scientific data

Check for updates

OPEN

DATA DESCRIPTOR

A global database of soil microbial phospholipid fatty acids and enzyme activities

Laura G. van Galen *et al.*#

Soil microbes drive ecosystem function and play a critical role in how ecosystems respond to global change. Research surrounding soil microbial communities has rapidly increased in recent decades, and substantial data relating to phospholipid fatty acids (PLFAs) and potential enzyme activity have been collected and analysed. However, studies have mostly been restricted to local and regional scales, and their accuracy and usefulness are limited by the extent of accessible data. Here we aim to improve data availability by collating a global database of soil PLFA and potential enzyme activity measurements from 12,258 georeferenced samples located across all continents, 5.1% of which have not previously been published. The database contains data relating to 113 PLFAs and 26 enzyme activities, and includes metadata such as sampling date, sample depth, and soil pH, total carbon, and total nitrogen. This database will help researchers in conducting both global- and local-scale studies to better understand soil microbial biomass and function.

Background & Summary

Soil microbes, particularly bacteria and fungi, are critical for biogeochemical cycling and Earth's planetary health[1]. Our understanding of soil microbes has been developed primarily through local and regional studies that use methods such as DNA metabarcoding and metagenomics, phospholipid fatty acid (PLFA) extractions, and enzyme activity assays to study microbial community composition and their role in ecosystem processes[2–4]. Efforts to collate global soil DNA metabarcoding databases have vastly improved our understanding of below-ground microbial diversity patterns and species' distributions[5–8]. However, quantifying microbial biomass and function remains challenging based on DNA sequencing alone[9]. PLFA and enzyme activity assays have been performed by scientists for decades, providing, among other things, the ability to estimate carbon stored in microbial biomass and measure *in situ* microbial community functioning[2,10]. These methods complement emerging DNA technologies and are crucial for understanding the potential impact of global change factors on carbon storage and other critical biogeochemical processes[9–11].

PLFA assays involve measuring fatty acids associated with phospholipids of cell membranes[3,9,12]. PLFA molecules differ in factors such as fatty acid chain length, degree of saturation, branching, and functional group modifications[11]. To a certain degree, different PLFAs can be used as biomarkers of taxonomic groups such as fungi, protozoa, Gram-negative bacteria, and Gram-positive bacteria[10,11,13,14], making it possible to use changes in the fatty acid composition of PLFAs to estimate broad shifts in microbial communities. Phospholipids are degraded rapidly after cell death, allowing PLFA assays to target living organisms[15]. Additionally, conversion factors have been developed that allow PLFAs to be used to estimate microbial carbon content[10,16,17]. Although PLFAs cannot provide taxonomic resolution equivalent to other quantitative biomass estimation methods, such as quantitative PCR, PLFA-based biomass estimates can be more reliable in many situations[9,18]. Additionally, PLFA analysis is a relatively cost-effective way to measure soil microbial community biomass and composition[10] and has been widely employed since the early 1990s[3,19].

Soil enzyme activity assays measure the oxidative or hydrolytic catalysis of organic matter substrates by enzymes in soils, the majority of which are thought to be extracellular[20,21] and largely released by microbes but also by plant roots. Extracellular soil enzymes degrade organic polymers to liberate bioavailable forms of nutrients required for metabolism and growth[22,23]. These assays aim to quantify the maximum potential enzyme activity by incubating soil samples in the lab and, most commonly, colourimetrically or fluorometrically measuring

#A full list of authors and their affiliations appears at the end of the paper.

the release of a chromophore or fluorophore from the oxidation or hydrolysis of dye-conjugated substrates[2,22]. It is important to note that these assays measure the maximum potential activity rather than *in situ* activity, which is influenced by temperature, soil pH, substrate availability, and other soil and ecosystem properties[22,23]. Due to substrate specificity, different enzyme activities relate to the acquisition of different products categorizable by macromolecular type and nutrient element. For example, β-glucosidase degrades cellobioside, and phosphomonoesterases — often referred to by the broader enzyme class term of phosphatases — hydrolyse phosphate monoesters[24]. Similar to PLFA assays, soil enzyme assays have been performed for many decades[25].

Despite the long history of PLFA and enzyme activity assays, data access has largely been restricted to relatively small local and regional scales. Larger-scale global research networks and analyses are becoming increasingly important for complementing local-scale research to better understand and tackle Earth's global-scale environmental challenges[26]. Global-scale meta-analyses of microbial biomass carbon (MBC) have revealed that MBC shows complex biogeographical patterns and is highly sensitive to environmental change and human land use[27–30]. However, the underlying data used to estimate microbial biomass and function (i.e., the raw PLFA and enzyme measurements) are rarely available in meta-analyses, making it difficult to use these datasets to answer new and targeted research questions.

To improve the availability of soil microbial data, Smith, *et al.*[31] released an open call for collaboration to develop a global database of soil PLFA and potential enzyme activity measurements. Many scientists with data collected from 89 countries answered this call and here we provide access to the developed database. We provide a full description of the database, along with analyses assessing the coverage of environmental space, remaining data gaps, and potential biases that users should be aware of. This database will allow researchers to investigate critical questions at both local and global scales to better understand patterns of microbial biomass and function. We also hope that data gaps revealed in this database will inspire further research in data-limited regions so that geographical biases can be reduced in the future.

Methods

Following the open call for collaboration by Smith, *et al.*[31], georeferenced data from PLFA and enzyme assays of soil samples were provided by interested collaborators. Additional PLFA data were sourced from the United States National Ecological Observatory Network[32]. We also added data from several sources that reported individual PLFA measurements[33–37], as well as a recent study with a large enzyme dataset[38]. Where necessary, data were extracted from figures using DataThief[39]. We did not perform an additional exhaustive formal literature search because very few studies have reported measurements of individual PLFA biomarkers. Only samples with geographical coordinates were included. Data from experimental plots were excluded, as well as those from samples solely consisting of leaf litter.

Authors performed PLFA and enzyme activity assays using numerous well-recognised methods, with the cited methods used by each study listed in the database. Full sample collection and processing methods can be found in the original publications for previously-published samples (DOIs provided in the database). Methods for unpublished samples are included in this publication as Supplementary Information. The majority of PLFA assays were performed using variants of Bligh and Dyer[12] lipid extraction methods and gas chromatography-mass spectrometry, following Frostegård, *et al.*[19]. Several contributed datasets used ester-linked fatty acid measurements, following Schutter and Dick[40]. Although these two methods recover comparable compositional signals, ester-linked fatty acid measurements have concentrations approximately twice as high[41]. Therefore, we divided values from these samples in half to scale them appropriately. Enzyme activities were assayed using colourimetric and/or fluorometric methods[22], and in the case of urease, with the natural substrate (i.e., urea). Assays were assumed to be performed under optimal conditions of substrate according to best practices, and varied in assay incubation temperatures from a standardised temperatures or a temperature that reflected *in situ* conditions (e.g., mean annual temperature).

We thoroughly cleaned and standardised the database by first converting all PLFA units to nmol g^{-1} soil and enzyme activity units to nmol h^{-1} g^{-1} soil. We also checked all other variables and converted variable categories and units where needed. Sample depths listed as O horizon, A horizon, or "organic" were classed as 0–10 cm. Then, where possible, missing important data (e.g., enzyme reaction temperature) was obtained by re-contacting data contributors or examining publications. We examined the range of values in all variables to look for errors and outliers. A small number of samples contained negative enzyme activity values, which we replaced with zero, and percentages greater than 100, which we capped at 100. It was not possible to evaluate the precision or accuracy of provided sample coordinates. However, we used the "coordinateCleaner" R package v3.0.1[42] to ensure all coordinates correspond to the correct coordinate reference system (WGS84), and to flag any potential errors such as those with equal absolute latitude and longitude or those within a 100 m radius of country centroids or capitals. All included samples passed coordinate validity checks.

We included all PLFAs and enzymes in the final database for which, once the data sources were merged, there were at least 100 data points available. We extracted the continent, country, and biome[43] information for each sample location using the "terra" (v1.7–78), "sf" (v1.0-19), and "rnaturalearth" (v1.0.1) packages in R version 4.4.1[44–47].

Data Records

The database is available as a .xlsx file on Figshare[48]. The file includes tabs with the names of all PLFAs and enzymes (including Enzyme Commission numbers) for which data are available, and the number of samples available for each. The metadata, PLFA data, and enzyme data for each sample are provided in separate tabs which can be linked by the "sampleID" column. All column names are programming-language friendly.

