Multiple elements of soil biodiversity drive ecosystem functions across biomes

Manuel Delgado-Baquerizo [1,2,3]*, Peter B. Reich [3,4], Chanda Trivedi[3], David J. Eldridge[5], Sebastián Abades[6], Fernando D. Alfaro[6], Felipe Bastida[7], Asmeret A. Berhe [8], Nick A. Cutler [9], Antonio Gallardo [1], Laura García-Velázquez[1], Stephen C. Hart [8], Patrick E. Hayes [10,11,12], Ji-Zheng He [13,14], Zeng-Yei Hseu [15], Hang-Wei Hu[13,14], Martin Kirchmair[16], Sigrid Neuhauser [16], Cecilia A. Pérez[17], Sasha C. Reed[18], Fernanda Santos[8], Benjamin W. Sullivan[19], Pankaj Trivedi[20], Jun-Tao Wang[3], Luis Weber-Grullon[21,22,23], Mark A. Williams[24] and Brajesh K. Singh [3,25]

The role of soil biodiversity in regulating multiple ecosystem functions is poorly understood, limiting our ability to predict how soil biodiversity loss might affect human wellbeing and ecosystem sustainability. Here, combining a global observational study with an experimental microcosm study, we provide evidence that soil biodiversity (bacteria, fungi, protists and invertebrates) is significantly and positively associated with multiple ecosystem functions. These functions include nutrient cycling, decomposition, plant production, and reduced potential for pathogenicity and belowground biological warfare. Our findings also reveal the context dependency of such relationships and the importance of the connectedness, biodiversity and nature of the globally distributed dominant phylotypes within the soil network in maintaining multiple functions. Moreover, our results suggest that the positive association between plant diversity and multifunctionality across biomes is indirectly driven by soil biodiversity. Together, our results provide insights into the importance of soil biodiversity for maintaining soil functionality locally and across biomes, as well as providing strong support for the inclusion of soil biodiversity in conservation and management programmes.

Belowground organisms comprise a large fraction of global terrestrial diversity and are responsible for essential ecosystem functions and services, such as plant productivity, nutrient cycling, organic matter (OM) decomposition, pollutant degradation and pathogen control[1–6], which are valued at trillions of US dollars annually. However, as most soil microorganisms and microfauna are difficult to observe directly, they are often neglected in global biodiversity surveys[7]. As a consequence, the roles played by biodiverse soil organisms (bacteria, fungi, protists and invertebrates; multidiversity[8]) for multiple types of ecosystem functions (ecosystem multifunctionality) remain largely unresolved. Multifunctionality is an important ecological and management concept and provides the basis for a solid statistical approach that enables the synthesis of the many diverse functions that soil organisms provide[2,6,8–10]. Although care must be taken in the development and interpretation of multifunctionality metrics, the approach is widely seen as important for creating a broad understanding of the linkages between diverse soil organisms and ecosystem functions.

Although relatively rare, experimental evidence suggests that soil biodiversity enhances the ability of ecosystems to maintain multifunctionality within controlled microcosm environments[2]. Experimental evidence also indicates that there are strong links between plant and soil biodiversity and function[6]. Moreover, observational studies within single biomes (for example, European temperate grasslands and drylands) and studies dedicated to the study of the biodiversity of a limited number of soil organism types

and biomes[9–11] suggest that soil biodiversity is correlated with the maintenance of numerous ecosystem functions. However, data are lacking for the relationship between the biodiversity of different groups of soil organisms (such as bacteria, fungi, protists and invertebrates) and multiple functions under natural conditions at the global scale across contrasting biomes. Moreover, experimental evidence evaluating how soil microbial diversity is associated with ecosystem functions is also lacking. Rigorous assessment of the role of soil biodiversity in regulating multifunctionality is urgently needed to better understand the potential consequences of soil biodiversity losses for the maintenance of multiple ecosystem functions and services that are critical for human wellbeing and global ecosystem sustainability.

It is also probable that different groups of soil organisms have different roles in maintaining multifunctionality. For example, larger soil invertebrates (such as annelids, tardigrades, arthropods and flatworms) are responsible for processing large amounts of plant and animal detritus[12,13], and might ultimately determine the amount of fresh resources and the potential functional rates in the soil food web. Analogous to the productivity of primary producers, the detrital products of large soil invertebrates help to regulate the functioning of terrestrial ecosystems. These organisms act as a manufacturing line that processes detritus and infuses the soil with physically smaller and chemically decomposed resources. We posit that the diversity of these soil invertebrates might therefore play critical roles in supporting multiple functions (that is, rates and availabilities) operating at high levels of functioning (relative to their maximum observed levels of functioning; as described previously[14]). By contrast, the biodiversity of soil microorganisms (such as protists, bacteria and fungi) might be fundamental for the maintenance of multiple functions and energy flow within the soil food web, but are still beholden to the activities of macrobiota. We therefore hypothesize that the smallest soil organisms are responsible for bottom-up (producers) and top-down (consumers) energy transfer by activating nutrients from the soil, and through predation, recirculating energy from larger to smaller organisms through the microbial loop[15,16]. In other words, these soil organisms recirculate the available resources in soils, ensuring the functioning of terrestrial ecosystems.

Moreover, soil organisms live within complex soil food webs, forming ecological clusters of strongly co-occurring phylotypes within ecological networks[17–19]. These ecological assemblages share similar environmental and resource preferences, and are expected to have important implications for ecosystem functioning[20]. Some of these assemblages—those that include a greater number of functionally important phylotypes—should also support higher levels of ecosystem functioning. However, in theory, the biodiversity of other assemblages dominated by low-functional phylotypes (that is, taxa that support low functional rates) might be less important for maintaining ecosystem functioning, ultimately challenging the hypothesis that all biodiversity is equally important for maintaining ecosystem functions. Furthermore, the degree of connectivity (for example, determined by co-occurrence) among soil phylotypes within these ecological networks might have consequences for ecosystem functioning. Some phylotypes are highly connected with multiple phylotypes within and across ecological clusters (hub phylotypes), whereas others are poorly connected (non-hub phylotypes)[21] within ecological networks. In plant communities, highly connected phylotypes are fundamental for maintaining ecosystem functions and services (such as pollination)[22,23]. Similarly, locations with a higher number of soil taxa—classified as hub phylotypes[21]—could, in theory, support greater levels of multifunctionality by facilitating the interconnection of multiple ecosystem processes (such as metabolic pathways). Evidence of the importance of diversity of soil taxa classified as hubs and within ecological clusters in regulating multifunctionality across the globe is lacking, yet could

lend insights into how community structure determines function, and is therefore in need of empirical study.

Here, we used a multicontinent observational field study and a controlled microcosm experiment to test the linkages among soil biodiversity and multifunctionality. We first conducted a soil analysis across 83 natural (unfertilized) terrestrial ecosystems on five continents and multiple ecosystem biomes (from arid ecosystems to tropical forests; Supplementary Fig. 1, Supplementary Table 1). Using marker-gene sequencing methods, we obtained plot-scale information on the richness (soil diversity) of 12 types of soil organisms, including bacteria, fungi (mycorrhizal and saprotrophic fungi), protists (Cercozoa, Ciliophora and Lobosa) and invertebrates (Annelida, Arthropoda, Nematoda, Rotifer, Tardigrada and Platyhelminthes) comprising around 45,000 soil phylotypes (taxa that share 100% sequence similarity across the amplified 16S rRNA gene for soil bacteria, and 18S rRNA gene for soil fungi, protists and invertebrates). We use the term soil biodiversity to refer to these different types of richness when speaking in general terms. We also obtained data for a set of 11 ecosystem functions (stocks and processes) that are influenced by soil organisms and that correspond to key components of ecosystem services—nutrient cycling, OM decomposition, plant net primary productivity (NPP), pathogen control (reduced relative abundance of potential fungal plant pathogens in soils) and control of antibiotic resistance genes (ARG; reduced abundance of soil ARGs). Together, these measurements represent core ecosystem functions that are both fundamental and quantifiable. We used four different metrics of richness (the most used and the simplest metric of biodiversity)[24,25]: the richness (that is, number of phylotypes or zero-radius operational taxonomic units (zOTUs)) within each of the 12 organismal types examined independently; a measure of their joint richness (using multidiversity indexes[8,14,25,26]); a measure of the richness of organismal types included within globally distributed ecological assemblages; and the richness of highly connected soil phylotypes within ecological networks. Given concerns regarding the interpretation of diversity metrics, we used multiple approaches to validate our findings. Thus, the results presented in this Article were robust to different analytical approaches to quantify multidiversity and multifunctionality.

