

SUPPLEMENTARY MATERIALS

Appendix A: Materials and Methods

The SPRUCE experiment is a climate change manipulation conducted in the Marcell experimental ombrotrophic bog forest in northern Minnesota (USA). The Deep Peat Heating (DPH) manipulation was a year-long deep soil heating at 5 warming levels (up to +9 °C), followed by the Whole Ecosystem Warming manipulation (WEW) that consists of air and belowground temperature treatment in open-topped chambers with elevated CO₂ treatments (+0 to +500 ppm) (Wilson *et al.*, 2016). The sampling for metagenomic analysis took place a year after the start of the DPH and WEW treatments. Metagenomic data are available for 10 experimental plots for the DPH treatment and 6 experimental plots for the WEW treatment at up to 4 depths per plot: 10, 40, 100, and 150 cm. In total, the metagenomic data of 53 samples were considered in this study, downloaded from the IMG/M website using the keywords 'Peatland microbial communities from SPRUCE experiment site at the Marcell Experimental Forest, Minnesota, USA'. Only samples with the term 're-annotation' in their names were considered. IMG accession numbers and related metadata are available in Supplementary Materials (Table S1).

Shotgun metagenomic sequencing data were used to calculate protein sequences. For each sample, the average amino acid composition of an overall protein at the microbial community level called hereafter the model protein, was calculated as the average of all inferred protein sequences weighted by the relative abundance of the protein-coding gene (Dick & Shock, 2011) (see Table S2 in Supplementary Materials). The amino acid composition of the nominally uncharged model proteins was used to calculate their average carbon oxidation state, determined by the covalent bonding structure of carbon as follows:

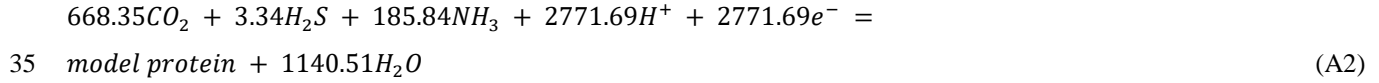
$$Z_c = \frac{2(n_O+n_S)+3n_N-n_H}{n_C} \quad (\text{A1})$$

where n is the number of atoms of C, H, N, O, and S. The average amino acid composition of model proteins and their Z_c values are available in Supplementary Materials (Table S2).

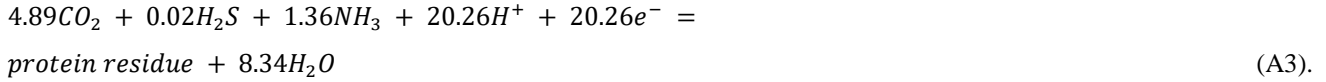
The *standard* molal thermodynamic properties of model proteins were calculated from the sequence of amino acids using a group additivity method (Dick *et al.*, 2006; Dick, 2008) (see Table S3 in Supplementary Materials). The method is restricted to unfolded aqueous proteins but the energy change of the folding reaction is negligible (Dick & Shock, 2011). Calculations were performed with the CHNOSZ package for R, which provides an integrated set of tools for thermodynamic calculations in aqueous geochemistry and biogeochemistry (Dick, 2008).

The *standard* molal Gibbs energy of the formation reaction of model proteins was calculated at a pressure of 1 bar and the annual average temperature of 3°C (Kolka *et al.*, 2011). The aqueous basis species CO₂, H₂O, H₂S, NH₃, H⁺ and e⁻ were considered; however, there is no thermodynamic restriction in the choice of basis species (Dick *et al.*, 2020). Because the model proteins of different samples have different lengths, the energy of the formation reaction cannot be directly compared in calculations of relative activity (Dick & Shock, 2011). To account for protein length, the *standard* molal Gibbs energy of formation reactions was normalized by the number of amino acid residues, i.e., the sequence lengths calculated with the