In total, there are data for 12,258 soil samples from 3,743 unique locations (Fig. 1). There are 6,923 samples with PLFA data (for 113 PLFAs), 6,657 samples with enzyme activity data (for 26 enzyme groups), with 1,322

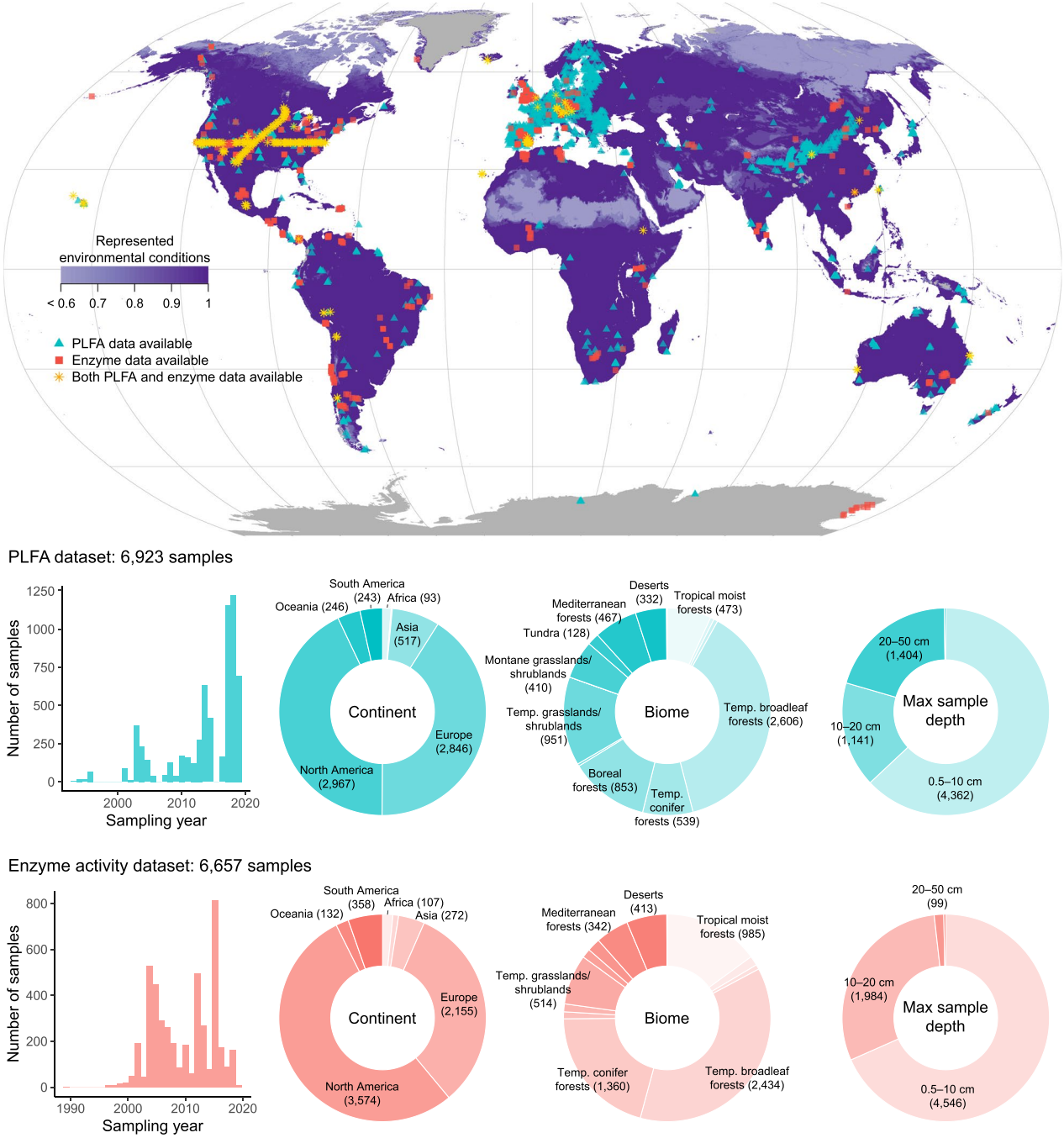

Fig. 1 The PLFA and enzyme activity database contains 12,258 samples from 3,743 locations. The purple colour gradient shows the degree to which climatic space is represented by the samples (scale of 0 to 1, see Technical Validation section). Grey regions do not have sufficient climate data available to evaluate climatic representation. The histogram and donut plots show number of samples (in parentheses) from the PLFA (blue) and enzyme (red) datasets collected in different years, from different continents, from different biomes[43], and within different categories of maximum sample depth. Only large segments of donut plots are labelled. Temp. = temperate.

samples containing both. Data from 627 samples (5.1%) have not previously been published. Samples were collected between 1989 and 2019 (Fig. 1). Of all samples, 10,436 (85%) are from natural ecosystems, and 1,822 (15%) are from managed (e.g., agriculture, plantation, urban) ecosystems. Data are predominantly from North America and Europe, but samples are available from all continents, including 70 samples from Antarctica (Fig. 1). Temperate and boreal biomes are the best represented, but many samples are also available from tropical moist broadleaf forests, particularly in the enzyme dataset. Montane grasslands, Mediterranean forests and deserts are

Name	Putative group	# samples	Metadata availability (%)								
			Depth	Year	Month	pH	C	N	Moisture	BD	Elev.
PLFAs											
c18:2ω6c	Fungi	6,612	100.0	89.2	85.5	46.5	4.0	29.8	2.4	1.3	32.8
cy17:0	GN bacteria	6,523	100.0	89.4	85.6	44.9	4.1	29.6	3.7	2.3	33.5
cy19:0		6,612	100.0	89.2	86.2	45.4	4.0	30.2	3.6	2.3	33.2
c16:1ω7c		5,600	100.0	88.7	87.8	45.9	4.7	31.8	2.9	2.7	36.1
c18:1ω7c		4,829	100.0	88.1	83.1	42.9	0.0	24.1	2.3	1.7	43.5
a15:0	GP bacteria	6,764	100.0	89.5	85.9	45.7	3.9	30.8	3.5	2.2	31.1
a17:0		4,962	100.0	89.6	84.7	44.4	4.3	27.4	0.6	3.1	43.7
i14:0		3,902	100.0	94.9	93.8	37.8	4.2	21.5	0.0	1.8	51.2
i15:0		6,883	100.0	89.7	86.1	47.0	3.9	32.0	3.5	2.2	32.5
i16:0		6,860	100.0	90.5	86.9	47.2	3.9	32.6	3.5	2.2	32.6
i17:0		6,665	100.0	89.6	85.9	45.3	4.0	30.7	3.6	2.3	31.4
c16:0–10Me	GP bacteria (act.)	5,401	100.0	97.8	93.4	46.1	3.1	31.2	0.6	1.9	41.4
c17:0–10Me		4,394	100.0	98.2	92.7	47.5	4.9	28.5	0.0	1.6	50.2
c18:0–10Me		5,694	100.0	87.6	83.4	39.1	2.9	24.8	0.6	1.8	39.3
All available PLFAs		6,923	100.0	89.7	86.2	46.8	3.8	32.3	3.5	2.2	32.3
Enzymes											
β-glucosidase	C acquisition	6,487	99.9	66.5	59.2	86.4	9.3	51.9	16.3	6.1	22.4
Cellobiohydrolase		2,638	100.0	75.6	75.3	89.2	0.0	56.4	23.2	11.2	4.1
Leucine aminopeptidase	N acquisition	3,278	99.8	76.9	62.4	92.6	17.4	59.8	30.5	6.3	20.9
N-acetylglucosaminidase		5,110	99.8	59.2	49.8	84.4	11.1	42.3	20.8	5.8	18.4
Acid phosphatase	P acquisition	4,445	100.0	85.3	74.7	90.8	12.4	67.4	23.9	8.9	12.2
All available enzymes		6,657	99.9	67.3	60.1	86.3	9.1	52.3	16.0	6.0	22.9

Table 1. Data available for 14 of the most commonly assessed PLFAs and five of the most commonly measured enzymes. Microbial groups and nutrient acquisition type classifications are putative only and specific to soil samples, based on information from Willers, *et al.*[11], Joergensen[57] and Dick[24]. GN = Gram-negative, GP = Gram-positive, act. differentiates Gram-positive bacteria that are actinobacteria. The number of samples available for each PLFA/enzyme and all available PLFAs/enzymes in the database is shown alongside the percentage of those samples for which metadata are available. C and N refer to soil carbon and nitrogen, moisture = gravimetric soil moisture content, BD = soil bulk density, and elev. = elevation.

also reasonably well represented. Most samples were collected with a maximum sampling depth between 0.5 and 20 cm, but the PLFA dataset also contains high representation of data from up to 50 cm deep (Fig. 1).

Fourteen of the most commonly assessed PLFAs and five of the most commonly measured enzymes are very well represented in the database, with between ~2,600 and ~6,800 samples available for each (Table 1). Data are also available for an additional 99 PLFAs and 21 enzyme categories. Metadata relating to sample depth are available for 100% of PLFA samples and 99.9% of enzyme samples, and data relating to sampling year, sampling month, and soil pH are available for 47–90% of PLFA samples and 60–86% of enzyme samples (Table 1). Metadata of other soil properties (carbon, nitrogen, moisture and bulk density) and elevation are also available for many samples (Table 1).

Technical Validation

We checked the database for erroneous outliers by calculating the interquartile range (IQR) of values within each biome for each PLFA/enzyme, and flagging values greater than 5 times the IQR. PLFA and enzyme data are often left-skewed, with our database being no exception, and many samples contained flagged values. We scanned the flagged values to look for patterns regarding the assay methods used or the study for which the samples were collected. No patterns were evident, and so all data were retained.

We assessed climatic space represented by the samples by assessing the degree of extrapolation in multi-dimensional space, following methods described in van den Hoogen, *et al.*[49]. Briefly, we extracted values for 19 bioclimatic layers from CHELSA[50] for each point location in the dataset, then transformed all values into principal component (PC) space. Next, we assessed whether each pixel value of the global bioclimatic layers fell within or outside convex hulls for each of the bivariate combinations from the first five principal components. These five PC axes collectively covered more than 90% of the sample space variation. We plotted the proportion of times that each pixel fell within the convex hulls on a map to evaluate the degree to which climate space is represented in the database (Fig. 1). As with many other large ecological datasets[51], global coverage of the data remains geographically and climatically uneven. In particular, substantial portions of Africa, South America, Asia, Antarctica, and ecosystems at high northern latitudes are currently under-sampled (Fig. 1).