Results

In soil samples from globally distributed ecosystems, we found significant positive relationships between the diversity of single groups of organisms and the multidiversity of all of the groups with averaging multifunctionality (Fig. 1). The richness of Ciliophora was the only exception, presenting a neutral relationship (Fig. 1). Importantly, the slope of the soil multidiversity–multifunctionality relationship was steeper than that of the richness of any individual type of soil phylotypes, and more variance was explained, suggesting that the diversity of multiple soil organisms fuels multifunctionality in terrestrial ecosystems (Fig. 2a). This positive association between soil biodiversity and multifunctionality was also found when using an alternative multifunctionality index that was weighted[26] by five groups of ecosystem services (plant productivity, ARG control, pathogen control, nutrient cycling and OM decomposition) such that functions from each ecosystem service contributed equally to multifunctionality[26] (Supplementary Fig. 2). Similarly, the relationship between soil biodiversity and multifunctionality was maintained when we used an alternative multidiversity index that was weighted equally by the four main groups of soil organisms included in this study (bacteria, fungi, protists and invertebrates; Supplementary Fig. 3). Our results from structural equation modelling (SEM; Supplementary Fig. 4, a priori model; Supplementary Table 2), performed as described previously[10], suggest that the positive effect of soil biodiversity on multifunctionality was maintained after accounting for key ecosystem factors such as geographic location, climate (temperature and aridity), plant attributes (perennial
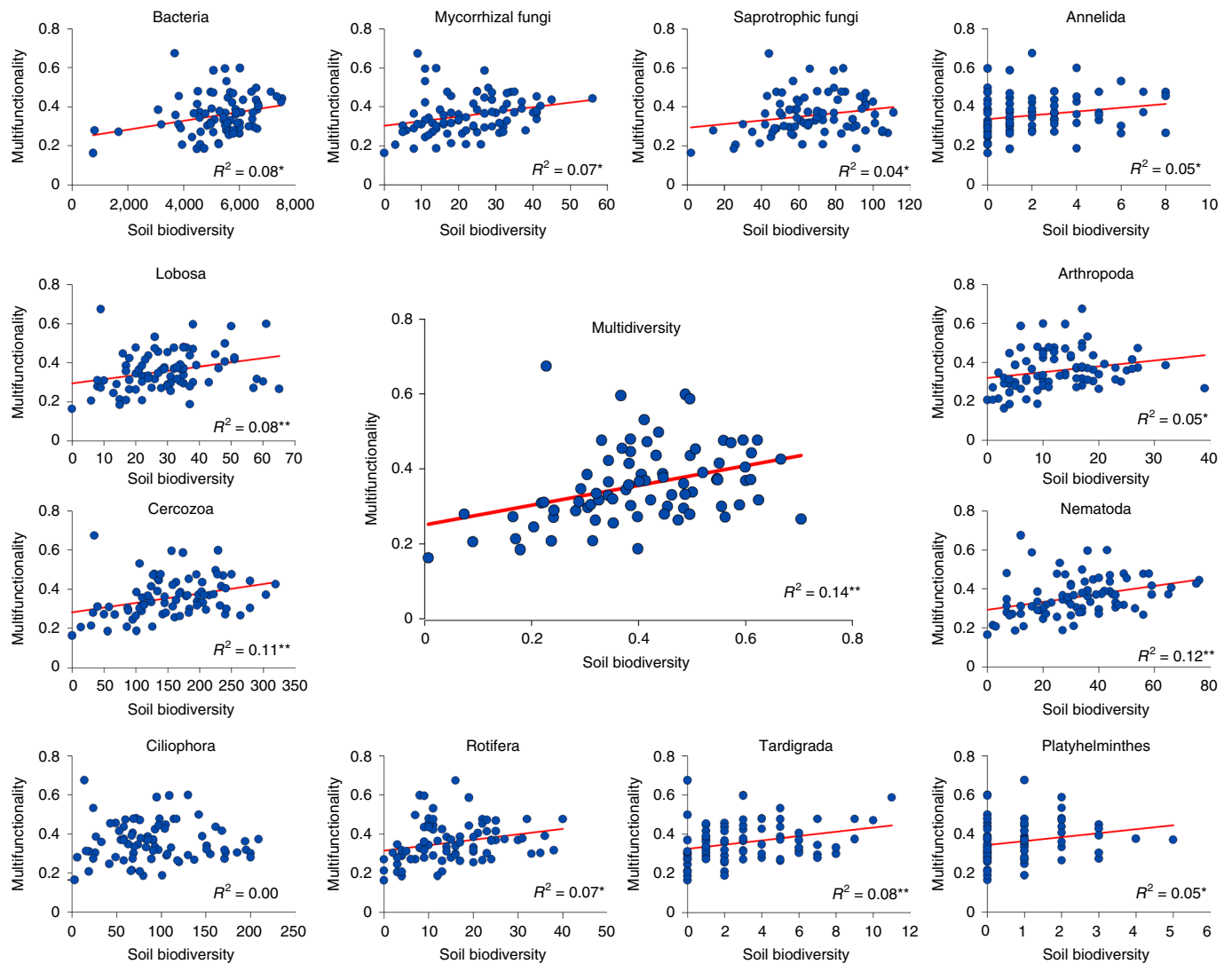
Fig. 1 | The relationship between multifunctionality and biodiversity of organisms. The linear relationships between multifunctionality and the biodiversity of selected groups of soil organisms (number of species, richness) or multidiversity (standardized between 0 and 1); $n = 81$ study sites. Statistical analysis was performed using ordinary least squares linear regressions; *$P \leq 0.05$; **$P < 0.01$.

plant richness and cover) and soil attributes (soil pH, total organic C and percentage of clay; Fig. 2b). The effects of plant diversity on multifunctionality were indirectly driven by changes in soil biodiversity (Fig. 2b). Our model goodness-of-fit was strong, indicating that patterns represent a causal scenario that is consistent with the data (Fig. 2b).

The positive association between soil multidiversity and multifunctionality was also observed for major biomes and ecosystem types when examined separately (Supplementary Fig. 5), and after accounting for sampling date in our statistical analyses (Spearman $\rho = 0.36$; $P < 0.001$)[24]. Moreover, our results were consistent, irrespective of multifunctionality index, including multiple single functions (Fig. 2c), the multithreshold approach[14] (Fig. 3, Supplementary Table 3) and multidimensional functionality[26] (Supplementary Table 4, Supplementary Fig. 6). In general, the richness of single soil organism types was consistently and positively correlated with multiple processes related to OM decomposition, reduced abundance of soil ARGs, nutrient cycling, plant productivity and reduced relative abundance of potential plant pathogens in soils (Fig. 2c) among the 12 soil group studies. For example, the positive relationship between soil biodiversity and lower abundance of the genes of ARGs

suggests that, in natural ecosystems at high ARG levels, lower diversity may be the result of outcompeting fast-growing highly competitive species through antibiotic production. Moreover, the diversity of nematodes (especially herbivores and bacterivores; Supplementary Table 5) and bacteria supported the highest number of single ecosystem functions (Fig. 2c). Soil biodiversity was also fundamental for maintaining the multiple dimensions of ecosystem functioning, mainly represented by plant productivity, OM decomposition, reduced abundance of ARGs (for example, as the result of the lack of fast growing highly competitive species) and enhanced nutrient cycling (Fig. 2c, Supplementary Table 4).

To provide a further test of the importance of soil biodiversity for ecosystem multifunctionality, we conducted a manipulative microcosm experiment using the dilution-to-extinction approach[27,28] with independent soil samples, at the local stand level. Our goal was to experimentally create a gradient of soil microbial diversity (Supplementary Fig. 7) while maintaining similar levels of microbial abundance (Supplementary Fig. 8) in independent soils from two eucalypt forests in eastern Australia[24]. Note that our study was not explicitly designed to provide a realistic expectation of biodiversity losses (such as those caused by soil degradation).
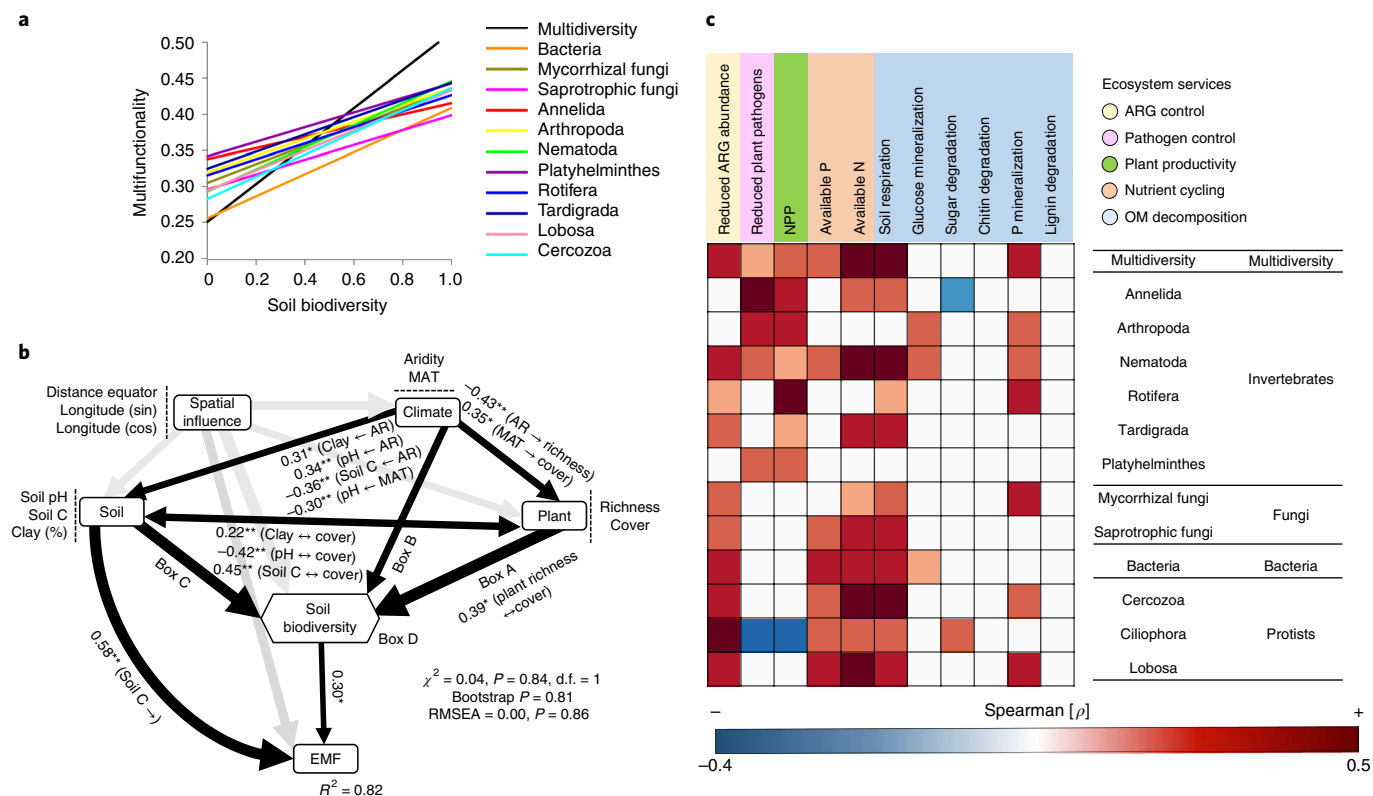
Fig. 2 | Links between soil biodiversity and ecosystem multifunctionality in a global field survey. a, The fitted linear relationships between average multifunctionality and the biodiversity of selected groups of soil organisms (number of species) and of a composite metric of their joint diversity (multidiversity; standardized between 0 and 1). Statistical analysis was performed using ordinary least squares linear regressions; $P \leq 0.05$; $n = 81$ study sites. **b**, Using a fitted SEM, we aimed to identify the direct relationship between the combined biodiversity of 12 groups of soil organisms and averaging ecosystem multifunctionality (EMF) ($n = 81$ study sites). We grouped the different categories of predictors (climate, soil properties, plants and spatial influence) into the same box in the model for graphical simplicity; however, these boxes do not represent latent variables. Soil biodiversity was included as a composite variable, including information about the biodiversity of 12 selected soil taxa. The rectangles are observable variables. Numbers adjacent to arrows are indicative of the effect size of the relationship. R^2 denotes the proportion of variance explained. Significance levels of each predictor (from SEM) are indicated by asterisks; $*P \leq 0.05$; $**P < 0.01$. AR, aridity; MAT, mean annual temperature; RMSEA, root mean square error of approximation. Information about boxes A–C and direct effects for other SEM arrows is provided in Supplementary Table 2. Information about our a priori model is provided in Supplementary Fig. 4 and Supplementary Table 2. **c**, Significant correlations (Spearman; $P \leq 0.05$) between the diversity of single groups of organisms and single ecosystem functions in the global field survey; $n = 81$ study sites.

