

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☐ ☒ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☐ ☒ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	No software was used
Data analysis	Formularity software provided by the Pacific Northwest National Laboratory, ftmsRanalysis R package (v1.1.0) , SPAdes (v3.12), UniteM (v0.0.18), DAS Tool (v1.1.1) , MetaWRAP (v1.0.6), CheckM (v1.0.12) , RefineM (v0.0.24) , MEGAHIT (v1.1.3), MetaBAT2 (v2.12.1) , GTDB toolkit (v1.5.1) , DRAM (1.4.0) , SingleM 111 (version 0.15.1), CoverM (v0.7.0), anvi'o (v.5.2), TranscriptM v0.3.1 with QC by KneadData (https://github.com/biobakery/biobakery/wiki/kneaddata), GraftM (v.0.15.0). R (v 4.1.0 and 4.2.1), picante (v1.8.2), rstatix (v0.7.2), vegan (v2.5-7), cluster (v 2.1.3), Hmisc (v5.0.1), igraph (v1.4.1), ggraph (v 2.1.0), tidygraph (v1.2.3), ggplot2 (v 3.4.2), patchwork (v.1.1.2) ggpubr (v.6.0) . Molecular Ecological Network Analyses pipeline (MENAP, available at the Institute for Environmental Genomics, University of Oklahoma http://ieg4.rccc.ou.edu/mena/main.cgi, no version. The greedy modularity optimization performed is part of the pipeline) METABOLIC pipeline version 4.0 All scripts for data processing and visualization are available on GitHub https://github.com/tfaily-lab/Metabolome_permafrost The following databases were used for annotation dbCAN (v11), KEGG/Kofam (release 104) as part of the DRAM pipeline (1.4.0)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The metagenomics and metatranscriptomics reads used in this manuscript are available at BioProject Accession: PRJNA386568 and previously published by 3.

The metagenomes assembled genomes used in this manuscript are available at Cronin, D., & NSF EMERGE Biology Integration Institute. (2023). Metagenome-assembled genomes (MAGs) from Stordalen Mire, Sweden (0.0.1-beta) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.10426238>. The annotation of the metagenome assembled genomes was performed using DRAM and can be found at: [10.5281/zenodo.7587534](https://doi.org/10.5281/zenodo.7587534)

FTICR-MS peak abundances for the extracts of the solid phase peat used in this study were described in reference 33 and are available from the EMERGE-DB at https://emerge-db.asc.ohio-state.edu/datasources/0141_Wilson-et-al-2022-STOTEN_ICR-plants.

Data of peat temperature, depth of each core, percentage of carbon and nitrogen content in the solid phase peat, and dissolved CO₂ and CH₄ concentrations used in this study was previously described in reference 91 and the Data is available at the EMERGE Database at https://emerge-db.asc.ohio-state.edu/datasources/0001_Coring2012, https://emerge-db.asc.ohio-state.edu/datasources/0006_GeochemPorewater20102012, and https://emerge-db.asc.ohio-state.edu/datasources/0005_GeochemSolid20102012.

Precipitation data used in this study was retrieved from Abisko Scientific Research Station, 2012. Meteorological data from Stordalen, Station ID 188790 (<https://www.smhi.se/data/meteorologi/ladda-ner-meteorologiska-observationer#param=airtemperatureInstant,stations=all,stationid=188790>).

(3) Woodcroft BJ, et al. Genome-centric view of carbon processing in thawing permafrost. Nature 560, 49-54 (2018).

(33) Wilson RM, et al. Plant organic matter inputs exert a strong control on soil organic matter decomposition in a thawing permafrost peatland. Sci Total Environ 820, 152757 (2022)

(91) Hodgkins SB. Changes in Organic Matter Chemistry and Methanogenesis Due to Permafrost Thaw in a Subarctic Peatland. Florida State University (2016).

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

This information is not applicable to this study thus it has not been collected

Reporting on race, ethnicity, or other socially relevant groupings

This information is not applicable to this study thus it has not been collected

Population characteristics

This information is not applicable to this study thus it has not been collected

Recruitment

This information is not applicable to this study thus it has not been collected

Ethics oversight

This information is not applicable to this study thus it has not been collected

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences

☐ Behavioural & social sciences

☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

Our study characterize microbial and metabolome composition from peat soil samples collected across a permafrost thaw gradient in Stordalen Mire, Sweden to assess environmental drivers and microbe-metabolite interactions. Samples were collected from distinct habitats representing a gradient of permafrost thaw: palsas (intact permafrost), bog (partially thawed permafrost), and fen (fully thawed permafrost). By sampling along this gradient, we could examine how microbiome composition and metabolite profiles differ in relation to varying permafrost states across the habitats. We applied microbial and metabolic null modeling, as well as random

matrix theory-based microbial-metabolite network analysis. The FT-ICR data was derived from 85 samples (palsa = 29, bog = 22 and fen = 29). Metagenomic analysis used an existing metagenome-assembled genomes database built from samples collected from 2010–2017 at the Stordalen Mire 39. Metatranscriptomics data, previously published by 3, was used to investigate the expression of the MAGs that significantly correlated with metabolome clusters and that were part of the microbe-metabolite networks.