PLFA contents and enzyme activities vary across biomes (Figs. 2, 3). Overall, PLFA values are relatively high in tropical conifer forests, tundra, and boreal forests (Fig. 2). Similarly, soil enzyme activities are high in tundra and boreal forest samples, but some enzyme activities are also high in soils from tropical dry forests and Mediterranean forests (Fig. 3).

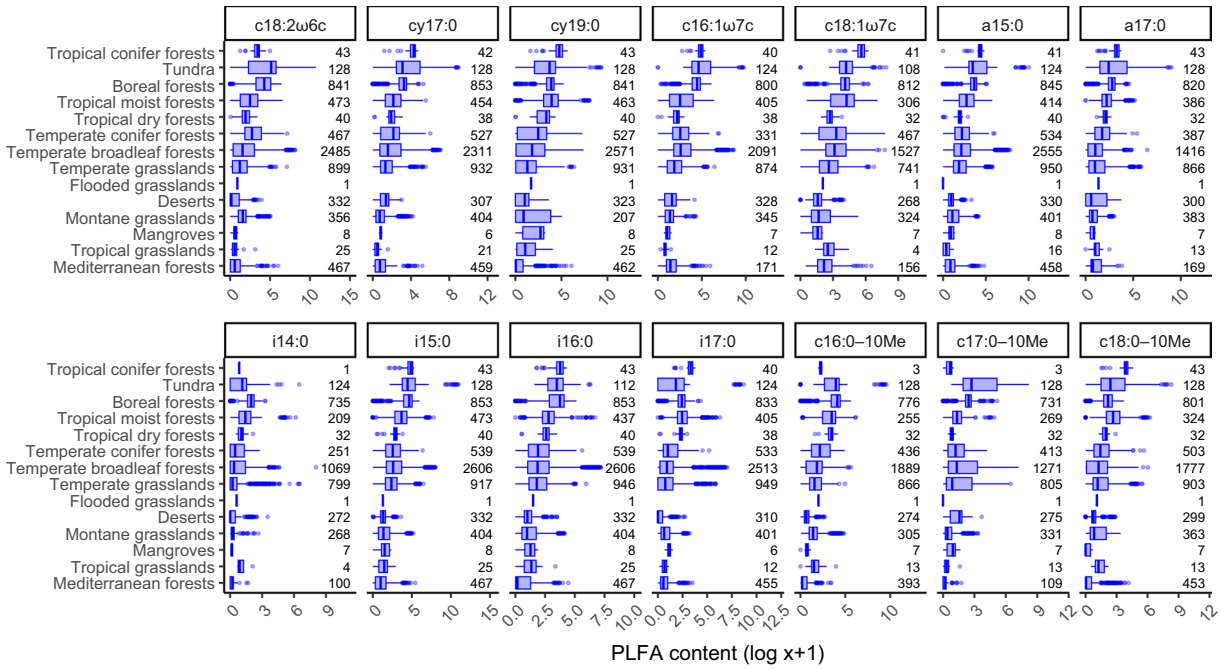

Fig. 2 Variation in PLFA content across biomes[43]. Boxes show median and interquartile range, with whiskers 1.5 times the interquartile range. Biomes are ordered according to the median value across all 14 PLFAs. Numbers show the sample size in each category.

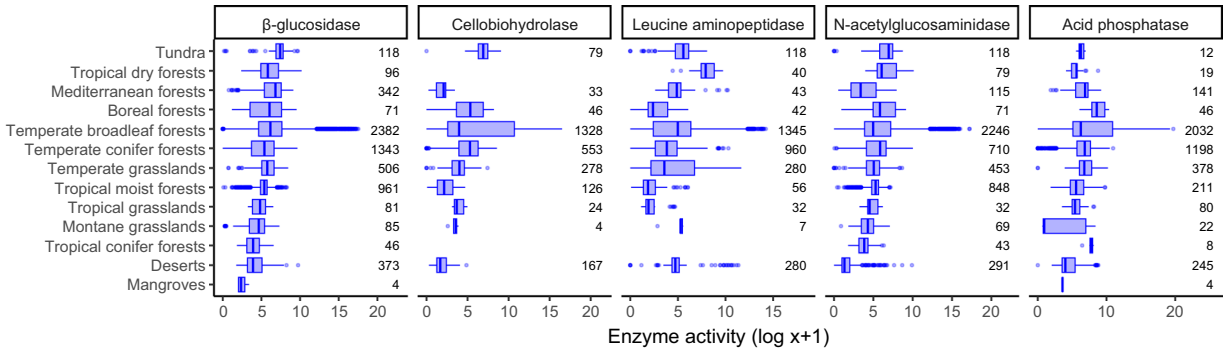

Fig. 3 Variation in soil enzyme activity across biomes[43]. Boxes show median and interquartile range, with whiskers 1.5 times the interquartile range. Biomes are ordered according to the median value across all five enzyme activities. Numbers show the sample size in each category.

Soil enzyme activities were measured using a variety of incubation temperatures ranging between 8 °C and 37 °C (Fig. 4). Assay incubation temperature data are available for 91% of samples (6,025). More details on the variation in assay temperature and its potential influence on activity rates are provided in the Usage Notes section.

Usage Notes

All samples are georeferenced, and so data at the pixel level relating to climate[50,52], soil properties[53], and land cover[54], for example, could be extracted from publicly-available global geospatial layers. Many samples contain field-collected metadata (Table 1), but for those that do not, available geospatial layers could also be used to fill in gaps. However, users should be aware that these geospatial layers are predictions associated with various uncertainties[55].

Because the database includes data collected over 30 years (1989 to 2019), there is potential to conduct time-series analyses. For example, some densely sampled countries contain samples spanning 6 to 25 years (Fig. 5). Additionally, 298 of the unique sampling locations (8%) contain data from more than one time point (unique month-year combinations), with the highest number of time points for a single location being 17. Time-series analyses are becoming increasingly important to track the response of organisms to global change factors, and such datasets are highly valuable[56].

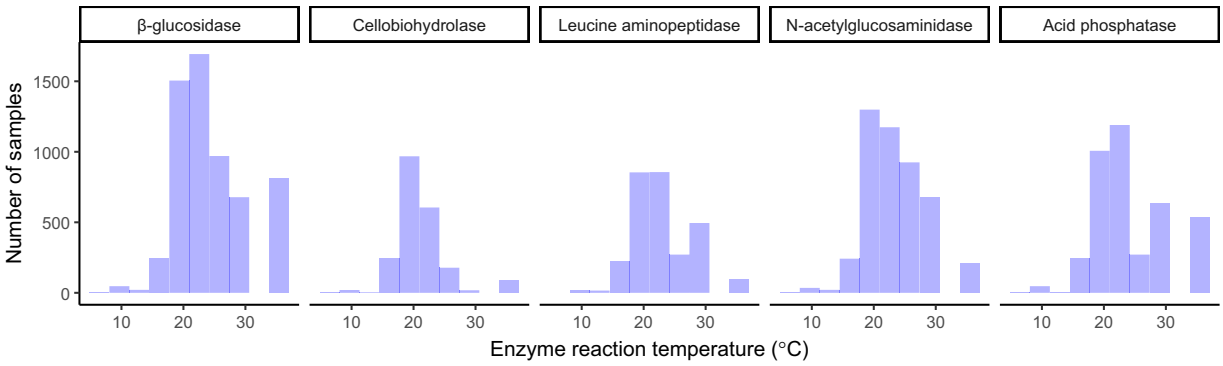

Fig. 4 The distribution of temperatures used to assay enzyme activity.

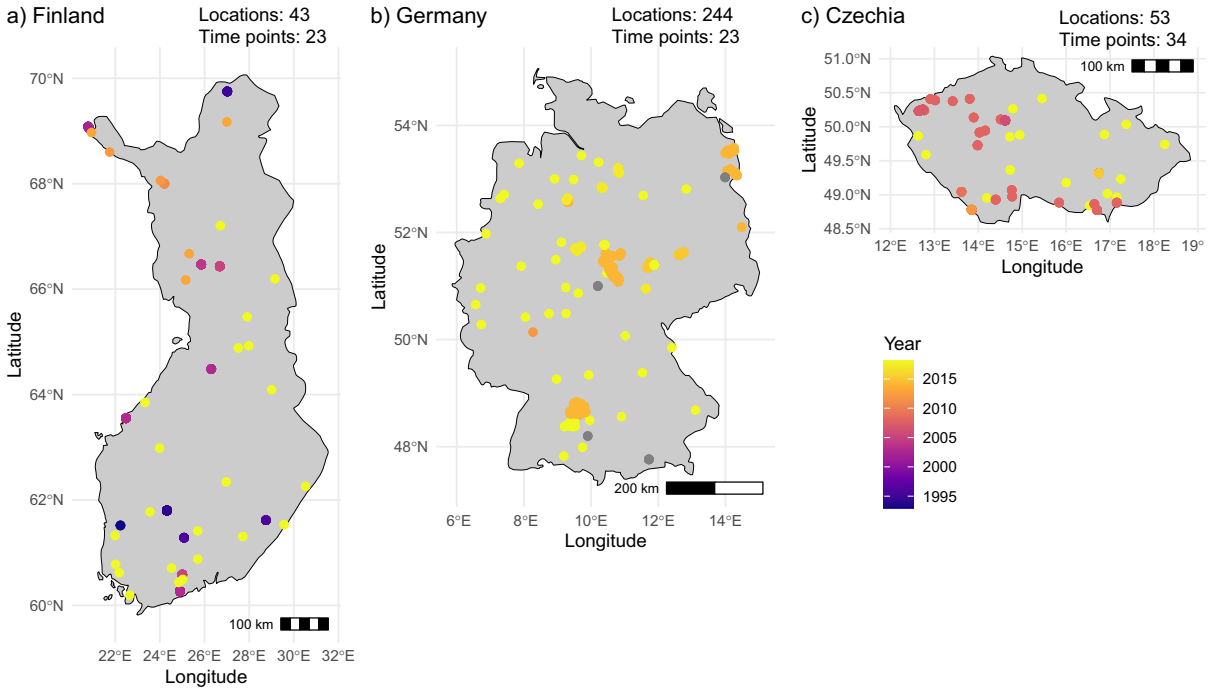

Fig. 5 Examples of countries that have data available for multiple years. Points are coloured according to the sampling year; points are grey where sampling year is unavailable. Text indicates the number of locations in each country, and the number of unique time points (year–month combinations) that samples were collected at.