In this microcosm, we assessed 8 out of the 11 key functions presented above, including N and P availability, P mineralization, chitin, sugar and lignin degradation, soil respiration and glucose mineralization, and their relationship with the diversity (richness of soil phylotypes) of microbial communities (fungi and bacteria)[24]. The results from this microcosm study provided independent and experimental verification of a significant and positive link between microbial richness (the number of phylotypes of fungi and bacteria) and multifunctionality (Fig. 4, Supplementary Figs. 9–11, Supplementary Table 6). We found that the positive effects of soil bacterial and fungal diversity on multifunctionality were independent of microbial abundance and community composition, as supported by partial-correlation analyses, which included community composition (first axis of a non-metric multi-dimensional scaling including the relative abundance of microbial taxa at the phylotypes level) and total abundance (measured using quantitative PCR) of fungi or bacteria (Supplementary Table 7).

The relationships between soil biodiversity and multiple functions at the global level depended on the type of organism and the identity and degree of connectivity of dominant soil phylotypes across globally distributed soil food webs. For example, the richness of larger soil invertebrates, such as tardigrades, annelids

(such as earthworms), platyhelminthes (flatworms) and arthropods, was especially positively associated with high functional thresholds (over 75% of their maximum observed levels of functioning; Fig. 3, Supplementary Table 3). By contrast, smaller soil taxa, such as bacteria, fungi, protists, and herbivorous and bacterivous nematodes, were positively associated with low functioning thresholds (<50% of their maximum rates/availabilities; Fig. 3, Supplementary Tables 3 and 5).

We next evaluated the importance of soil biodiversity for predicting multifunctionality within key ecological clusters using a global soil correlation network. These ecological clusters represent ecological assemblages of soil phylotypes that strongly co-occur. Note that one location can have more than one ecological cluster and that the number of phylotypes within these clusters differs across soil samples. We found five dominant ecological clusters that included >97% of the soil phylotypes that strongly co-occurred within the soil network (Fig. 5). Conceptually, clusters probably have similar ecological preferences, and can support similar functions. Taxa within a common cluster were more strongly correlated with other taxa within that cluster than with taxa from other clusters. A complete list of phylotypes within each ecological cluster is provided in Supplementary Table 8.
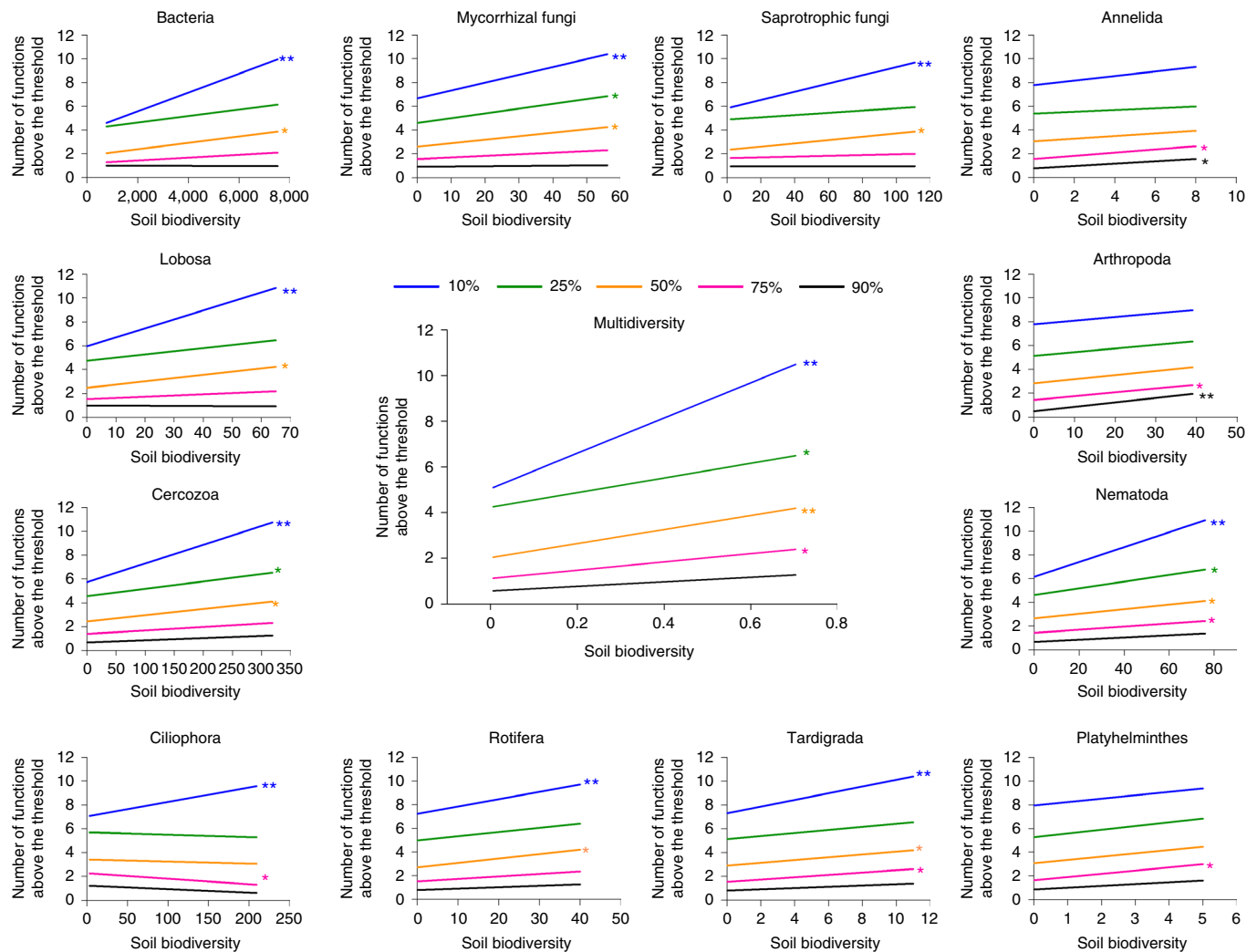
Fig. 3 | The relationship between multithreshold functioning and biodiversity of soil taxa. The relationships between multithreshold functioning and the biodiversity of selected groups of soil taxa (number of phylotypes) or a composite metric of their joint diversity (multidiversity; standardized between 0 and 1) in a global field survey; $n = 81$ study sites. Fitted linear regressions between the diversity of single groups of soil organisms and the number of functions above multiple thresholds. Different colours represent different thresholds of functioning. Statistical analysis was performed using ordinary least squares linear regressions; P values are indicated by asterisks: *$P \leq 0.05$; **$P < 0.01$.

As described above, the number of phylotypes within each ecological cluster changed across soil samples, as not all of the soil phylotypes occurred in every soil sample. We found a positive correlation between the richness of soil phylotypes within three of these ecological clusters (clusters 2, 4 and 5) and multifunctionality (Fig. 5, Supplementary Fig. 12). Nematode phylotypes were always present in those functionally important ecological clusters (Supplementary Table 8), and their richness was positively associated with multifunctionality (clusters 2 and 4; Fig. 5, Supplementary Figs. 12 and 13). We also tested the associations between the richness of soil phylotypes within the two dominant ecological clusters 2 and 4 and multifunctionality in our microcosm experiment[24], and also found positive associations between the richness of phylotypes within these ecological clusters and multifunctionality, providing independent evidence for the importance of these dominant soil phylotypes in regulating multifunctionality (Fig. 5, Supplementary Figs. 12 and 13, Supplementary Tables 9 and 10; taxonomic information on these soil phylotypes is provided in Supplementary Table 8). We also detected two additional ecological clusters (clusters 1 and 3; Supplementary Fig. 14), for which increases in the richness of soil phylotypes resulted in either no correlation

(cluster 3) or negative association (cluster 1, which included multiple Ciliophora taxa; Supplementary Table 8) with multifunctionality (Supplementary Fig. 14).