(3) Woodcroft BJ, et al. Genome-centric view of carbon processing in thawing permafrost. *Nature* 560, 49–54 (2018).

(39) Cronin, D., & NSF EMERGE Biology Integration Institute. (2023). Metagenome-assembled genomes (MAGs) from Stordalen Mire, Sweden (0.0.1-beta) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.10426238>

Research sample	The sample unit derived from peat cores (11-cm diameter) collected between June and August in 2012 in a permafrost peat plateau located in Northern Sweden. Cores were divided into 3 depths (surface, middle and deep, specific depths are described in Supplementary table 2). Cored peat was divided for metagenomics and metabolomics analysis.
Sampling strategy	No a priori sample-size calculation was performed before sampling. Sample sizes are sufficient and all statistical tests included proper number of randomizations and p-values were subjected to multiple test corrections.
Data collection	Data collection of peat temperature, depth of each core, percentage of carbon and nitrogen content in the solid phase peat, and dissolved CO₂ and CH₄ concentrations used in this study was collected by Suzanne Hodgkins and previously described in reference 91. Data is available at the EMERGE Database at https://emerge-db.asc.ohio-state.edu/datasources/0001_Coring2012, https://emerge-db.asc.ohio-state.edu/datasources/0006_GeochemPorewater20102012, and https://emerge-db.asc.ohio-state.edu/datasources/0005_GeochemSolid20102012. Ultrahigh characterization of metabolites extracted from peat was collected by Malak Tfaily and Rachel Wilson (33) using a 12 Tesla Bruker Fourier Transform Ion Cyclotron Resonance (FTICR) mass spectrometer (Bruker, Solarix, Billerica, MA) located at the Pacific Northwest National Laboratory (PNNL). The metagenome assembled genome (MAG) database used in this study was created by Dylan Cronin (https://zenodo.org/records/10426238), reference 39. The metatranscriptomics data was previously published by 3. (3) Woodcroft BJ, et al. Genome-centric view of carbon processing in thawing permafrost. <i>Nature</i> 560, 49–54 (2018). (33) Wilson RM, et al. Plant organic matter inputs exert a strong control on soil organic matter decomposition in a thawing permafrost peatland. <i>Sci Total Environ</i> 820, 152757 (2022) (39) Cronin, D., & NSF EMERGE Biology Integration Institute. (2023). Metagenome-assembled genomes (MAGs) from Stordalen Mire, Sweden (0.0.1-beta) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10426238 (91) Hodgkins SB. Changes in Organic Matter Chemistry and Methanogenesis Due to Permafrost Thaw in a Subarctic Peatland. Florida State University (2016).
Timing and spatial scale	Peat samples were collected in June, July and August of 2012. Details about sampling date, time, GPS location, air and soil temperature, active layer depth, and notes related to the sampling can be found at the EMERGE database https://emerge-db.asc.ohio-state.edu/datasources/12 . The link to this information was provided in the Material and methods section and as the Data availability section.
Data exclusions	A total of 8 samples were excluded from the metabolomics data (available from the EMERGE-DB at https://emerge-db.asc.ohio-state.edu/datasources/0141_Wilson-et-al-2022-STOTEN_ICR-plants) because preliminary analysis demonstrated that these samples had a lower number of metabolite peaks compared with the rest.
Reproducibility	There were no attempts to reproduce the experiment. All statistical tests include randomization and multiple testing correction.
Randomization	Samples were not randomized into groups due to the nature of field collection.
Blinding	Given the nature of the study blinding was not considered relevant.
Did the study involve field work?	☒ Yes ☐ No

Field work, collection and transport

Field conditions	Peat samples were collected in June, July and August of 2012. Details about sampling date, time, GPS location, air and soil temperature, active layer depth, and notes related to the sampling can be found at the EMERGE database https://emerge-db.asc.ohio-state.edu/datasources/0001_Coring2012 . The link to this information was provided in the Material and methods section and as the Data availability section.
Location	The Stordalen Mire, a peat plateau characterized by discontinued permafrost is located southeast of the Abisko Scientific Research Station (ANS) in Northern Sweden (68.35 ° N, 19.05 ° E; 351 m above sea level).
Access & import/export	The field site at Stordalen Mire is under the jurisdiction of the Abisko Scientific Research Station, which requires applications and documentation of all sampling. Our team filed the standard application with them to sample peat and porewater
Disturbance	Disturbance was minimal, the peat coring process was designed to minimize disturbance by including accessing the site via and already-installed boardwalk, minimizing footsteps off the boardwalk and re-filling the core holes with any leftover material.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks

This information is not applicable to this study thus it has not been collected

Novel plant genotypes

This information is not applicable to this study thus it has not been collected

Authentication

This information is not applicable to this study thus it has not been collected