One important property of PLFA content measurements is that they are additive. As mentioned earlier, different PLFAs can be associated with different microbial functional groups, such as fungi, Gram-negative bacteria, actinobacteria, and other Gram-positive bacteria[11,13,57]. PLFA measurements that are markers of the same group can therefore be summed to estimate biomass of those groups and the ratios of different groups examined. However, the accuracy of some commonly used classifications has been questioned[10,11]. For example, it is recognised that the PLFA 16:1ω5, sometimes used as a marker of arbuscular mycorrhizal fungi, also occurs in significant amounts in bacteria[58,59]. Also, some PLFAs can be good indicators of certain groups in some (agro) ecosystems, but not others[10]. For example, 18:1ω9 can be a good indicator of fungi in soils under forest but not agricultural land use[10]. It is important that data users review the latest literature and be aware of any potential errors with classifications.

The magnitude of enzyme activities measured in soils can be strongly influenced by several methodological parameters, in particular assay temperature, pH, and substrate concentration[60–64]. It is important that the impact of different assay conditions on enzyme activity is considered when analysing and interpreting these data, because whilst there are recommendations for best practices in soil enzyme activity assays[65] these are not always strictly adhered to[62]. We made the assumption that soil enzyme activities were assayed at non-rate limiting substrate concentrations (i.e., activities approximate maximum catalysis rate (V_{max})), which is recommended[22,66] to ensure that the activity assayed is independent of substrate-concentration[67]. As a result, soil enzyme activities assayed at non-rate liming activities reflect inherent differences in activity (V_{max}) of a soil sample[62]. Confirming

that substrate concentrations approximate V_{max} requires soil sample-specific substrate saturation curves to be calculated[62,68]. We assume in good faith that individual labs have confirmed that the substrate concentrations used achieve V_{max}, which can vary by soil as well as assay conditions such as buffer and temperature[22,68,69].

Similarly, as enzyme activities are standardised by maximizing activity, we assume that individual researchers confirmed for their soils that the use of assay pH[70], substrate concentration[22,62] and matrix (e.g., buffer)[68] ensure maximization of assayed activity. Though there are multiple issues with assumptions made in assay conditions that maximize activities[68,70], such assumptions are ubiquitous in soil enzyme activity assays. Thus, this is a potential issue that impacts all soil enzyme activity data and not just our dataset. Our dataset reflects the best possible quality to-date in the field, even though there are methodological improvements needed[71].

Though most researchers standardise enzyme activity assay by using assumed assay conditions that maximise activities to measure the maximum potential activity[23,72], others perform assays under temperature and pH conditions that match those at the locations at which the samples were collected, in an attempt to better assess *in situ* activities[22,23]. For some research questions, this may be the more appropriate approach[71]. The methods used to assay soil enzyme activities in our database include a mix of both approaches, which reflects the reality of methodology diversity in soil enzyme activity assays used by researchers globally. To ensure transparency and enable interpretation of enzyme activities based on assay conditions, we have provided all the metadata available for each sample (e.g., assay temperature and pH) as well as soil properties so that database users can incorporate these variables into models in a way that is most appropriate for the analyses being performed. For example, it is possible to use temperature sensitivity models (e.g., Arrhenius equations) to normalize activity for different enzymes based on known enzyme kinetics[60]. To date, no well-established standardisation methods currently exist that take into account the full complexity of the assay parameters of temperature, matrix type and substrate concentration that may impact absolute values of assayed enzyme activities[62,64,73]. This is a clear need for soil enzymology.

Finally, users should be aware of the biases in the database towards certain geographical regions and climatic zones. The impact of both geographical and climatic biases on model outputs should be carefully explored when conducting any analysis[74]. Data thinning or other bias correction approaches may be required. Additionally, many PLFAs and enzyme activities contain measurements from multiple samples collected at the same location, and so users must decide on the most appropriate way to treat these values.

Code availability

Code used to conduct technical validation analyses and create the figures is available on Figshare[48].

Received: 31 March 2025; Accepted: 1 August 2025;
Published online: 26 September 2025

References

1. Crowther, T. W. *et al*. Scientists' call to action: Microbes, planetary health, and the Sustainable Development Goals. *Cell* **187**, 5195–5216, https://doi.org/10.1016/j.cell.2024.07.051 (2024).
2. Caldwell, B. A. Enzyme activities as a component of soil biodiversity: A review. *Pedobiologia* **49**, 637–644, https://doi.org/10.1016/j.pedobi.2005.06.003 (2005).
3. Frostegård, Å., Tunlid, A. & Bååth, E. Microbial biomass measured as total lipid phosphate in soils of different organic content. *Journal of Microbiological Methods* **14**, 151–163, https://doi.org/10.1016/0167-7012(91)90018-L (1991).
4. Semenov, M. V. Metabarcoding and Metagenomics in Soil Ecology Research: Achievements, Challenges, and Prospects. *Biology Bulletin Reviews* **11**, 40–53, https://doi.org/10.1134/S2079086421010084 (2021).
5. Ma, B. *et al*. A genomic catalogue of soil microbiomes boosts mining of biodiversity and genetic resources. *Nature Communications* **14**, 7318, https://doi.org/10.1038/s41467-023-43000-z (2023).
6. Tedersoo, L. *et al*. The Global Soil Mycobiome consortium dataset for boosting fungal diversity research. *Fungal Diversity* **111**, 573–588, https://doi.org/10.1007/s13225-021-00493-7 (2021).
7. Thompson, L. R. *et al*. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457–463, https://doi.org/10.1038/nature24621 (2017).
8. Větrovský, T. *et al*. GlobalFungi, a global database of fungal occurrences from high-throughput-sequencing metabarcoding studies. *Scientific Data* **7**, 228, https://doi.org/10.1038/s41597-020-0567-7 (2020).
9. Siles, J. A., Gómez-Pérez, R., Vera, A., García, C. & Bastida, F. A comparison among EL-FAME, PLFA, and quantitative PCR methods to detect changes in the abundance of soil bacteria and fungi. *Soil Biology and Biochemistry* **198**, 109557, https://doi.org/10.1016/j.soilbio.2024.109557 (2024).
10. Frostegård, Å., Tunlid, A. & Bååth, E. Use and misuse of PLFA measurements in soils. *Soil Biology and Biochemistry* **43**, 1621–1625, https://doi.org/10.1016/j.soilbio.2010.11.021 (2011).
11. Willers, C., Jansen van Rensburg, P. J. & Claassens, S. Phospholipid fatty acid profiling of microbial communities–a review of interpretations and recent applications. *Journal of Applied Microbiology* **119**, 1207–1218, https://doi.org/10.1111/jam.12902 (2015).
12. Bligh, E. G. & Dyer, W. J. A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology* **37**, 911–917 https://doi.org/10.1139/o59-099 %M13671378 (1959).
13. Ruess, L. & Chamberlain, P. M. The fat that matters: Soil food web analysis using fatty acids and their carbon stable isotope signature. *Soil Biology and Biochemistry* **42**, 1898–1910, https://doi.org/10.1016/j.soilbio.2010.07.020 (2010).
14. Zelles, L. Fatty acid patterns of phospholipids and lipopolysaccharides in the characterisation of microbial communities in soil: a review. *Biology and Fertility of Soils* **29**, 111–129, https://doi.org/10.1007/s003740050533 (1999).
15. Zhang, Y. *et al*. High turnover rate of free phospholipids in soil confirms the classic hypothesis of PLFA methodology. *Soil Biology and Biochemistry* **135**, 323–330, https://doi.org/10.1016/j.soilbio.2019.05.023 (2019).
16. Joergensen, R. G. & Emmerling, C. Methods for evaluating human impact on soil microorganisms based on their activity, biomass, and diversity in agricultural soils. *Journal of Plant Nutrition and Soil Science* **169**, 295–309, https://doi.org/10.1002/jpln.200521941 (2006).
17. Klamer, M. & Bååth, E. Estimation of conversion factors for fungal biomass determination in compost using ergosterol and PLFA 18:2ω6,9. *Soil Biology and Biochemistry* **36**, 57–65, https://doi.org/10.1016/j.soilbio.2003.08.019 (2004).
18. Baldrian, P. *et al*. Estimation of fungal biomass in forest litter and soil. *Fungal Ecology* **6**, 1–11, https://doi.org/10.1016/j.funeco.2012.10.002 (2013).