We identified the soil phylotypes that were highly connected with other phylotypes within the ecological network[24] (Fig. 5, Supplementary Figs. 15 and 16, Supplementary Tables 9 and 10). A total of 76 bacterial phylotypes were classified as hub phylotypes (as described previously[21]; Supplementary Figs. 15 and 16, Supplementary Tables 9 and 10). These phylotypes were highly connected among or within ecological clusters within our soil global ecological network. Interestingly, no fungal, protist or invertebrate phylotypes were selected as hub phylotypes. We found a strong and positive association between the richness of soil hub phylotypes and multifunctionality in both observational and microcosm studies (Fig. 5, Supplementary Fig. 13, Supplementary Tables 9 and 10). Finally, further statistical analyses suggested that the different soil biodiversity indices described above (multidiversity and diversity of taxa within ecological clusters and classified as hub phylotypes) are all important predictors of multifunctionality, and are needed to predict multiple ecosystem functions simultaneously (Supplementary Fig. 17).
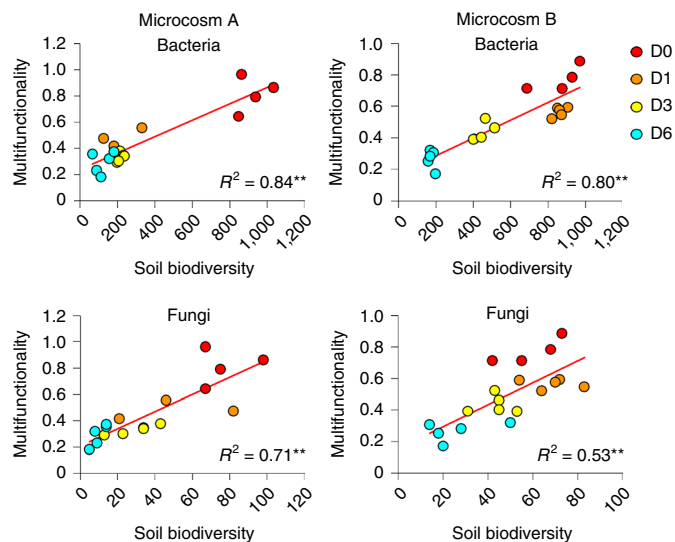
Fig. 4 | Linkages between soil biodiversity and ecosystem multifunctionality in a microcosm study. The linear relationships between the diversity of single groups of soil organisms (number of phylotypes) and average multifunctionality for microcosms of two soils (microcosms A and B) from eastern Australia. The different colours represent different dilutions from our dilution-to-extinction approach (D0–D6; $n = 5$ replicates). Statistical analysis was performed using ordinary least squares linear regressions; P values are indicated by asterisks: $**P < 0.01$.

Discussion

The importance of soil biodiversity as a major driver of multiple ecosystem functions is often assumed[1–6], yet is often undervalued, as microorganisms are usually regarded as highly functionally redundant in their environments[28]. However, the reality is that evidence for the link between cross-biome soil biodiversity and multiple ecosystem functions is lacking at the global scale, and experimental evidence linking soil microbial diversity to multifunctionality is scarce. Here, we provide solid evidence—from a global survey and a microcosm experiment—that multiple elements of soil biodiversity are necessary to maintain multiple ecosystem functions globally. In particular, we found a positive link between soil biodiversity and ecosystem functions across globally distributed biomes. Such positive associations were also observed for major biomes and ecosystem types (Supplementary Fig. 5) and when studying the associations between the diversity of individual taxa (bacteria, fungi, protists and invertebrates) and multiple individual functions (Fig. 2c). Our results further suggest that the effects of plant (perennial) diversity on multifunctionality, across contrasting biomes, are indirectly driven by changes in soil biodiversity (Fig. 2b) and by plant cover (plant cover ↔ plant richness SEM standardized effect = 0.39; $P < 0.001$; Supplementary Table 2). Moreover, from our microcosm experiment, we provide compelling experimental evidence that soil microbial diversity is positively associated with multifunctionality, with no evidence of functional redundancy in these relationships. Finally, our research highlights the importance of soil invertebrates, highly connected taxa and key globally distributed dominant phylotypes within the soil ecological network for simultaneously maintaining multiple ecosystem functions. Our study highlights the value of including soil biodiversity in the political and management agenda to protect the functioning of terrestrial ecosystems worldwide.

Our experimental tests support the observed soil-biodiversity–ecosystem-function relationships across terrestrial ecosystems using laboratory manipulations, which held most environmental sources of variation relatively constant. Notably, although the results of the global survey were consistent with the laboratory experiment results, associations between soil biodiversity and multifunctionality in this microcosm study were, as expected, always stronger than those in our global survey. This suggests that: (1) soil abiotic properties and climatic conditions do influence the biodiversity–ecosystem-function relationships (Fig. 2b); and (2) the observed relationships among soil biodiversity and functions that occur in nature can be a combination of direct diversity effects offset by covariance among other ecological factors that can covary with diversity, and can cause simultaneous positive and negative functional feedbacks.

Despite the overall positive relationships between soil biodiversity and multifunctionality, we also found that not all of the soil organisms were equally important for maintaining multifunctionality. First, our results indicated that diversity of larger soil invertebrates seems to be essential for maintaining multiple ecosystem functions operating at high levels of functioning (>75% threshold), meaning that locations with higher diversity of biodiversity of tardigrades, annelids (such as earthworms), platyhelminthes (flatworms) and arthropods support a greater number of functions working close to their highest (reported) levels of functioning (maximum rates/availabilities). For example, relatively large soil invertebrates comminute large amounts of animal and plant litter, regulating the flow of resources to microorganisms and, therefore, controlling the potential rates of multiple ecosystem functions. However, the biodiversity of smaller soil organisms, such as bacteria, fungi and protists, has a major role in supporting multiple ecosystem functions working at low levels of functioning (<50% of their maximum rates/availabilities). These results support the idea that larger invertebrates are especially important for maintaining multiple soil functions operating near peak capacity, whereas smaller invertebrates are critical for the fine-tuning of multifunctionality (for example, by nutrient recycling). Moreover, we found multiple potential associations between the biodiversity of soil organisms that might be positively influencing ecosystem multifunctionality. For example, the biodiversity of nematodes and protists was positively associated with bacterial diversity suggesting potential predator–prey associations (Supplementary Table 3) that could potentially positively influence multifunctionality.

We further investigated the importance of dominant taxa within the food web as controllers of ecosystem multifunctionality and found significant positive associations among the richness of soil phylotypes within three of these ecological clusters (clusters 2, 4 and 5) and multifunctionality (Fig. 5, Supplementary Fig. 11). In other words, soils that have a larger number of phylotypes belonging to these three ecological clusters (Supplementary Table 3) also have greater levels of multifunctionality. Importantly, we found that nematode phylotypes were always present in these functionally important ecological clusters. Nematodes have recently been reported to have a considerable role in controlling carbon fluxes in terrestrial ecosystems across the globe[6]. Notably, we also detected two additional ecological clusters (clusters 1 and 3; Supplementary Fig. 14), for which increases in the richness of soil phylotypes resulted in either no correlation (cluster 3) or negative association (cluster 1, which included multiple Ciliophora taxa; Supplementary Table 8) with multifunctionality (Supplementary Fig. 14). Thus, these soil phylotypes might not contribute appreciably to multifunctionality. This result suggests that it is crucial to know the identity of the phylotypes within soil ecological clusters to understand biodiversity–function relationships, and ultimately to challenge the common misconception that all biodiversity is equally needed to maintain ecosystem functioning. Nonetheless, the richness of soil phylotypes within ecological clusters 1 and 3 was positively correlated with specific groups associated with nutrient cycling, OM decomposition and reduced abundance of antibiotic resistance genes, suggesting that phylotypes included within these
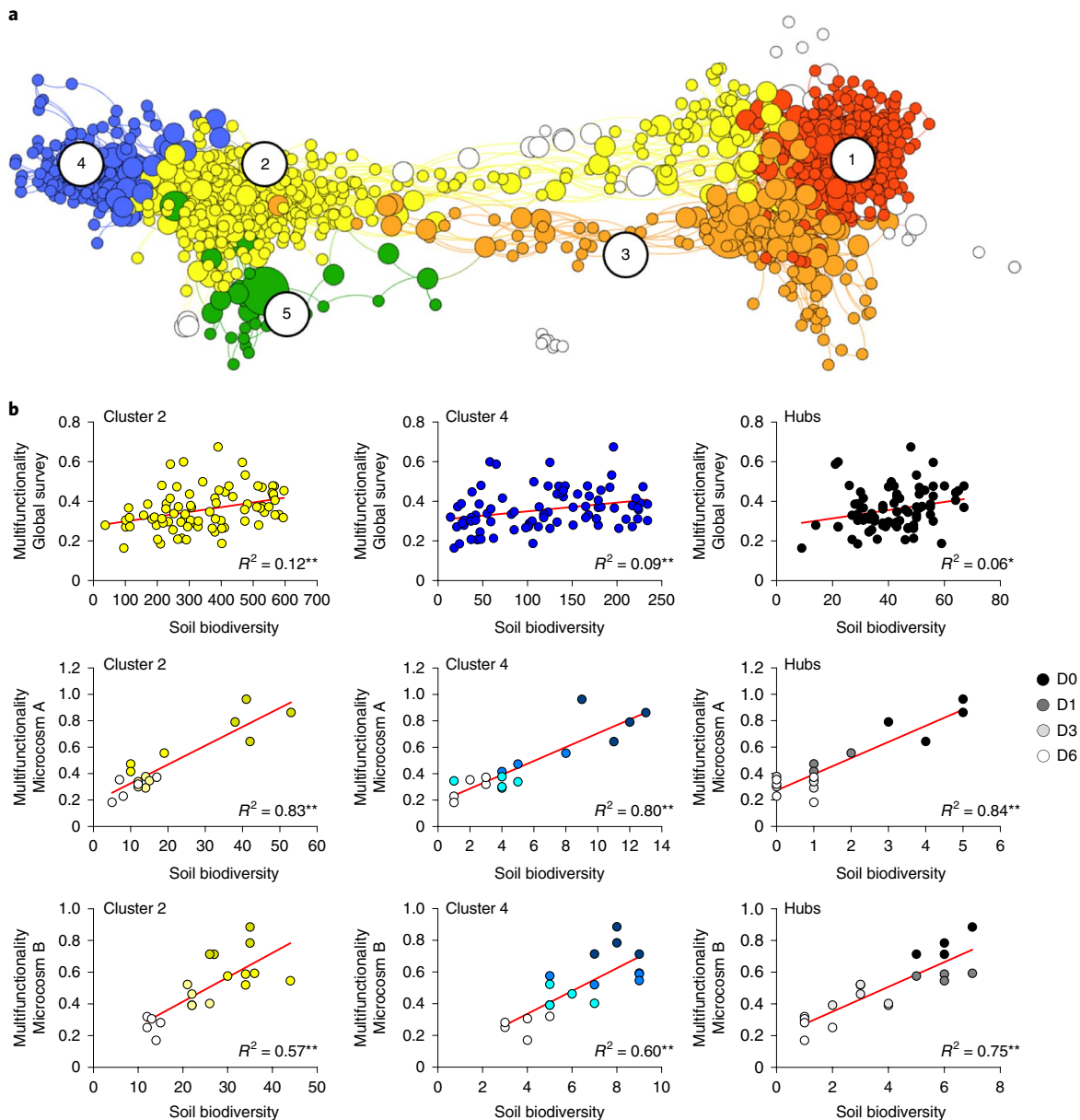
Fig. 5 | Linkages between the soil biodiversity within ecological networks and multifunctionality. a, Correlation network including multiple ecological clusters. **b**, The linear relationships between average multifunctionality and the diversity (number of phylotypes) of soil phylotypes within ecological clusters 2 and 4 and highly connected hub phylotypes within a global-scale soil ecological network; $n = 81$ study sites. Microcosms A and B were conducted in two different soils from eastern Australia. The different colours represent different dilutions from our dilution-to-extinction approach (D0–D6; $n = 5$ replicates). Statistical analysis was performed using ordinary least squares linear regressions; P values are indicated by asterisks: *$P \leq 0.05$; **$P < 0.01$.