CHNOSZ package (Dick, 2008). Considering the formation reaction of the model protein at 10 cm depth in the experimental plot 6 after DPH treatment such as:



and the length, l , of the model protein of 136.81, the formation reaction of the protein residue is:



40 Deviations of environmental composition from standard composition were accounted for by considering the *actual* molal Gibbs energy change of the formation reaction of protein residues such as:

$$\Delta G_r = RT \ln \frac{Q}{K} \quad (A4)$$

where R is the gas constant, T is the absolute temperature in Kelvin, K is the equilibrium constant of the reaction, and Q is the reaction quotient as follows:

$$Q = \frac{a_{protein\ residue} a_{H_2O}^g}{a_{CO_2}^b a_{H_2S}^c a_{NH_3}^d a_{H^+}^e a_{e^-}^f} \quad (A5)$$

45 where a is the chemical activity of the reactants and products and $b, c, d, e, f,$ and g are the reaction coefficient of CO₂, H₂O, H₂S, NH₃, H⁺ and e⁻, respectively. Considering an equal-affinity state in which $\Delta G_r = 0$, the activity of each protein residue can be calculated from Equation (A5) as a function of the chemical composition of the environment such as:

$$a_{protein\ residu} = \frac{K_r a_{CO_2}^b a_{H_2S}^c a_{NH_3}^d a_{H^+}^e a_{e^-}^f}{a_{H_2O}^g} \quad (A6).$$

50 The activity of protein residues at thermodynamic equilibrium was calculated for a range of porewater compositions. The Eh value was limited to the stability domain of water. The pH value ranged from 3 to 6.6, which are the minimum and maximum pH values measured in the SPRUCE experiments. The activity of CO₂, H₂S, and NH₃ ranged from 10⁻⁹ to 10⁻², and that of H₂O from 10⁻¹ to 1. An annual average temperature of 3 °C was considered (Kolka *et al.*, 2011).

The activity of model proteins was calculated from that of its residue by dividing it by the protein length as follows:

$$a_{protein} = \frac{a_{residue}}{l_{protein}} \quad (A7).$$

55 Considering an Eh value of 101 mV, and logarithmic values for a_{H^+} , a_{CO_2} , a_{H_2S} , a_{NH_3} , and a_{H_2O} of -4, -3, -6, -9, and 0, respectively, the equilibrium activity of the model protein at 10 cm depth in the experimental plot 6 after DPH treatment is of 10^{-103.49}. This unrealistically low value reflects the low stability of the model protein in these relatively oxidizing conditions compared to the aqueous basis species CO₂. Applying the same calculation to the model protein at 150 cm depth indicates an activity of 10^{-104.29}, nearly ten times lower than the model protein at 10 cm depth.

60 To identify the least unstable proteins across the chemical domain considered, the relative equilibrium activities were calculated by normalizing the activity of each model protein by the sum of the activities of all model proteins as follows:

$$a_{protein}^* = \frac{a_{protein}}{\sum_{i=0}^n a_{protein}} \quad (A8).$$

In all experimental plots, the relative activities of model proteins at 10, 40, 100, and 150 cm depth, hereinafter referred to as target proteins, were compared with each other to assess the impact of the redox gradient. Despite this, normalization did not only consider this subset of 4 proteins, but also included the other 49 model proteins of the study, considered as background proteins. In this calculation, background proteins compete for formation with target proteins, which allows the chemical conditions of highest chemical activity of the target proteins to be resolved in finer details (Dick, 2022). Relative equilibrium activities were specifically calculated for chemical activities of aqueous basis species consistent with measured porewater composition. Speciation calculation was performed with the geochemical code Crunchflow (Steefel, 2009) on time series of porewater chemical composition from August 2013 to June 2016, i.e., the date of sampling for metagenomic analysis for the WEW treatment (Griffiths *et al.*, 2016, 2019). For comparison with the equilibrium activity of model proteins, the redox potential at equilibrium with the NO_3^- - N_2 , Fe^{2+} -goethite, SO_4^{2-} - H_2S , and CO_2 - CH_4 redox couples was calculated considering the Fe^{2+} and NO_3^- activity obtained from the speciation calculation, the N_2 activity at equilibrium with the atmosphere, equiactivity for the SO_4^{2-} - H_2S and CO_2 - CH_4 redox couples and the average measured pH value.