19. Frostegård, Å., Bååth, E. & Tunlio, A. Shifts in the structure of soil microbial communities in limed forests as revealed by phospholipid fatty acid analysis. *Soil Biology and Biochemistry* **25**, 723–730 (1993).
20. Burns, R. G. Enzyme activity in soil: Location and a possible role in microbial ecology. *Soil Biology and Biochemistry* **14**, 423–427, https://doi.org/10.1016/0038-0717(82)90099-2 (1982).
21. Skujiņš, J. & Burns, R. G. Extracellular Enzymes in Soil. *CRC Critical Reviews in Microbiology* **4**, 383–421, https://doi.org/10.3109/10408417609102304 (1976).
22. German, D. P. *et al*. Optimization of hydrolytic and oxidative enzyme methods for ecosystem studies. *Soil Biology and Biochemistry* **43**, 1387–1397, https://doi.org/10.1016/j.soilbio.2011.03.017 (2011).
23. Wallenstein, M. D. & Weintraub, M. N. Emerging tools for measuring and modeling the *in situ* activity of soil extracellular enzymes. *Soil Biology and Biochemistry* **40**, 2098–2106, https://doi.org/10.1016/j.soilbio.2008.01.024 (2008).
24. Dick, R. P. *Methods of Soil Enzymology*. (Soil Science Society of America, 2011).
25. Tabatabai, M. A. & Bremner, J. M. Use of p-nitrophenyl phosphate for assay of soil phosphatase activity. *Soil Biology and Biochemistry* **1**, 301–307, https://doi.org/10.1016/0038-0717(69)90012-1 (1969).
26. Maestre, F. T. & Eisenhauer, N. Recommendations for establishing global collaborative networks in soil ecology. *Soil Organisms* **91**, 73–85, https://doi.org/10.25674/so91iss3pp73 (2019).
27. He, L. *et al*. Global biogeography of fungal and bacterial biomass carbon in topsoil. *Soil Biology and Biochemistry* **151**, 108024, https://doi.org/10.1016/j.soilbio.2020.108024 (2020).
28. Serna-Chavez, H. M., Fierer, N. & van Bodegom, P. M. Global drivers and patterns of microbial abundance in soil. *Global Ecology and Biogeography* **22**, 1162–1172, https://doi.org/10.1111/geb.12070 (2013).
29. Wan, X., Chen, X., Huang, Z. & Chen, H. Y. H. Global soil microbial biomass decreases with aridity and land-use intensification. *Global Ecology and Biogeography* **30**, 1056–1069, https://doi.org/10.1111/geb.13282 (2021).
30. Xu, X., Thornton, P. E. & Post, W. M. A global analysis of soil microbial biomass carbon, nitrogen and phosphorus in terrestrial ecosystems. *Global Ecology and Biogeography* **22**, 737–749, https://doi.org/10.1111/geb.12029 (2013).
31. Smith, G. R., Crowther, T. W., Eisenhauer, N. & van den Hoogen, J. Building a global database of soil microbial biomass and function: a call for collaboration. *Soil Organisms* **91**, 139–142, https://doi.org/10.25674/so91iss3pp140 (2020).
32. NEON: National Ecological Observatory Network. Soil microbe biomass (DP1.10104.001), provisional data. Dataset accessed from https://data.neonscience.org/data-products/DP1.10104.001 on March 13 2020 (2025).
33. Hu, L. *et al*. Changes in soil physicochemical and microbial properties along elevation gradients in two forest soils. *Scandinavian Journal of Forest Research* **31**, 242–253, https://doi.org/10.1080/02827581.2015.1125522 (2016).
34. McGuire, K. L., Zak, D. R., Edwards, I. P., Blackwood, C. B. & Upchurch, R. Slowed decomposition is biotically mediated in an ectomycorrhizal, tropical rain forest. *Oecologia* **164**, 785–795, https://doi.org/10.1007/s00442-010-1686-1 (2010).
35. Ushio, M., Wagai, R., Balser, T. C. & Kitayama, K. Variations in the soil microbial community composition of a tropical montane forest ecosystem: Does tree species matter? *Soil Biology and Biochemistry* **40**, 2699–2702, https://doi.org/10.1016/j.soilbio.2008.06.023 (2008).
36. Wang, X., Sasaki, A., Toda, M. & Nakatsubo, T. Changes in soil microbial community and activity in warm temperate forests invaded by moso bamboo (*Phyllostachys pubescens*). *Journal of Forest Research* **21**, 235–243, https://doi.org/10.1007/s10310-016-0533-6 (2016).
37. Xu, M. *et al*. Soil microbial community structure and activity along a montane elevational gradient on the Tibetan Plateau. *European Journal of Soil Biology* **64**, 6–14, https://doi.org/10.1016/j.ejsobi.2014.06.002 (2014).
38. Zhou, L. *et al*. Soil extracellular enzyme activity and stoichiometry in China's forests. *Functional Ecology* **34**, 1461–1471, https://doi.org/10.1111/1365-2435.13555 (2020).
39. Tummers, B. DataThief III. https://datathief.org/ (2006).
40. Schutter, M. E. & Dick, R. P. Comparison of Fatty Acid Methyl Ester (FAME) Methods for Characterizing Microbial Communities. *Soil Science Society of America Journal* **64**, 1659–1668, https://doi.org/10.2136/sssaj2000.6451659x (2000).
41. Li, C., Cano, A., Acosta-Martinez, V., Veum, K. S. & Moore-Kucera, J. A comparison between fatty acid methyl ester profiling methods (PLFA and EL-FAME) as soil health indicators. *Soil Science Society of America Journal* **84**, 1153–1169, https://doi.org/10.1002/saj2.20118 (2020).
42. Zizka, A. *et al*. CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution* **10**, 744–751 (2019).
43. Dinerstein, E. *et al*. An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm. *BioScience* **67**, 534–545, https://doi.org/10.1093/biosci/bix014 (2017).
44. Hijmans, R. terra: Spatial Data Analysis. R package version 1.7-78, https://CRAN.R-project.org/package=terra (2024).
45. Massicotte, P. & South, A. rnaturalearth: World Map Data from Natural Earth. R package version 1.0.1, https://CRAN.R-project.org/package=rnaturalearth (2023).
46. Pebesma, E. & Bivand, R. *Spatial Data Science: With Applications in R (1st ed.)*. (Chapman and Hall/CRC, 2023).
47. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/ (2024).
48. van Galen, L. G. Data for "A global database of soil microbial phospholipid fatty acids and enzyme activities". https://doi.org/10.6084/m9.figshare.28684586 (2025).
49. van den Hoogen, J. *et al*. Soil nematode abundance and functional group composition at a global scale. *Nature* **572**, 194–198, https://doi.org/10.1038/s41586-019-1418-6 (2019).
50. Karger, D. N. *et al*. Climatologies at high resolution for the earth's land surface areas. *Scientific Data* **4**, 170122, https://doi.org/10.1038/sdata.2017.122 (2017).
51. Cameron, E. K. *et al*. Global gaps in soil biodiversity data. *Nature Ecology & Evolution* **2**, 1042–1043, https://doi.org/10.1038/s41559-018-0573-8 (2018).
52. Lembrechts, J. J. *et al*. Global maps of soil temperature. *Global Change Biology* **28**, 3110–3144, https://doi.org/10.1111/gcb.16060 (2022).
53. Poggio, L. *et al*. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL* **7**, 217–240, https://doi.org/10.5194/soil-7-217-2021 (2021).
54. Tuanmu, M.-N. & Jetz, W. A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. *Global Ecology and Biogeography* **23**, 1031–1045, https://doi.org/10.1111/geb.12182 (2014).
55. Koldasbayeva, D. *et al*. Challenges in data-driven geospatial modeling for environmental research and practice. *Nature Communications* **15**, 10700, https://doi.org/10.1038/s41467-024-55240-8 (2024).
56. Faust, K., Lahti, L., Gonze, D., de Vos, W. M. & Raes, J. Metagenomics meets time series analysis: unraveling microbial community dynamics. *Current Opinion in Microbiology* **25**, 56–66, https://doi.org/10.1016/j.mib.2015.04.004 (2015).
57. Joergensen, R. G. Phospholipid fatty acids in soil—drawbacks and future prospects. *Biology and Fertility of Soils* **58**, 1–6, https://doi.org/10.1007/s00374-021-01613-w (2022).
58. Ngosong, C., Gabriel, E. & Ruess, L. Use of the Signature Fatty Acid 16:1ω5 as a Tool to Determine the Distribution of Arbuscular Mycorrhizal Fungi in Soil. *Journal of Lipids* **2012**, 236807, https://doi.org/10.1155/2012/236807 (2012).
59. Olsson, P. A. & Lekberg, Y. A critical review of the use of lipid signature molecules for the quantification of arbuscular mycorrhiza fungi. *Soil Biology and Biochemistry* **166**, 108574, https://doi.org/10.1016/j.soilbio.2022.108574 (2022).