ecological clusters are important drivers of ecosystem functioning (Supplementary Tables 9 and 10).

Finally, our research provides further evidence that the level of connectivity of taxa within the soil food web strongly influences ecosystem multifunctionality. In particular, we found that the richness of highly connected (hub) phylotypes within the ecological network was positively associated with multiple ecosystem functions in soils across the globe as well as in our microcosm experiment. Highly connected and globally distributed bacteria constituted the foundation for the soil food webs from our sites across the globe. Hub phylotypes contained some functionally important phylotypes from the order Nitrospirales, family Beijerinckiaceae, genus *Pedomicrobium* and family Methylocystaceae (Supplementary Table 8), and are known to include soil phylotypes that are involved in important soil processes such as nitrification, free-living N_2 fixation, biofilm

formation and methane consumption, respectively. Hub phylotypes also included multiple phylotypes from orders Actinomycetales and Rhizobiales and the phyla Verrucomicrobia, which have been previously postulated to be potential keystone taxa[29]. Critically, we found a strong and positive association between the richness of soil hub phylotypes and multifunctionality in both observational and microcosm studies (Fig. 5, Supplementary Fig. 13).

Conclusions

Our findings provide observational and experimental evidence that soil biodiversity is critically important for maintaining ecosystem function across the globe. It should be noted that we see similar patterns for single metrics of diversity or function as with those that are combined into multimetrics; this is true in both our cross-continent study and the manipulated experiment. Furthermore, our results

highlight the fact that, although the positive relationship between soil biodiversity and multifunctionality is a general one, the specific nature of this relationship depends on the type of soil organisms and on the identity and degree of connectivity of dominant soil phylotypes within the food web. Our results indicate that the richness of larger soil invertebrates (such as annelids, arthropods, tardigrade and flatworms) is especially important for maintaining multiple soil functions operating near peak capacity. Moreover, our findings provide evidence that a subset of globally distributed dominant phylotypes co-occurring within food webs is critically important for maintaining multiple ecosystems functions across the globe. Finally, highly connected phylotypes within ecological networks were found to be especially important for maintaining multiple ecosystem functions. Together, our research represents an important step for soil biology and ecosystem ecology. Our collective results suggest that multiple ecosystem functions and services that are supported by soil biodiversity should not be overlooked, as they probably have key roles for human wellbeing and ecosystem sustainability. Locally and across biomes, increasing knowledge of soil biodiversity could provide an emerging cornerstone for biodiversity, conservation and, with time, become a key component of management decision making.

Methods

Global survey. *Field survey.* Soil and vegetation data were collected between 2016 and 2017 from 83 locations across five continents (Supplementary Fig. 1). The field survey was designed to include globally distributed locations spanning a wide range of climate (tropical, temperate, continental, polar and arid) and vegetation types (including grasslands, shrublands, forests and forblands). By doing so, we aimed to maximize the inclusion of a wide range of environmental conditions (such as edaphic characteristics; examples are provided Supplementary Fig. 18), soil biodiversity and ecosystem functioning. Field surveys were conducted according to a standardized sampling protocol[25]. In each location, we surveyed a 50 m × 50 m plot using 3 parallel transects of the same length, spaced 25 m apart. The cover of perennial vegetation was measured in each transect using the line–intercept method[25]. Perennial plant richness (number of species) was estimated at the plot level. Our sampling design covered wide gradients in key environmental factors. For example, mean annual temperature at our sites was between −1.8 °C and 21.6 °C and mean annual precipitation was between 104 mm and 2,833 mm. Plant cover ranged between 0% and 100%, pH ranged from 3.19 to 9.45 and soil C ranged from 0.3 to 473.6 g kg⁻¹, providing a good representation of the most common environmental conditions found on Earth.

Soil sampling. Our sampling was explicitly designed to assess soil biodiversity and ecosystem functions at the plot level (50 m × 50 m resolution; Supplementary Fig. 19). Five composite topsoil samples from five 0–10 cm soil cores were collected under the dominant vegetation within each location, meaning that 25 cores were collected in each plot, and five composite samples were analysed for functions and soil biodiversity. A total of 415 soil samples were analysed in this study. We calculated site-level estimates of soil biodiversity and ecosystem functions as explained below.

After field sampling, soils were sieved (<2 mm) and separated into two portions. After soil sampling, one portion was air-dried and used for soil biochemical analyses. The second portion of soil was immediately frozen at −20 °C for molecular analyses. This storage approach is commonly used in global surveys[25,30]. Frozen soil sample (10 g; from composite soil samples as described above) was ground using a mortar and liquid N, aiming to homogenize the soils and obtain a representative sample for sequencing analysis.

Soil properties. Soil properties were determined using standardized protocols[25]. pH was measured in all of the soil samples using a pH metre, in a 1:2.5 mass:volume suspension of soil and water. Total organic carbon in soil was determined as described previously[25]. Texture (percentage of clay) was determined on a composite sample from each site as described previously[31]. pH, C and clay content ranged between 4.1 and 9.1, 0.1% and 25.7%, and 0.1% and 23.4%, respectively.

Diversity measures. The diversity of soil bacteria, fungi, protists and invertebrates was measured by amplicon sequencing using an Illumina MiSeq platform. Soil DNA was extracted using the Powersoil DNA Isolation Kit (MoBio Laboratories) according to the manufacturer's instructions. A portion of the bacterial 16S and eukaryotic 18S rRNA genes were sequenced using the 515F/806R[32] and Euk1391f/EukBr[33] primer sets, respectively. Bioinformatics processing was performed using a combination of QIIME[20], USEARCH[34] and UNOISE3[35]. Sequences were clustered into soil phylotypes (known as zOTUs) using a 100% identity level. Annotation of

the representative sequences of zOTUs was performed against the Greengenes (16S gene) and PR2 (18S gene) databases[20,36]. The zOTU approach used here is expected to provide similar results to those using an OTU approach[37]. Before we calculated the richness of soil organisms (explained below), the zOTU abundance tables were rarefied at 5,000 (bacteria, 16S rRNA gene), 2,000 (fungi, 18S rRNA gene), 800 (protists, 18S rRNA gene) and 300 (invertebrates, 18S rRNA gene) sequences per sample to ensure an even sampling depth within each belowground group of organisms (Supplementary Fig. 20). Protists were defined as all eukaryotic taxa, except fungi, invertebrates (Metazoa) and vascular plants (Streptophyta). Note that not all of the samples passed our rarefaction cut-off. We obtained information for 81 out of 83 study sites. This information was used for the downstream analyses. The ranges of soil biodiversity are similar to those found in previous global studies[20,33]. Moreover, the choice of rarefaction level did not impact our results, as we found highly statistically significant correlations between the number of soil phylotypes of bacteria (rarefied at 5,000 versus 18,000 sequences per sample), fungi (rarefied at 2,000 versus 10,000 sequences per sample), protists (rarefied at 800 versus 4,000 sequences per sample) and invertebrates (rarefied at 300 versus 1,800 sequences per sample) (Pearson $r > 0.96$; $P < 0.001$) across different rarefaction levels. On average, bacterial communities were dominated by Proteobacteria, Actinobacteria and Acidobacteria; fungal communities were dominated by Ascomycota, Basidiomycota and Mucoromycota; protist communities were dominated by Cercozoa, Ciliophora and Lobosa; and invertebrate communities were dominated by Nematoda, Arthropoda and Rotifera in this order.

Here, we used richness (that is, number of soil phylotypes) as a metric of soil biodiversity. Richness is the most used, as well as the simplest, metric of biodiversity. Before calculating the richness of different groups of soil organisms, the information on the relative abundance of soil phylotypes (zOTU abundance tables) from five soil replicates (five composite samples per plot) was averaged. Using these averaged zOTU tables, we then calculated the richness of the 12 most prevalent prokaryotic and eukaryotic organisms in our soil samples: bacteria, mycorrhizal and saprophytic fungi, protists (Cercozoa, Ciliophora and Lobosa) and invertebrates (Annelida, Arthropoda, Nematoda, Rotifer, Tardigrada and Platyhelminthes). This approach enabled us to obtain site-level estimates of the total number of phylotypes within each 50 m × 50 m plot. Even so, we highlight the potential limitation of sequencing approaches for quantifying the biodiversity of soil invertebrates; larger soil organisms are possibly underrepresented with this approach. The identity of saprophytic and mycorrhizal fungi, and animal predator, herbivore and bacterivore nematodes were identified using FUNguild and NEMAguild, respectively[38]. We used only highly probable and probable guilds for these analyses. Moreover, we focused on those taxa with an identified single trophic mode.