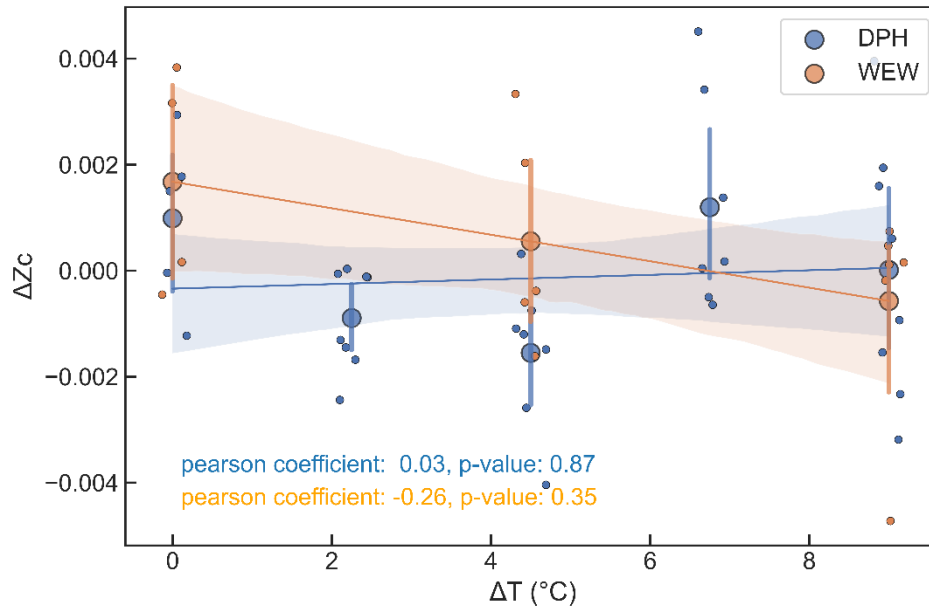
75

References

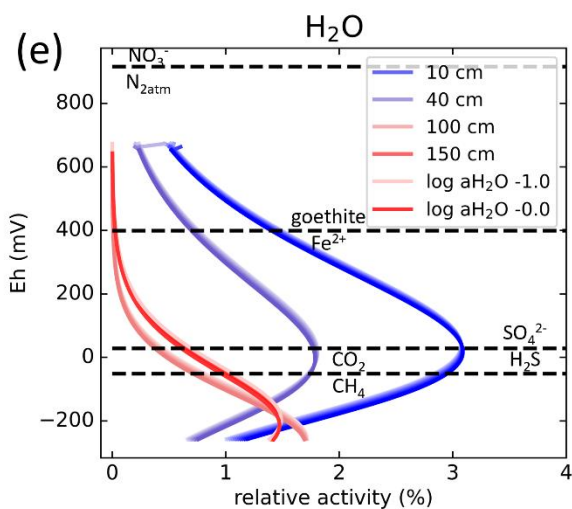
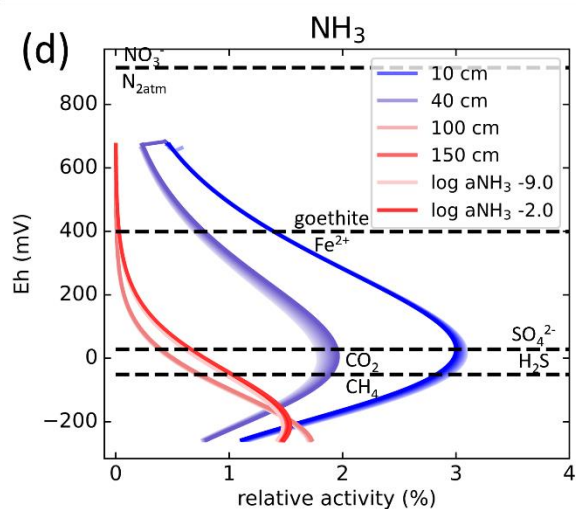
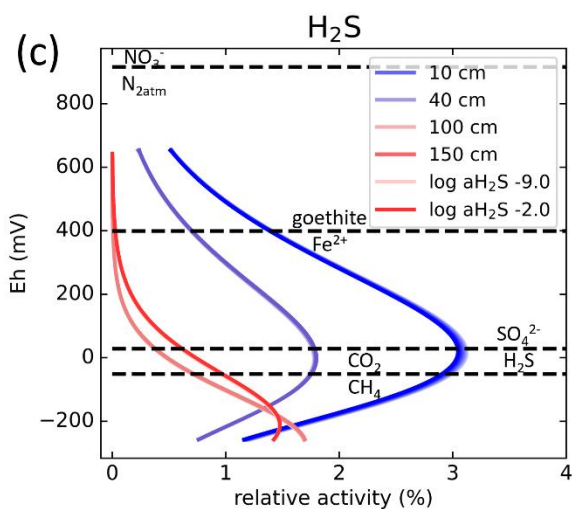
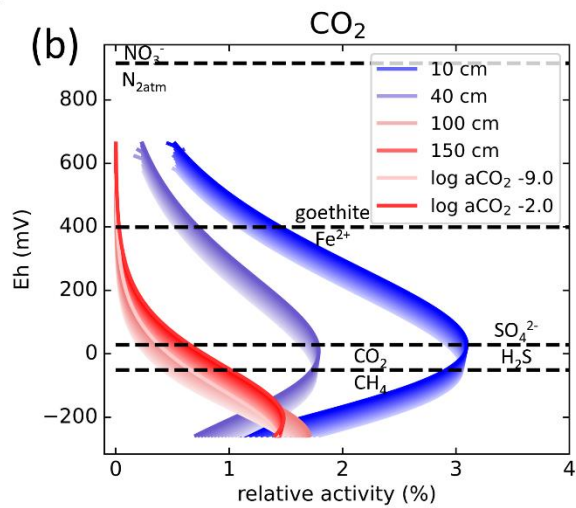
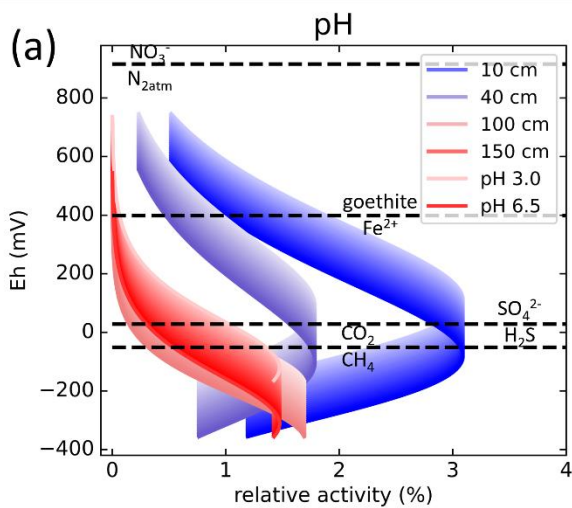
- Dick JM (2008) Calculation of the relative metastabilities of proteins using the CHNOSZ software package. *Geochemical Transactions* **9**, 1–17.
- Dick JM (2022) A thermodynamic model for water activity and redox potential in evolution and development. *Journal of Molecular Evolution* **90**, 182–199.
- Dick JM, LaRowe DE, Helgeson HC (2006) Temperature, pressure, and electrochemical constraints on protein speciation: Group additivity calculation of the standard molal thermodynamic properties of ionized unfolded proteins. *Biogeosciences* **3**, 311–336.
- Dick JM, Shock EL (2011) Calculation of the relative chemical stabilities of proteins as a function of temperature and redox chemistry in a hot spring. *PLoS One* **6**, e22782.
- Dick JM, Yu M, Tan J (2020) *Uncovering chemical signatures of salinity gradients through compositional analysis of protein sequences* (preprint). Biodiversity and Ecosystem Function: Microbial Ecology & Geomicrobiology.
