107. Ramm E, Liu C, Ambus P, Butterbach-Bahl K, Hu B, Martikainen PJ, Marshchak ME, Mueller CW, Rennenberg H, Schloter M, Siljanen HMP, Voigt C, Werner C, Biasi C, Dannenmann M. 2022. A review of the importance of mineral nitrogen cycling in the plant-soil-microbe system of permafrost-affected soils—changing the paradigm. *Environ Res Lett* 17:013004. <https://doi.org/10.1088/1748-9326/ac417e>
 108. Kuzyakov Y, Xu X. 2013. Competition between roots and microorganisms for nitrogen: mechanisms and ecological relevance. *New Phytol* 198:656–669. <https://doi.org/10.1111/nph.12235>
 109. Fierer N, Lauber CL, Ramirez KS, Zaneveld J, Bradford MA, Knight R. 2012. Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. *ISME J* 6:1007–1017. <https://doi.org/10.1038/ismej.2011.159>
 110. Weintraub MN, Schimel JP. 2003. Interactions between carbon and nitrogen mineralization and soil organic matter chemistry in Arctic tundra soils. *Ecosystems* 6:129–143. <https://doi.org/10.1007/s10021-002-0124-6>
 111. Gentsch N, Mikutta R, Alves RJE, Bárta J, Čapek P, Gittel A, Hugelius G, Kuhry P, Lashchinskiy N, Palmtag J, Richter A, Šantrůčková H, Schnecker J, Shibistova O, Urich T, Wild B, Guggenberger G. 2015. Storage and transformation of organic matter fractions in cryoturbated permafrost soils across the Siberian Arctic. *Biogeosciences* 12:4525–4542. <https://doi.org/10.5194/bg-12-4525-2015>
 112. Sullivan PF, Stokes MC, McMillan CK, Weintraub MN. 2020. Labile carbon limits late winter microbial activity near Arctic treeline. *Nat Commun* 11:4024. <https://doi.org/10.1038/s41467-020-17790-5>