Importantly, the richness of soil bacteria, fungi, protists and invertebrates was highly correlated with Shannon diversity in all cases (Pearson $r = 0.80–0.95$; $P < 0.001$). Moreover, the richness of soil bacteria, fungi, protists and invertebrates calculated at the plot scale (from averaged zOTU tables) was highly correlated with the richness of soil organisms calculated as the average of five soil replicates (Pearson $r = 0.88–0.93$; $P < 0.001$). These analyses suggest that the choice of diversity metric do not alter our results.

Ecosystem functions. Eleven ecosystem functions regulated by soil organisms and belonging to a wide range of ecosystem services were included in this study: nutrient cycling (soil N and P availability), OM decomposition (soil extracellular enzyme activities related to P mineralization, chitin and sugar degradation, and also measurements of lignin degradation, soil respiration and glucose mineralization), primary production (aboveground NPP) and pathogens (reduced relative abundance of fungal plant pathogens in soils), and ARG control (reduced abundance of antibiotic resistance genes in soils). In all of the soil samples, the availability of N (ammonium and nitrate) and P was obtained from K_2SO_4 and bicarbonate extracts, respectively, using colorimetric assays as described previously[39]. The measure of available P used here (Olsen P) was significantly positively correlated with other commonly used measures of soil P (resin-P) (Spearman $\rho = 0.64$; $P < 0.001$), suggesting that the choice of available P did not influence our results. The activities of β-glucosidase (sugar degradation), N-acetylglucosaminidase (chitin degradation) and phosphatase (P mineralization) were measured from 1 g of soil using fluorometry as described previously[40]. Moreover, we used the MicroResp approach[41] to measure lignin-induced respiration (calculated from basal respiration measurements using this method). The total abundance of 285 unique ARGs encoding resistance to all of the major categories of antibiotics was obtained using the high throughput qPCR, from soil samples as described previously[42]. The inversed abundance of ARGs (reduced abundance of ARGs) was obtained by calculating the inverse of this variable (−1 × total abundance of ARGs). Antibiotic resistance regulates soil processes such as microbial competition and productivity[30], and is important in natural ecosystem at the large spatial scale[42]. The relative abundance of potential fungal plant pathogens in soils was obtained from the amplicon sequencing analyses (as explained above) and was inferred by parsing the soil phylotypes using FUNguild[38]. We used only highly probable and probable guilds for these analyses. The inverse abundance (reduced relative abundance) of potential fungal plant pathogens was obtained by calculating the inverse of this variable (total

relative abundance of fungal plant pathogens × −1). Soil respiration (the basal flux of CO_2) as well as glucose-C mineralization were estimated in a composite soil sample per plot using an isotope approach. In brief, two parallel sets of dry soil samples (1 g) were placed in 20 ml glass vials at 50% of the water-holding capacity, sealed with a rubber septum and pre-incubated for 1 week at 28 °C in the dark. During this time, microorganisms readapted to the water conditions and released a pulse of CO_2 due to the new moisture conditions. After that, the glass vials were opened and the atmosphere was refreshed. The mineralization of fresh C (glucose mineralization) was assayed by adding ^{13}C-glucose (99 atom% U-^{13}C, Cambridge Isotope Laboratories) dissolved in water to one of the vial series at a dose of 250 µg of glucose-C per gram of soil (which is commonly used in incubation studies)[43–46]. In parallel, the second sample set was processed using the same procedure, adding water without glucose; this sample set was used for measuring soil respiration rates. Soils were then incubated for 16 d at 28 °C in the dark. After incubation, 4 ml of headspace gas from each vial was transferred to pre-evacuated glass vials (Labco), and the quantity and isotopic composition of released CO_2 was then determined. Soil respiration and glucose-C mineralization were estimated from these analyses. We used the normalized difference vegetation index (NDVI) as our proxy for NPP during the sampling dates. This index provides a measure of the 'greenness' of vegetation across Earth's landscapes. NDVI data were obtained from the Moderate Resolution Imaging Spectroradiometer aboard NASA's Terra satellites at a resolution of 250 m. The NDVI during the sampling dates was highly correlated with monthly averages for this variable during the 2008–2017 period (Spearman $\rho = 0.83$; $P < 0.001$), suggesting that the choice of productivity period should not alter our results.

Microcosm study. *Field survey and soil sample collection.* This microcosm study was conducted in soils independent from the global survey presented above, explaining the slight methodological differences between these two studies, and enabled us to test relationships between soil diversity and function independently of the data used to assess the global patterns. This microcosm experiment further enabled us to account for any effects of community composition and abundance of fungi and bacteria in our conclusions.

Soil sampling was performed in March 2014 at two locations in eastern Australia (microcosm A: New South Wales 33.9867° S, 145.7115° E; and Microcosm B: New South Wales, 33.7035° S, 148.2612° E) with contrasting precipitation regimes—an important environmental factor that often leads to contrasting environmental conditions[25]. Soil samples were collected from the top 10 cm. Locations were both open forests dominated by *Eucalyptus* spp., and were selected owing to their contrasting precipitation regimes: 400 (site A) and 657 mm (site B). The percentage of clay and total soil organic C and pH (estimated as described above) were 32% and 37%, 1.7% and 1.8%, and 6.0 and 5.6 for soils obtained from sites A and B, respectively.

Microcosm preparation. Soil samples from each site were sieved to <2 mm and divided in two portions: (1) soil for sterilization and (2) soil for microbial inoculum and experimental controls (non-sterilized original soils). The first portion was sterilized using a double dose of gamma radiation (50 kGy each) at ANSTO Life Sciences facilities. Gamma radiation was used as it is known to cause minimal changes to the physical and chemical properties of soils compared with other methods, such as autoclaving[47,48]. The dilution-to-extinction approach was used to prepare soil microcosms[27,28]. A parent inoculum suspension was prepared by mixing 25 g soil in 180 ml of sterilized PBS. The mixture was vortexed at high speed for 5 min to mix the contents. The sediment was then allowed to settle for 1 min and serial dilutions were prepared from the suspension. For each soil (soils A and B), five dilutions were used as the microbial inoculum to create a diversity gradient. These dilutions were: undiluted (10^0; Dx), 1/10 dilution (D1), $1/10^3$ dilution (D3) and $1/10^6$ dilution (D6). A total of 40 microcosms (500 g each; 4 dilutions × 5 replicates × 2 soil types) were prepared. The moisture contents in these microcosms were adjusted to 50% water holding capacity to allow microbial activities to be maintained (by adding sterile water if needed) during the incubation period. These microcosms were established under sterile conditions; aseptic techniques were used throughout the experiment to avoid contamination.

Soil microcosms were incubated at 20 °C for 6 weeks for microbial colonization and biomass recovery as described previously[28]. Microcosms with the highest dilution are expected to have the lowest microbial biomass initially, which may affect any interpretation regarding the relationship between microbial diversity and ecosystem functioning. Biomass recovery is needed to properly address the link between soil microbial diversity and ecosystem functioning by controlling for biomass interferences. We therefore started measuring soil microbial diversity and functions only after the microbial biomass had recovered across all dilutions of the microcosm (Supplementary Fig. 6).

Diversity measurement. Total genomic DNA was extracted using the MoBio PowerSoil DNA Isolation Kit (MoBio Laboratories) according to the manufacturer's instructions. To quantify the abundance of bacteria and fungi in our microcosms, and then be able to statistically account for any effect of microbial biomass on our biodiversity–function conclusions, the abundances of total bacteria (using the 16S rRNA gene; primer set Eub338/Eub518) and fungi (using the

internal transcribed spacer region (ITS); primer set ITS1-5.8S) were quantified using a CFX-96 thermocycler (Bio-Rad) as described previously[48]. Standard curves were generated using tenfold serial dilutions of plasmids containing the correct insert of each gene. The diversity of soil bacteria and fungi was measured using amplicon sequencing with the Illumina MiSeq platform. Bacterial 16S rRNA gene and fungal ITS region were sequenced using the 341F/805R and FITS7/ITS4 primer sets[10], respectively. Bioinformatics and rarefaction analyses were performed as described above for the cross-biome study. Note that not all of the samples passed our rarefaction cut-off. We obtained information for 17 out of 20 microcosms for soil A, and in 19 out of 20 microcosms for soil B. We calculated the richness of bacteria and fungi in each soil replicate from rarefied zOTU tables.

Ecosystem functions. Eight out of the eleven functions explained above were available for this microcosm study including N and P availability, P mineralization, chitin degradation and glucose mineralization, lignin degradation, soil respiration and glucose mineralization. All of the functions except for soil respiration and glucose mineralization were measured as described above. In the case of glucose mineralization, here we used the MicroResp approach[41] to measure glucose-induced respiration (calculated from basal respiration measurements using this method). Soil respiration (CO_2 fluxes) was monitored by placing 20 g of soil from each microcosm into a glass jar (12 cm depth, 75 cm diameter, Ball), and then sealed with a gas-tight lid, which had a rubber stopper in the middle. Gas samples were collected in 25 ml gas-tight syringes at 0 min, 30 min and 60 min after sealing. Soil gas flux for CO_2 was measured using an Agilent-7890a gas chromatograph (Agilent Technologies). Soil respiration was estimated from these analyses.