- Griffiths N, Sebestyen S, Oleheiser KC, Stelling JM, Pierce CE, Nater EA, Toner BM, Kolka RK (2016) SPRUCE Porewater Chemistry Data for Experimental Plots, Beginning in 2013, <https://doi.org/10.3334/CDIAC/spruce.028>.
- Griffiths NA, Sebestyen SD, Oleheiser KC (2019) Variation in peatland porewater chemistry over time and space along a bog to fen gradient. *Science of the Total Environment* **697**, 134152.
- Kolka R, Sebestyen S, Verry ES, Brooks K (2011) *Peatland Biogeochemistry and Watershed Hydrology at the Marcell Experimental Forest*. CRC Press.
- Lin X, Tfaily MM, Steinweg JM, Chanton P, Esson K, Yang ZK, Chanton JP, Cooper W, Schadt CW, Kostka JE (2014) Microbial community stratification linked to utilization of carbohydrates and phosphorus limitation in a boreal peatland at Marcell Experimental Forest, Minnesota, USA. *Applied and environmental microbiology* **80**, 3518–3530.
- Steefel CI (2009) CrunchFlow. *Software for modeling multicomponent reactive flow and transport. User's manual* 12–91.
- Wilson RM, Hopple AM, Tfaily MM, Sebestyen SD, Schadt CW, Pfeifer-Meister L, Medvedeff C, McFarlane KJ, Kostka JE, Kolton M (2016) Stability of peatland carbon to rising temperatures. *Nature Communications* **7**, 13723.