60. Allison, S. D., Romero-Olivares, A. L., Lu, Y., Taylor, J. W. & Treseder, K. K. Temperature sensitivities of extracellular enzyme V_{max} and K_m across thermal environments. *Global Change Biology* **24**, 2884–2897, https://doi.org/10.1111/gcb.14045 (2018).
61. Fraser, T. D., Duddigan, S., Diaz, A., Green, I. & Tibbett, M. Optimizing pH for Soil Enzyme Assays Reveals Important Biochemical Functions in Low pH Soil. *Journal of Soil Science and Plant Nutrition* **24**, 6236–6247, https://doi.org/10.1007/s42729-024-01866-y (2024).
62. Margenot, A. J., Nakayama, Y. & Parikh, S. J. Methodological recommendations for optimizing assays of enzyme activities in soil samples. *Soil Biology and Biochemistry* **125**, 350–360, https://doi.org/10.1016/j.soilbio.2017.11.006 (2018).
63. Puissant, J. *et al*. The pH optimum of soil exoenzymes adapt to long term changes in soil pH. *Soil Biology and Biochemistry* **138**, 107601, https://doi.org/10.1016/j.soilbio.2019.107601 (2019).
64. Wallenstein, M., Allison, S. D., Ernakovich, J., Steinweg, J. M. & Sinsabaugh, R. in *Soil Enzymology* (eds Shukla, G. & Varma, A.) 245–258 (Springer Berlin Heidelberg, 2011).
65. Malcolm, R. E. Assessment of phosphatase activity in soils. *Soil Biology and Biochemistry* **15**, 403–408, https://doi.org/10.1016/0038-0717(83)90003-2 (1983).
66. Bell, C. W. *et al*. High-throughput fluorometric measurement of potential soil extracellular enzyme activities. *Journal of Visualized Experiments*, e50961 https://doi.org/10.3791/50961 (2013).
67. Tabatabai, M. A. in *Methods of Soil Analysis* (eds Weaver, R. W. *et al*.) 775–833 (Soil Science Society of America, 1994).
68. Li, C. *et al*. Do chromogenic assays of soil enzyme activities need buffers? More disadvantages than advantages of modified universal buffer in the para-nitrophenyl-based assay of phosphomonoesterase and β-glucosidase activities. *Soil Biology and Biochemistry* **202**, 109704, https://doi.org/10.1016/j.soilbio.2024.109704 (2025).
69. Li, C. & Margenot, A. J. Apparent kinetic properties of soil phosphomonoesterase and β-glucosidase are disparately influenced by pH. *Soil Science Society of America Journal* **85**, 2007–2018, https://doi.org/10.1002/saj2.20332 (2021).
70. Wade, J. *et al*. Prescribed pH for soil β-glucosidase and phosphomonoesterase do not reflect pH optima. *Geoderma* **401**, 115161, https://doi.org/10.1016/j.geoderma.2021.115161 (2021).
71. Margenot, A. J. & Wade, J. Getting the basics right on soil enzyme activities: A comment on Sainju *et al*. (2022). *Agrosystems, Geosciences & Environment* **6**, e20405, https://doi.org/10.1002/agg2.20405 (2023).
72. Nannipieri, P., Trasar-Cepeda, C. & Dick, R. P. Soil enzyme activity: a brief history and biochemistry as a basis for appropriate interpretations and meta-analysis. *Biology and Fertility of Soils* **54**, 11–19, https://doi.org/10.1007/s00374-017-1245-6 (2018).
73. Fanin, N. *et al*. Soil enzymes in response to climate warming: Mechanisms and feedbacks. *Functional Ecology* **36**, 1378–1395, https://doi.org/10.1111/1365-2435.14027 (2022).
74. Cosentino, F. & Maiorano, L. Is geographic sampling bias representative of environmental space? *Ecological Informatics* **64**, 101369, https://doi.org/10.1016/j.ecoinf.2021.101369 (2021).

Acknowledgements

We are grateful to Mark Bradford, Qiang Gao, Diana Wall and Deli Wang for providing data and advice. LGvG, JvdH, GRS and TWC acknowledge funding from DOB Ecology (EC-2021-GEM003), the Marc R. Benioff Revocable Trust in collaboration with the World Economic Forum, and the Bernina Initiative (2022-FS-318). GRS was additionally supported by a Graduate Research Fellowship from the United States National Science Foundation, a ThinkSwiss Research Scholarship from the Embassy of Switzerland in the United States of America, and an Ambizione fellowship from the Swiss National Science Foundation (PZ00P3_216194). KGP is a Canadian Institute for Advanced Research (CIFAR) Fellow in the program Fungal Kingdom: Threats and Opportunities, and is supported by the U.S. Department of Energy (DE-SC0023661), and the U.S. National Science Foundation (DEB-1845544 and DEB-1926335). JMA and TAA acknowledge funding from Qatar Petroleum (QUEX-CAS-QP-RD-18/19). ASFA acknowledges the National Council for Scientific and Technological Development (CNPq-Brazil) for his Fellowship of Research Productivity (Grant 301755/2022-1). MB was supported by the Swedish Research Councils Vetenskapsrådet (Grant 2021–03724). PVM and RF acknowledge support by national funds through FCT (UID/EMS/00285/2020). LGB and SLutz acknowledge financial support from a Helmholtz Recruiting Initiative award (I-044-16-0) and a Baillet Latour Antarctica Fellowship (2016–2018) to LAZ that supported part of the soil sample collection in Antarctica. NE, SC and CGu acknowledge support of iDiv funded by the German Research Foundation (DFG– FZT 118, 202548816). NE acknowledges funding by the DFG (Ei 862/29-1 and Ei 862/31-1) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (677232). CGa acknowledges funding by the Portuguese Science and Technology Foundation (PTDC/BIA-CBI/2340/2020). HF and KPe acknowledge funding from the Academy of Finland (315415). FBa thanks the project PID2020-11942RB-I00 funded by MCIN/AEI/10.13039/501100011033, the ILINK program from CSIC (ILINK23065), the project SOMMIT funded by the European Joint Programme SOIL (862695), and Fundación Séneca (19896/GERM/15). MGB and HI acknowledge funding from the Fonds zur Förderung der wissenschaftlichen Forschung (FWF) Austria (I989-B16). MGB also acknowledges support from the Spanish State Agency for Research (AEI) for her Ramón y Cajal grant (RYC-2016-21231). VEJJ acknowledges financial support from the French National Research Agency (ANR; MIXOPEAT; ANR-17-CE01-0007). JZL and SSh acknowledge support from the German Research Foundation through the RTG 2300 (316045089). MZH and SSu acknowledge support from JSPS KAENHI (15658101 and 15380181). VK, SSh and AMP acknowledge the Deutsche Forschungsgemeinschaft (DFG; 192626868–SFB 990). SDF acknowledges support from the U.S. National Science Foundation Long-term Ecological Research (LTER) and Long-term Research in Environmental Biology (LTREB) Programs. MAA acknowledges support from a National Science Foundation Graduate Research Fellowship (DGE 1450271). RDB, AADB and WP acknowledge funding from the UK Natural Environment Research Council (NE/N009452/1). EJS was funded by a European Research Council Starting Grant (FP/2007-2013/ERC 307888). JP was supported by a European Union Horizon 2020 Marie Sklodowska-Curie grant (892654). ATN was supported by a UK NERC grant (NE/T012226). AJM was supported by Illinois Nutrient Research and Education Council (2021-4-360731-469), USDA NIFA (2021-67019-35068), and US NSF (2125626). CR acknowledges funding from the COST Action FP1305 BioLink and from the program "Sustainability and resilience – Tackling climate and environmental changes" (Vetenskapsrådet 2016-06327) jointly supported by the Swedish Research Council, the Swedish Research Council Formas, and Sida. RM acknowledges the funding from the University of Kassel and Valli Sustainability Research and Education. FTM acknowledges support by the King Abdullah University of Science and Technology (KAUST) and the KAUST Climate and Livability Initiative,

the European Research Council (647038) and Generalitat Valenciana (CIDEGENT/2018/041). KZ acknowledges support from the US National Science Foundation (2244711). IH, MMo, JO, MÖ, SP, TV, and MZ acknowledge support from the Estonian Ministry of Education and Research (Centre of Excellence AgroCropFuture) and the Estonian Research Council (PRG1065, PRG1789, PSG784). CW was funded by the National Natural Science Foundation of China (32101286) and China Postdoctoral Science Foundation (2021M693360). CCD was supported by the FCT (BIPD_01_2021_FCT-PTDC/BIA-CBI/2340/2020, UIDB/05937/2020, UIDP/05937/2020). MNBG was supported by Agencia Nacional de Promoción Científica y Tecnológica of Argentina (PICT 2014-2838). KEM and NO were funded by the UK Natural Environment Research Council (NE/I027037/1). SYuE is grateful to the budgetary funding from the state assignments FWES-2024-0023. LB acknowledges financial support from the Swiss National Science Foundation (315260_149807, SPHAGNOL). NAS thanks NWO foundation for Vidi grant (016.161.318). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. ZZ was supported by the National Natural Science Foundation of China (32061143027). XLi was supported by the Jilin Province Science and Technology Development Plan Project (2020020s1003JC). LJ was supported by the Second Tibetan Plateau Scientific Expedition and Research (STEP) program (2019QZKK0302), and the National Natural Science Foundation of China (31872994). XLu was supported by the National Natural Science Foundation of China (41922056). PB was supported by the Ministry of Education, Youth and Sports of the Czech Republic (CZ.02.01.01/00/22_008/0004635). AP was supported by the Danish National Research Foundation within the Center for Volatile Interactions (VOLT, DNRF168). SMBR acknowledges FAPEPI and CAPES (code 001) for her PhD scholarship. JPS was supported by the strategic plan of the Centre for Functional Ecology - Science for People and the Planet (CFE) (UIDB/04004/2020), funded by FCT/MCTES through national funds (PIDDAC). PK was supported by the Czech Science Foundation (21-20802 M). EK acknowledges the DFG Priority Program 1374 'Infrastructure-Biodiversity-Exploratories' (KA 1590/8-2, KA 1590/8-3) and AA acknowledges the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001 – process 88881.172163/2018-01 for her fellowship. NF acknowledges the support of the Region Nouvelle-Aquitaine (AAPR2020A-2019–8472310). MMR acknowledges support from the Spanish Ministry of Science and Innovation for the I + D + i Project PID2021-123097OA-I00 funded by MCIN/ AEI /10.13039/501100011033/ and FEDER Una manera de hacer Europa and for the Project TED2021-132332A-C22, funded by MCIN/AEI /10.13039/501100011033 and Unión Europea NextGenerationEU/ PRTR. JV was supported by the Czech Science Foundation project (21-19209 M). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Author contributions

G.R.S., J.v.d.H., K.G.P., R.B.J., K.Y. and T.W.C. conceived the study. L.G.v.G., G.R.S. and J.v.d.H. led compiling and cleaning the database with advice from A.J.M. and M.P.W. L.G.v.G. and J.vd.H. performed the analyses. L.G.v.G. wrote the initial manuscript draft with input from J.v.d.H. and G.R.S. All other authors cleaned and curated data for the database, or provided substantial advice regarding the database and manuscript structure. All authors edited the manuscript.