Ecosystem multifunctionality and multidiversity. To obtain a quantitative multifunctionality index for each site from the global survey and replicate from the microcosm study, we used four independent multifunctionality approaches: (1) the averaging multifunctionality index[25], (2) the multithreshold multifunctionality index[14], (3) multiple single functions and (4) the principal coordinate multifunctionality index[26]. To obtain an averaging ecosystem multifunctionality index, we first standardized all individual ecosystem functions between 0 and 1 (rawFunction − min(rawFunction)/(max(rawFunction) − min(rawFunction)) and then calculated their average. A similar approach was used to calculate multidiversity (using the richness of individual groups of soil organisms). In the case of the global survey, before this analysis, we averaged the soil variables observed in the five replicates (five composite samples per plot) collected within each plot to obtain site-level estimates. This multidiversity index is largely used and accepted in the current biodiversity–function literature[2,8,11].

Moreover, we used multifunctionality (multiple individual functions and using three state-of-the-art multifunctionality indices)[14,25,26] to denote both a set of functions examined individually and their joint actions when described with a single multifunctionality index, and do not argue that one is better or more appropriate than the other. The multithreshold approach[14] evaluates the linkage between biodiversity and the number of functions (rate or availability) that simultaneously exceed a critical threshold (>10%, 25%, 50%, 75% and 90% of the maximum observed level of functioning for a given function). Finally, for the global survey, we used principal coordinate analyses to identify the different dimensions of multifunctionality[26].

To obtain a multidiversity index[8], we first standardized the site-estimated richness of each soil group between 0 and 1, and then averaged them so that the richness of each soil group contributed equally to this multidiversity index. In general, the 11 functions and the 12 soil biodiversity (richness of bacteria, fungi, protists and invertebrates) indices included in the averaging index were not strongly multicollinear ($r < 0.8$).

Statistical analyses. *Linking soil biodiversity to multifunctionality.* We first conducted ordinary least squares linear regressions between soil multidiversity (standardized averaged of the diversity of 12 soil organisms) and single soil organisms with multifunctionality, multidimensional functioning (axes of a principal coordinate analysis including 11 functions) and the number of functions above the threshold. We then conducted Spearman correlations between the diversity of single soil organisms and single functions. In the global survey, and to account for any influence of sampling dates in our statistical analyses, we conducted an analysis of variance (ANOVA) using sampling year, season (summer, spring, winter and fall) and trimester (1, January–March; 2, April–June; 3, July–September; and 4, October–December) as fixed factors and multifunctionality as a response variable. We then correlated (Spearman) the residuals of this ANOVA (the portion of variation in multifunctionality that was not explained by sampling date) with multidiversity.

SEM. We used SEM[10] to evaluate the direct link between soil biodiversity and multifunctionality (averaging) in our global survey after accounting for multiple key ecosystem factors such as spatial influence (distance from equator and sine and cosine of longitude), climate (mean annual temperature and aridity), plant (richness and cover) and soil (soil pH, total organic C content and percentage of clay) attributes simultaneously (Supplementary Fig. 4, a priori model; Supplementary Table 2). Mean annual temperature and aridity index (AI = precipitation/potential evapotranspiration) were obtained from data

derived from WorldClim (http://www.worldclim.org) at a resolution of 1 km. Aridity was calculated as the inverse of the AI ($-1 \times$ AI). A useful characteristic of SEM for the purposes of our investigation is its utility for partitioning the effects that a variable may have on another, and for estimating the strengths of these multiple effects. In contrast to regression or ANOVA, SEM offers the ability to separate multiple pathways of influence and view them as parts of a system and is therefore useful for investigating the complex relationships among predictors that are commonly found in natural ecosystems[10]. All of the variables were included as independent observable variables. The diversity of 12 soil organisms was included as a composite variable in our SEM model because, together, they determine ecosystem multifunctionality. The use of composite variables does not alter the underlying SEM model, but collapses the effects of multiple conceptually related variables into a single composite effect, aiding interpretation of model results. Moreover, we identified curvilinear relationships between environmental factors and multifunctionality (Supplementary Fig. 21). We found that multifunctionality was associated with aridity in a hump-shaped manner, and that this relationship was well described by a second-order polynomial. To introduce polynomial relationships into our model, we calculated the square of aridity and introduced it into our model using a composite variable approach as described above. SEM models were conducted using the software AMOS 20 (IBM SPSS).

Correlation networks. To identify ecological clusters of strongly associated soil taxa, including unique soil phylotypes, a correlation network—that is, a co-occurrence network—was established. We conducted these analyses with 81 globally distributed locations for which we have information on soil organisms. We used the site-level estimated zOTU tables described above for these analyses. We focused on the most dominant phylotypes—those that were both abundant (top 10% of all identified prokaryotes and eukaryotes in terms of relative abundance) and ubiquitous (>25% of all locations) across all globally distributed soils, and identified ecological clusters of strongly co-occurring soil phylotypes within this network. Using this filtering, we aimed to reduce potential spurious correlation from the rare taxa. We used a definition of dominant phylotype explained previously[20] to apply an additional constraint to ensure that we identified dominant phylotypes. Although many bacterial taxa are globally distributed[20], this is unlikely to be the case for eukaryotic organisms. Owing to this, we applied a ubiquity threshold of >25%. We focused on these dominant soil phylotypes because they are expected to have a disproportionate functional importance in their ecosystems, and are globally distributed, reinforcing the global perspective of our conclusions. Our network included 1,782 dominant soil phylotypes strongly co-occurring with each other. These soil phylotypes were dominated by 1,674 bacteria, 53 fungi, 77 protists and 5 nematodes.

We used a correlation cut-off of Spearman $\rho > 0.65$, $P < 0.001$, which is often used in the current literature, and is comparable across studies[18], to generate statistically robust correlations and control the false-positive rate as much as possible. We expected that this cut-off, which is frequently used in the microbial literature[18], would have both a mathematical and biological meaning, as we only focused on organisms that are strongly correlated with each other. Even so, we reinforce the notion that correlation network analyses are only a simplistic representation of a complex microbial system. Moreover, ecological networks that are based on correlations can yield spurious results, and associations between taxa within these networks cannot be directly interpreted as interactions. This is particularly true for microbial community data (on the basis of relative abundance) in which data (the relative abundance of different taxa) are not completely independent. However, the information derived from these networks is essential for generating novel hypothesis and ecological frameworks (to be tested in future experiments) about the role of highly connected taxa and dominant taxa within food webs in controlling multifunctionality.

The network was visualized using the interactive platform Gephi (https://gephi.org). We identified the ecological clusters and hub taxa within our ecological network using the R packages (https://cran.r-project.org/web/packages/) igraph[49] and brainGraph[50]. We then computed the richness of soil organisms within each ecological cluster, and that of highly connected soil taxa (classified as hubs; figure 2 in ref. [21]) across 81 globally distributed locations.

We also estimated the richness of dominant taxa within ecological clusters, as well as that of hub taxa within the ecological network, in our microcosm experiment to cross-validate our observational data using an independent approach. We focused on bacterial communities for these analyses because: (1) the 16S rRNA gene region amplified in both the observational (515F/806R) and experimental (341F/805R) study overlaps, enabling us to match (>97% similarity) representative sequences for bacterial soil phylotypes found in both databases; and (2) on the basis of global survey, bacterial taxa accounted for 94% of all taxa included in our correlation network (on the basis of our global survey), and was the only group of organisms that included highly connected (hub) taxa. We focused on the two dominant ecological clusters in our network (2 and 4; Fig. 4). About 70% of all bacterial taxa within ecological clusters 2 and 4 were present in our microcosm study (>97% similarity; Supplementary Table 9). Moreover, 71% of taxa classified as hub taxa were detected in our microcosm study (>97% similarity; Supplementary Table 9).

Semi-partial correlations. In our microcosm study, and to test for the influence of community composition and abundance in our biodiversity–function conclusions, we conducted partial correlation analysis between soil biodiversity and multifunctionality, accounting for microbial abundance (quantitative PCR data) and community composition (main axes of a non-metric multidimensional scaling analysis; see ref. [28] for a similar approach). We did not conduct these analyses for the observational database because obtaining absolute information for the abundance of all multiple soil taxa (bacteria, fungi, protist and soil invertebrates) at the global scale was not possible.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
Soil biodiversity and functional data from the global field survey and the microcosm experiment are publicly available in Figshare[51].