100

Appendix B: Additional figures



105 **Figure B1.** Z_c change (ΔZ_c) as a function of warming level (ΔT) in the Deep Peat Heating (DPH – blue circles) and Whole Ecosystem
Warming (WEW – orange circles) manipulation of the SPRUCE experiment. ΔZ_c is calculated as the Z_c value of each model protein
minus the average Z_c values of all model proteins at a given depth. This normalization allows cancelling the influence of depth to
evaluate the impact of the warming level on the Z_c value. The small dots are individual model proteins. The larger circles and
110 associated error bars are the average ($\pm 2\sigma$) of model proteins for each warming level. The absolute value of the Pearson coefficient
below 0.2 and the p-value above 0.5 indicate no significant correlation between ΔZ_c and ΔT .



115 **Figure B2.** Relative equilibrium activity (%) of model proteins at 10 (blue), 40 (purple), 100 (pink), and 150 (red) cm depth in experimental plot 6 after DPH treatment as a function of redox potential and (a) pH values, (b) CO₂ activity, (c) H₂S activity, (d) NH₃ activity and (e) H₂O activity. The dark-to-light gradation represents pH values varying from 3 to 6.5 pH units (a), CO₂, H₂S, and NH₃ activity varying from 10⁻⁹ to 10⁻² (b, c, d), and H₂O activity varying from 0.1 to 1 (e).