Funding

Open access funding provided by Swiss Federal Institute of Technology Zurich.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41597-025-05759-2.

Correspondence and requests for materials should be addressed to L.G.v.G. or J.v.D.H.

Laura G. van Galen[1]✉, Gabriel Reuben Smith[1,2], Andrew J. Margenot[3,4], Mark P. Waldrop[5], Thomas W. Crowther[1], Kabir G. Peay[2,6], Robert B. Jackson[6], Kailiang Yu[7,8], Anna Abrahão[9], Talaat A. Ahmed[10], Juha M. Alatalo[10], Sten Anslan[11,12], Mark A. Anthony[13], Ademir Sergio Ferreira Araujo[14], Judith Ascher-Jenull[15], Elizabeth M. Bach[16], Mohammad Bahram[17,18], Christopher C. M. Baker[19], Petr Baldrian[20], Richard D. Bardgett[21,22], M. Noelia Barrios-Garcia[23,24], Felipe Bastida[25], Francesca Beggi[26], Liane G. Benning[27,28], Luca Bragazza[29], Arthur A. D. Broadbent[21,30], Concha Cano-Díaz[31], Anna M. Cates[32], Carlos E. P. Cerri[33], Simone Cesarz[34,35], Baodong Chen[36,37], Aimeé T. Classen[38], Mathilde Borg Dahl[39], Manuel Delgado-Baquerizo[40], Nico Eisenhauer[34,35], Svetlana Yu. Evgrafova[41,42], Nicolas Fanin[43], Flavio Fornasier[44], Romeu Francisco[45], André L. C. Franco[46], Serita D. Frey[47], Hannu Fritze[48], Carlos García[25], Pablo García-Palacios[49], María Gómez-Brandón[50], Marina Gonzalez-Polo[51], Beatriz Gozalo[52,53], Robert Griffiths[54], Carlos Guerra[35,55], Moritz Hallama[56], Inga Hiiesalu[57], Mohammad Zabed Hossain[58], Yajun Hu[36,37,59], Heribert Insam[60], Vincent E. J. Jassey[61], Lili Jiang[62], Ellen Kandeler[56], Petr Kohout[63], Urmas Kõljalg[57], Valentyna Krashevska[64], Xiaofei Li[65], Jing-Zhong Lu[66], Xiankai Lu[67], Shan Luo[22,68], Stefanie Lutz[27,69], Kathleen Allison Mackie-Haas[70], Fernando T. Maestre[71], Minna Malmivaara-Lämsä[48], Kai Mangelsdorf[72], Maria Manjarrez[73], Sven Marhan[56], Ashley Martin[74,75], Kelly E. Mason[76], Jordan Mayor[77], Rebecca L. McCulley[78], Mari Moora[57], Paula V. Morais[45], Miriam Muñoz-Rojas[40,79], Rajasekaran Murugan[80,81], Andrew T. Nottingham[82], Victoria Ochoa[83], Raúl Ochoa-Hueso[84,85], Jane Oja[57], Pål Axel Olsson[86], Maarja Öpik[57], Nick Ostle[22], Krista Peltoniemi[48], Taina Pennanen[48], David S. Pescador[87], G. Kenny Png[21,88], Christian Poll[56], Sergei Põlme[57,89], Anton M. Potapov[34,90], Anders Priemé[91], William Pritchard[21], Jeremy Puissant[92], Sandra Mara Barbosa Rocha[14], Christoph Rosinger[93], Liliane Ruess[94], Emma J. Sayer[95], Stefan Scheu[66,96], Robert L. Sinsabaugh[97], Lindsey C. Slaughter[98], Nadejda A. Soudzilovskaia[99], José Paulo Sousa[100], Lee Stanish[101], Shu-ichi Sugiyama[102], Leho Tedersoo[57,103,104], Pankaj Trivedi[105,106], Tanel Vahter[57], Jana Voriskova[20], Dirk Wagner[107], Cong Wang[108], David A. Wardle[109], Jeanette Whitaker[76], Yuanhe Yang[37,110], Zhiwei Zhong[111,112,113], Kai Zhu[114], Lori A. Ziolkowski[115], Martin Zobel[57] & Johan van den Hoogen[1]✉

[1]Institute of Integrative Biology, Department of Environmental Systems Sciences, ETH Zurich, Zurich, Switzerland. [2]Department of Biology, Stanford University, Stanford, California, USA. [3]Department of Crop Sciences, University of Illinois Urbana-Champaign, Urbana, IL, 61801, USA. [4]Agroecosystem Sustainability Center (ASC), Institute for Sustainability, Energy and Environment (iSEE), Urbana, IL, 61801, USA. [5]US Geological Survey, Geology, Minerals, Energy, and Geophysics Science Center, Menlo Park, California, USA. [6]Department of Earth System Science, Stanford University, Stanford, California, USA. [7]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA. [8]High Meadows Environmental Institute, Princeton University, Princeton, NJ, USA. [9]Department of Biology, Science Center, Federal University of Ceará – UFC, Fortaleza, CE, 60440-900, Brazil. [10]Environmental Science Center, Qatar University, Doha, Qatar. [11]Department of Biological and Environmental Science, University of Jyväskylä, Jyväskylä, Finland. [12]Department of Biology, College of Science, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. [13]Center for Microbiology and Environmental Systems Science, University of Vienna, 1030, Vienna, Austria. [14]Soil Microbial Ecology Group, Universidade Federal do Piauí (UFPI), 64019-550, Teresina, Piauí, Brazil. [15]Department of Experimental Architecture, Integrative Design Extremes, University of Innsbruck, Technikerstrasse 21, Innsbruck, A-6020, Austria. [16]The Nature Conservancy in Illinois, Nachusa Grasslands, Franklin Grove, IL, USA. [17]Department of Ecology, Swedish University of Agricultural Sciences, Uppsala, Ulls väg 16, 756 51, Sweden. [18]Department of Agroecology, Aarhus University, Slagelse, 4200, Denmark. [19]US Army ERDC Cold Regions Research and Engineering Laboratory, Hanover, NH, 03755, USA. [20]Laboratory of Environmental Microbiology, Institute of Microbiology of the Czech Academy of Sciences, Videnska 1083, 14200 Praha 4, Prague, Czech Republic. [21]Department of Earth and Environmental Sciences, The University of Manchester, Manchester, M13 9PT, UK. [22]Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ, UK. [23]Rubenstein School of Environment and Natural Resources, University of Vermont, Burlington, VT, USA. [24]CENAC-APN, CONICET, Bariloche, Argentina. [25]CEBAS-CSIC Department of Soil and Water Conservation Campus Universitario de Espinardo, 30100, Murcia, Spain. [26]Alliance Bioversity International - CIAT, Bengaluru, India. [27]GFZ Helmholtz Centre for Geosciences, Interface Geochemistry Section, Telegrafenberg, 14473, Potsdam, Germany. [28]Department of Earth Sciences, Freie Universität Berlin, 12249, Berlin, Germany. [29]Field-Crop Systems and Plant Nutrition, Agroscope, Route de Duillier 60, CH-1260, Nyon, Switzerland. [30]Biological & Environmental Sciences, Faculty of Natural Sciences, University of Stirling, FK9 4LA, Stirling, UK. [31]CISAS - Centre for Research and Development in Agrifood Systems and Sustainability, Polytechnic Institute of Viana do Castelo, Rua Escola Industrial e Comercial Nun'Álvares, 34, 4900-347, Viana do Castelo, Portugal. [32]Department of Soil, Water, and Climate, University of Minnesota, 1991 Upper Buford Circle, St Paul, MN, 55108, USA. [33]University of São Paulo ESALQ/USP Av Pádua Dias 11, 13418-900, Piracicaba, SP, Brazil. [34]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstrasse 4, 04103, Leipzig, Germany.