References
1. Holzman, D. C. Accounting for nature's benefits: the dollar value of ecosystem services. *Environ. Health Perspect.* **120**, a152–a157 (2012).
2. Wagg, C. et al. Soil biodiversity and soil community composition determine ecosystem multifunctionality. *Proc. Natl Acad. Sci. USA* **111**, 5266–70 (2014).
3. Van Elsas, J. D. et al. Microbial diversity determines the invasion of soil by a bacterial pathogen. *Proc. Natl Acad. Sci. USA* **109**, 1159–1164 (2012).
4. Bardgett, R. D. & van der Putten, W. H. Belowground biodiversity and ecosystem functioning. *Nature* **515**, 505–11 (2014).
5. Wall, D. H., Nielsen, U. N. & Six, J. Soil biodiversity and human health. *Nature* **528**, 69–76 (2015).
6. van den Hoogen, J. et al. Soil nematode abundance and functional group composition at a global scale. *Nature* **572**, 194–198 (2019).
7. Troudet, J. et al. Taxonomic bias in biodiversity data and societal preferences. *Sci. Rep.* **7**, 9132 (2017).
8. Allan, E. et al. Interannual variation in land-use intensity enhances grassland multidiversity. *Proc. Natl Acad. Sci. USA* **111**, 308–313 (2014).
9. Bradford, M. A. Discontinuity in the responses of ecosystem processes and multifunctionality to altered soil community composition. *Proc. Natl Acad. Sci. USA* **111**, 14478–14483 (2014).
10. Delgado-Baquerizo, M. et al. Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nat. Commun.* **7**, 10541 (2016).
11. Soliveres, S. et al. Biodiversity at multiple trophic levels is needed for ecosystem multifunctionality. *Nature* **536**, 456–9 (2016).
12. Hättenschwiler, S. & Gasser, P. Soil animals alter plant litter diversity effects on decomposition. *Proc. Natl Acad. Sci. USA* **102**, 1519–24 (2005).
13. García-Palacios, P. et al. Climate and litter quality differently modulate the effects of soil fauna on litter decomposition across biomes. *Ecol. Lett.* **16**, 1045–53 (2013).
14. Byrnes, J. E. et al. Investigating the relationship between biodiversity and ecosystem multifunctionality: challenges and solutions. *Meth. Ecol. Evol.* **5**, 111–124 (2014).
15. Geisen, S. The bacterial-fungal energy channel concept challenged by enormous functional versatility of soil protists. *Soil Biol. Biochem.* **102**, 22–25 (2016).
16. Bonkowski, M. Protozoa and plant growth. *New Phytol.* **162**, 617–631 (2004).
17. Menezes, A. B. et al. Network analysis reveals that bacteria and fungi form modules that correlate independently with soil parameters. *Environ. Microbiol.* **17**, 2677–2689 (2015).
18. Barberán, A. et al. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.* **6**, 343–351 (2012).
19. de Vries, F. T. et al. Soil bacterial networks are less stable under drought than fungal networks. *Nat. Commun.* **9**, 3033 (2018).
20. Delgado-Baquerizo, M. et al. A global atlas of the dominant bacteria found in soil. *Science* **359**, 320–325 (2018).
21. Guimerà, R. & Amaral, L. A. Functional cartography of complex metabolic networks. *Nature* **433**, 895–900 (2005).
22. Jens, M. & Olesen, J. M. et al. The modularity of pollination networks. *Proc. Natl Acad. Sci. USA* **104**, 19891–19896 (2007).
23. Bascompte, J. & Stouffer, D. B. The assembly and disassembly of ecological networks. *Proc. R. Soc. B* **364**, 1781–1787 (2009).
24. Gotelli, N. J. & Colwell, R. K. in *Biological Diversity: Frontiers in Measurement and Assessment* (eds Magurran, A. E. & McGill, B. J.) 39–54 (Oxford Univ. Press, 2011).
25. Maestre, F. T. et al. Plant species richness and ecosystem multifunctionality in global drylands. *Science* **335**, 214–218 (2012).
26. Manning, P. et al. Redefining ecosystem multifunctionality. *Nat. Ecol. Evol.* **2**, 427–436 (2018).

27. Philippot, L. et al. Loss in microbial diversity affects nitrogen cycling in soil. *ISME J.* **7**, 1609–19 (2013).
28. Delgado-Baquerizo, M. et al. Lack of functional redundancy in the relationship between microbial diversity and ecosystem functioning. *J. Ecol.* **104**, 936–946 (2016).
29. Banerjee S., Schlaeppi K. & van der Heijden, M. G. A. Keystone taxa as drivers of microbiome structure and functioning. *Nat. Rev. Microbiol.* **16**, 567–576 (2018).
30. Bahram, M., Hildebrand, F., Forslund, S. K., Anderson, J. L., Soudzilovskaia, N. A. & Bodegom, P. M. et al. Structure and function of the global topsoil microbiome. *Nature* **560**, 233–237 (2018).
31. Kettler, T. A. et al. Simplified method for soil particle-size determination to accompany soil-quality analyses. *Soil Sci. Soc.Am. J* **65**, 849–852 (2001).
32. Fierer, N. et al. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc. Natl Acad. Sci. USA* **109**, 21390–5 (2012).
33. Ramirez, K. S. et al. Biogeographic patterns in below-ground diversity in New York City's Central Park are similar to those observed globally. *Proc. R. Soc. B* **281**, 1795 (2014).
34. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460 (2010).
35. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10**, 996–998 (2013).
36. Guillou, L. et al. The protist ribosomal reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **41**, 597–604 (2013).
37. Glassman S. I. & Martiny J. B. H. Broadscale ecological patterns are robust to use of exact sequence variants versus operational taxonomic units. *mSphere* **3**, e00148-18 (2018).
38. Nguyen, N. H. et al. FUNGuild: an open annotation tool for parsing fungal community datasets by ecological guild. *Fung. Ecol.* **20**, 241–248 (2016).
39. Delgado-Baquerizo, M. et al. Decoupling of soil nutrient cycles as a function of aridity in global drylands. *Nature* **502**, 672–676 (2013).
40. Bell, C. W. et al. High-throughput fluorometric measurement of potential soil extracellular enzyme activities. *J. Vis. Exp.* **81**, e50961 (2013).
41. Campbell, C. D. et al. A rapid microtiter plate method to measure carbon dioxide evolved from carbon substrate amendments so as to determine the physiological profiles of soil microbial communities by using whole soil. *Appl. Env. Microbiol.* **69**, 3593–3599 (2013).
42. Hu, H.-W. et al. Diversity of herbaceous plants and bacterial communities regulates soil resistome across forest biomes. *Envir. Microbiol.* **20**, 3186–3200 (2018).
43. Bastida, F. et al. Phylogenetic and functional changes in the microbial community of long-term restored soils under semiarid climate. *Soil Biol. Biochem.* **65**, 12–21 (2013).
44. Derrien, D. et al. Does the addition of labile substrate destabilise old soil organic matter? *Soil Biol. Biochem.* **76**, 149–160 (2014).
45. Hopkins, F. M. et al. Increased belowground carbon inputs and warming promote loss of soil organic carbon through complementary microbial responses. *Soil Biol. Biochem.* **76**, 57–69 (2014).
46. Kuzyakov, Y. Priming effects: interactions between living and dead organic matter. *Soil Biol. Biochem.* **42**, 1363–1371 (2010).
47. Wolf, D. C. et al. Influence of sterilization methods on selected soil microbiological, physical, and chemical properties. *J. Environ. Qual.* **18**, 39–44 (1989).
48. Lotrario, J. B. et al. Effects of sterilization methods on the physical characteristics of soil: implications for sorption isotherm analyses. *Bull. Environ. Contam. Toxicol.* **54**, 668–675 (1995).
49. Csárdi, G. igraph, network analysis and visualization. R package version 1.2.2. R package (2018).
50. Watson, C. G. brainGraph, graph theory analysis of brain MRI data. R package version 2.2.0 (2018).
51. Delgado-Baquerizo, M. et al. Data from: Multiple elements of soil biodiversity drive ecosystem functions across biomes. *Figshare* https://doi.org/10.6084/m9.figshare.9976556 (2020).

Author contributions

M.D.-B., P.B.R. and B.K.S. developed the original ideas presented in the manuscript. M.D.-B. designed the global field study and coordinated all of the field operations. P.B.R., B.K.S. and M.D.-B. designed the microcosm experiment. Field data were collected by M.D.-B., C.T., D.J.E., S.A., F.D.A., A.A.B., N.A.C., A.G., L.G.-V., S.C.H., P.E.H., Z.-Y.H., M.K., S.N., C.A.P., S.C.R., F.S., B.W.S., J.-T.W., L.W.-G. and M.A.W. Functional analyses were performed by M.D.-B., A.G., L.G.-V., P.T., C.T., J.-Z.H., H.-W.H. and F.B. Bioinformatics analyses were performed by M.D.-B. and J.-T.W. Statistical modelling and network analyses were performed by M.D.-B. The first draft of the paper was written by M.D.-B., and further drafts were written by M.D.-B., P.B.R., D.J.E. and B.K.S., and all of the authors contributed to those subsequent drafts.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-019-1084-y.

Correspondence and requests for materials should be addressed to M.D.-B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Manuel Delgado-Baquerizo
M.delgadobaquerizo@gmail.com

Last updated by author(s): Nov 20, 2019

Reporting Summary

Nature Research 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 Research policies, see Authors & Referees 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	We conducted a global field survey (83 locations in five continents) and a microcosm experiment. Data collection information is included in the method section of our manuscript.
Data analysis	For Bioinformatic analysis, a combination of Qiime and USEARCH, UPARSE and UCLUST were used. For the Co-ocurrence Network, R (https://cran.r-project.org) was used for analyses, and Gephi for visualization. The rest of the analysis were made with R 3.4.0.

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/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

All the materials, raw data, and protocols used in the article are available upon request, and all data will be made publicly available in a public repository (Figshare) upon publication.

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

Ecological, evolutionary & environmental sciences study design

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

Study description	Here, we combined a global field study with microcosm experiments to provide a novel test of the linkages between soil biodiversity and multifunctionality.
Research sample	A total of 415 soil samples were analyzed in this study for soil biodiversity and ecosystem functions.
Sampling strategy	Our sampling was designed to assess soil biodiversity and ecosystem functions at the plot level (50 m × 50 m resolution). Five composite topsoil samples from five 0-10 cm soil cores were collected under the dominant vegetation within each location, meaning that twenty five cores were collected in each plot, and five composite samples were analyzed for functions and soil biodiversity. A total of 415 soil samples were analyzed in this study. We calculated site-level estimates of soil biodiversity and ecosystem functions as explained in the Material and Methods of our paper.
Data collection	We combined a global field survey (87 locations in five continents), across a wide range of ecosystem types, with a microcosm experiment. Our field study included information on ~45,000 soil species, and eleven ecosystem functions associated with multiple ecosystem services including nutrient cycling, organic matter decomposition, primary production and plant pathogen and antibiotic resistance gene control.
Timing and spatial scale	Sample collection of soils took place between 2014 and 2017
Data exclusions	N/A
Reproducibility	Information about the sampled locations and methods used in this paper are included in our material and methods
Randomization	N/A
Blinding	N/A

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions	We conducted a global field survey (83 locations in five continents) and a microcosm experiment. These global locations were selected to include a wide range of climates (tropical, temperate, continental, polar and arid) and vegetation types (including grasslands, shrublands, forests, and forblands) in order to represent the wide gradients of soil biodiversity and multifunctionality found across the globe.
Location	We conducted a global field survey (83 locations in five continents). See Extended Data Figure 1.
Access and import/export	Samples were collected by all authors in their respective locations and using local permits.
Disturbance	This study did not cause any environmental disturbance

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
☒	Animals and other organisms
☒	Human research participants
☒	Clinical data

Methods

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