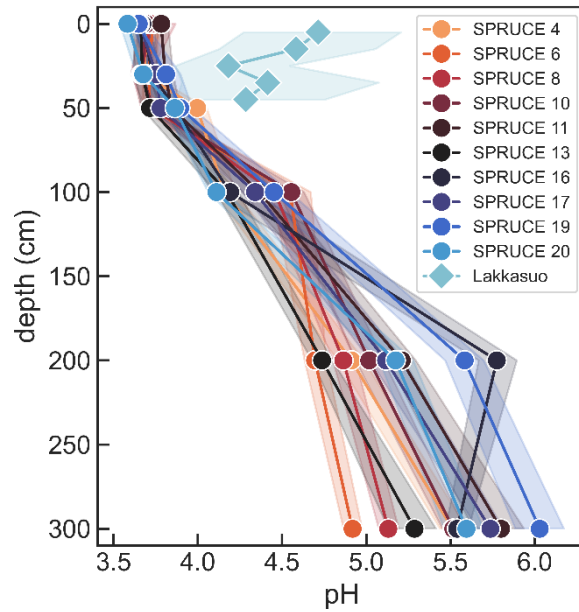


Figure B3. Average pH values as a function of depth in the SPRUCE experimental plots and at the Lakkasuo ombrotrophic peatland. For the SPRUCE experiments, only time series from August 2013 to June 2016 are considered for consistency with metagenomic sampling (Griffiths *et al.*, 2016). Error bars are $\pm\sigma$.

120

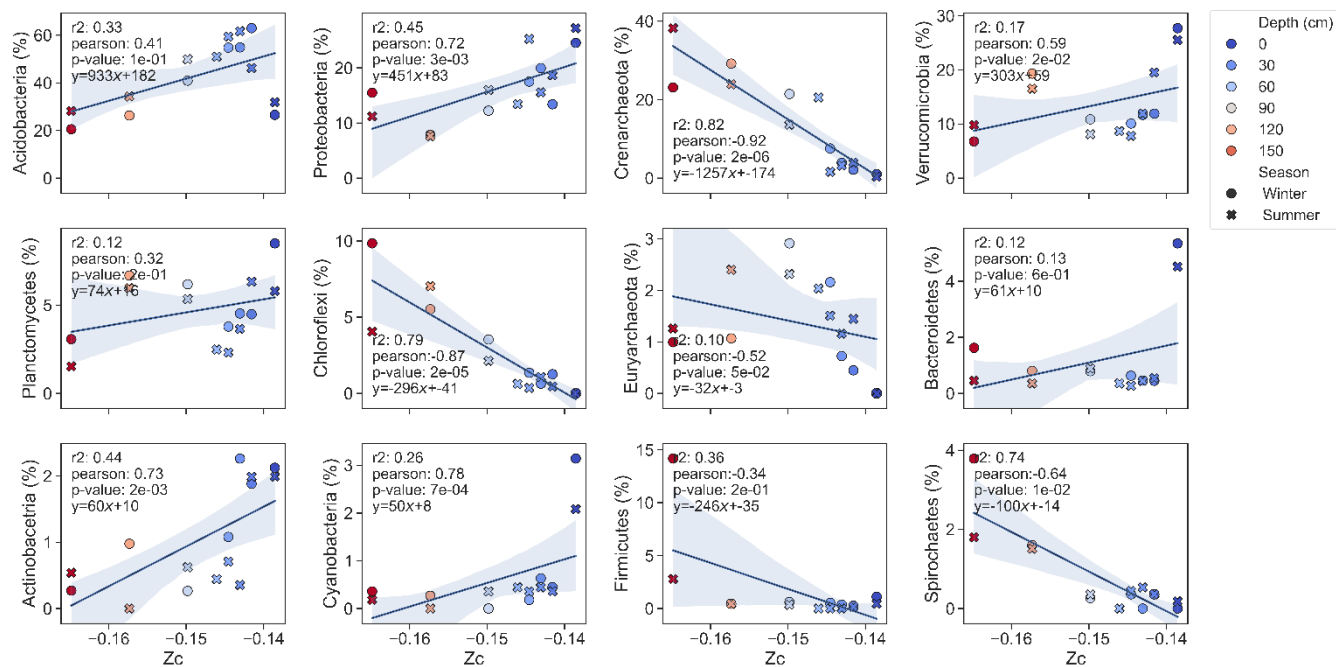


Figure B4. Relative taxonomic abundance (Lin *et al.*, 2014) as a function of Z_c values for different taxonomic groups. The Z_c value at the depth of measured taxonomic abundance was calculated using the correlation identified in Figure 1.