[35]Institute of Biology, Leipzig University, Puschstrasse 4, 04103, Leipzig, Germany. [36]State Key Laboratory of Regional and Urban Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, China. [37]College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China. [38]Ecology and Evolutionary Biology Department, University of Michigan, Ann Arbor, MI, 48109, USA. [39]Institute of Microbiology, Center for Functional Genomics of Microbes, University of Greifswald, 17489, Greifswald, Germany. [40]Laboratorio de Biodiversidad y Funcionamiento Ecosistémico, Instituto de Recursos Naturales y Agrobiología de Sevilla (IRNAS), Consejo Superior de Investigaciones Científicas (CSIC), Av., Reina Mercedes 10, E-41012, Sevilla, Spain. [41]Sukachev Institute of Forest SB RAS, Federal Research Center "Krasnoyarsk Science Center SB RAS", Krasnoyarsk, 660036, Russia. [42]Melnikov Permafrost Institute, Siberian Branch, Russian Academy of Sciences, 677010, Yakutsk, Russia. [43]INRAE, Bordeaux Sciences Agro, UMR 1391 Interaction Soil Plant Atmosphere (ISPA), 71 avenue Edouard Bourlaux, 33882, Villenave-d'Ornon cedex, France. [44]CREA Research Center for Viticulture and Enology, Gorizia, Italy. [45]University of Coimbra, Centre for Mechanical Engineering, Materials and Processes, ARISE, Department of Life Sciences, 3000-456, Coimbra, Portugal. [46]O'Neill School of Public and Environmental Affairs, Indiana University, Bloomington, IN, 47405, USA. [47]Department of Natural Resources and the Environment, University of New Hampshire, Durham, NH, 03824, USA. [48]Natural Resources Institute Finland, Luke, Latokartanonkaari 9, 00790, Helsinki, Finland. [49]Instituto de Ciencias Agrarias, Consejo Superior de Investigaciones Científicas, Serrano 115 bis, 28006, Madrid, Spain. [50]Animal Ecology Group (GEA), University of Vigo, E-36310, Vigo, Spain. [51]INIBIOMA, CONICET- Universidad Nacional del Comahue, Bariloche, Argentina. [52]Instituto Multidisciplinar para el Estudio del Medio "Ramón Margalef", Universidad de Alicante, Alicante, Spain. [53]Departamento de Ecología, Universidad de Alicante, Alicante, Spain. [54]School of Environmental and Natural Sciences, Bangor University, Gwynedd, LL57 2DG, UK. [55]University of Coimbra, Department of Geography, Colégio de São Jerónimo, Largo D. Dinis, 3004-530, Coimbra, Portugal. [56]Institute of Soil Science and Land Evaluation, University of Hohenheim, D-70593, Stuttgart-Hohenheim, Germany. [57]Institute of Ecology and Earth Sciences, University of Tartu, 2 J. Liivi St, 50409, Tartu, Estonia. [58]Department of Botany, University of Dhaka, Dhaka, 1000, Bangladesh. [59]College of Agronomy, Hunan Agricultural University, Changsha, 410128, P.R. China. [60]BioTreaT GmbH, Technikerstrasse 21, 6020, Innsbruck, Austria. [61]Centre de Recherche sur la Biodiversité et l'Environnement (CRBE), Université de Toulouse, CNRS, IRD, Toulouse INP, Université Toulouse 3 – Paul Sabatier (UT3), Toulouse, France. [62]State Key Laboratory of Tibetan Plateau Earth System, Environment and Resources (TPESER), Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, 100101, China. [63]Laboratory of Microbial Ecology and Biogeography, Institute of Microbiology of the Czech Academy of Sciences, Videnska 1083, 14200 Praha 4, Prague, Czech Republic. [64]Senckenberg Biodiversity and Climate Research Centre, Functional Environmental Genomics, Frankfurt, Germany. [65]College of Resources and Environmental Sciences/Key Laboratory of Sustainable Utilization of Soil Resources in the Commodity Grain Bases in Jilin Province, Jilin Agricultural University, Changchun, 130118, China. [66]Johann Friedrich Blumenbach Institute of Zoology and Anthropology, University of Göttingen, Untere Karspüle 2, 37073, Göttingen, Germany. [67]Key Laboratory of Vegetation Restoration and Management of Degraded Ecosystems, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China. [68]Department of Evolution, Ecology, and Behaviour, University of Liverpool, Crown street, Liverpool, L69 7BE, UK. [69]Plant-Soil Interactions, Agroecology and Environment, Agroscope, Reckenholzstrasse 191, 8046, Zurich, Switzerland. [70]Agroscope, Viticulture, 8820, Wädenswil, Switzerland. [71]Environmental Sciences and Engineering, Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal, 23955-6900, Kingdom of Saudi Arabia. [72]GFZ Helmholtz Centre for Geosciences, Organic Geochemistry Section, Telegrafenberg, 14473, Potsdam, Germany. [73]Microbiology Laboratories Australia, St Marys, 5042, Australia. [74]Microbiology Laboratories Australia, Adelaide, Australia. [75]Benevita Biotech, Adelaide, Australia. [76]UK Centre for Ecology & Hydrology, Library Avenue, Bailrigg, Lancaster, LA1 4AP, UK. [77]ICF, 201 Mission St. Suite 1500, San Francisco, CA, 94105, USA. [78]Department of Plant & Soil Sciences, University of Kentucky, Lexington, KY, 40546-0312, USA. [79]Centre for Ecosystem Science, School of Biological, Earth & Environmental Sciences, UNSW, Sydney, 2052, NSW, Australia. [80]Valli Sustainability Research and Education Foundation, Kanchipuram, India. [81]Institute of Soil Research, Department of Ecosystem Management, Climate and Biodiversity, BOKU University, 1190, Vienna, Austria. [82]School of Geography, University of Leeds, Leeds, LS2 9JT, UK. [83]Instituto Univeritario de Investigación en Olivar y Aceite de Oliva-INUO, Universidad de Jaén, Jaén, Spain. [84]Department of Biology, IVAGRO, University of Cádiz, Campus de Excelencia Internacional Agroalimentario (CeiA3), Campus del Rio San Pedro, 11510 Puerto Real, Cádiz, Spain. [85]Department of Terrestrial Ecology, Netherlands Institute of Ecology (NIOO-KNAW), P.O. Box 50, 6700 AB, Wageningen, the Netherlands. [86]Department of Biology, Lund University, SE-223 62, Lund, Sweden. [87]Departamento de Farmacología, Farmacognosia y Botánica, Facultad de Farmacia, Universidad Complutense de Madrid, Madrid, Spain, Pl. de Ramón y Cajal, s/n, Moncloa - Aravaca, 28040, Madrid, Spain. [88]Singapore Botanic Gardens, National Parks Board, Singapore, 259569, Republic of Singapore. [89]Natural History Museum of Tartu University, 46 Vanemuise St., 51014, Tartu, Estonia. [90]Senckenberg Museum of Natural History Gorlitz, Am Museum 1, D-02826, Görlitz, Germany. [91]Department of Biology and Center for Volatile Interactions (VOLT), University of Copenhagen, Copenhagen, Denmark. [92]Laboratoire d'Ecologie Alpine, Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LECA, Grenoble, France. [93]Institute of Agronomy, Department of Agricultural Sciences, BOKU University, 3430, Tulln an der Donau, Austria. [94]Institute of Biology, Ecology Group, Humboldt-Universität zu Berlin, Philippstraße 13, 10115, Berlin, Germany. [95]Institute of Botany, Ulm University, 89081, Ulm, Germany. [96]Centre of Biodiversity and Sustainable Land Use, University of Göttingen, Büsgenweg 1, 37077, Göttingen, Germany. [97]Biology Department, University of New Mexico, Albuquerque, NM, 87131, USA. [98]Texas Tech University, Department of Plant and Soil Science, Lubbock, TX, 79409, USA. [99]Centre for Environmental Sciences, Hasselt University - Campus Diepenbeek Agoralaan Building D, Units AB, B-3590, Diepenbeek, Belgium. [100]Centre for Functional Ecology, Associate Laboratory TERRA, Department of Life Sciences, University of Coimbra, 3000-456, Coimbra, Portugal. [101]Institute of Arctic and Alpine Research University of Colorado, Boulder, CO, USA. [102]Faculty of Agriculture and Life Science, Hirosaki University, Bunkyocho 1, Hirosaki 036-8560, Aomori, Japan. [103]Mycology and Microbiology Center, University of Tartu, Tartu, Estonia.

[104]Department of Zoology, College of Science, King Saud University, Riyadh, Saudi Arabia. [105]Microbiome Network and Department of Agricultural Biology, Colorado State University, Fort Collins, CO, 80523, USA. [106]Institute of Genomics for Crop Abiotic Stress Tolerance, Department of Plant and Soil Sciences, Texas Tech University, Lubbock, TX, 79409, USA. [107]GFZ Helmholtz Centre for Geosciences, Geomicrobiology Section, Telegrafenberg, 14473, Potsdam, Germany. [108]State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, 100101, China. [109]Department of Ecology and Environmental Science, Umeå University, Umeå, Sweden. [110]State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing, China. [111]Key Laboratory of Vegetation Ecology, Ministry of Education/Jilin Songnen, Grassland Ecosystem National Observation and Research Station, Northeast Normal University, Changchun, 130024, China. [112]Heilongjiang Xingkai Lake Wetland Ecosystem National Observation and Research Station & Key Laboratory of Wetland Ecology and Environment, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun, 130102, China. [113]Key Laboratory of Grassland Resources (Inner Mongolia Agricultural University), Ministry of Education, Hohhot, 010021, China. [114]School for Environment and Sustainability, University of Michigan, Ann Arbor, MI, USA. [115]School of the Earth, Ocean and Environment, University of South Carolina, Columbia, SC, 29208, USA. ✉e-mail: laura.vangalen9@gmail.com; johan.vandenhoogen@usys.ethz.ch