Multiple elements of soil biodiversity drive ecosystem functions across biomes

Authors: 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[10], Laura García-Velázquez[10], Stephen C. Hart[8], Patrick E. Hayes[11-13], Ji-Zheng He[14,15], Zeng-Yei Hseu[16], Hang-Wei Hu[14,15], Martin Kirchmair[17], Sigrid Neuhauser[17], Cecilia A. Pérez[18], Sasha C. Reed[19], Fernanda Santos[8], Benjamin W. Sullivan[20], Pankaj Trivedi[21], Jun-Tao Wang[3], Luis Weber-Grullon[22-24], Mark A. Williams[25], Brajesh K. Singh[3,26].

Affiliations:

1. Departamento de Ecología and Instituto Multidisciplinar para el Estudio del Medio "Ramon Margalef", Universidad de Alicante, Alicante, Spain.
2. Departamento de Biología y Geología, Física y Química Inorgánica, Escuela Superior de Ciencias Experimentales y Tecnología, Universidad Rey Juan Carlos, Calle Tulipán Sin Número, Móstoles 28933, Spain.
3. Hawkesbury Institute for the Environment, Western Sydney University, Penrith, 2751, New South Wales, Australia.
4. Department of Forest Resources, University of Minnesota, St. Paul, MN 55108, USA.
5. School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales 2052, Australia.
6. GEMA Center for Genomics, Ecology & Environment, Universidad Mayor, Camino La Pirámide 5750, Huechuraba, Santiago, Chile.
7. CEBAS-CSIC. Department of Soil and Water Conservation. Campus Universitario de Espinardo, 30100, Murcia, Spain.
8. Department of Life and Environmental Sciences and Sierra Nevada Research Institute, University of California Merced, Merced, California, 95343 USA.
9. School of Geography, Politics and Sociology, Newcastle University, UK.
10. Departamento de Sistemas Físicos, Químicos y Naturales, Universidad Pablo de Olavide, 41013 Sevilla, Spain.
11. School of Biological Sciences, The University of Western Australia, Perth, WA, 6009, Australia.
12. Centre for Microscopy, Characterisation and Analysis, The University of Western Australia, Perth, WA, 6009, Australia.
13. Crop, Livestock and Environment Division, Japan International Research Centre for Agricultural Sciences, Tsukuba, Ibaraki, 305-8656, Japan.
14. Key Laboratory for Humid Subtropical Eco-geographical Processes of the Ministry of Education, School of Geographical Science, Fujian Normal University, Fuzhou 350007, China.
15. Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Parkville, Victoria 3010, Australia.
16. Department of Agricultural Chemistry, National Taiwan University, Taipei 10617, Taiwan.
17. Institute of Microbiology, University of Innsbruck, Innsbruck, Austria.
18. Instituto de Ecología y Biodiversidad, Las Palmeras 3425, Santiago, Chile.
19. U.S. Geological Survey, Southwest Biological Science Center, Moab, UT, USA.
20. Department of Natural Resources and Environmental Science, University of Nevada, Reno, NV, 89557, USA.

21. Department of Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, 80523, CO, USA.
22. Global Drylands Center, Arizona State University, Tempe, AZ, USA.
23. School of Life Sciences, Arizona State University, Tempe, AZ, USA.
24. School of Sustainability, Arizona State University, Tempe, AZ, USA.
25. School of Plant and Environmental Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA.
26. Global Centre for Land Based Innovation, Western Sydney University, Building L9, Locked Bag 1797, Penrith South, New South Wales 2751, Australia.

***Authors for correspondence:**
Manuel Delgado-Baquerizo. E-mail: M.DelgadoBaquerizo@gmail.com

Abstract

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 well-being and ecosystem sustainability. Combining a global observational study with an experimental microcosm study, we provide compelling 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 via soil biodiversity. Together our results provide insights into the importance of soil biodiversity for maintaining soil functionality locally, and across biomes, and strong support for the inclusion of soil biodiversity in conservation and management programs.

Introduction

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 dollars annually. However, as most soil microorganisms and micro fauna are difficult to observe directly, they are often neglected in global biodiversity surveys[7]. Consequently, the roles played by biodiverse soil organisms (bacteria, fungi, protists, and invertebrates; multidiversity; *sensu*[8]), for multiple kinds 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 allows for the synthesis of the many diverse functions 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 strong links between plant and soil biodiversity and function[6]. Moreover, observational studies within single biomes (e.g., 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, the relationship between the biodiversity of different groups of soil organisms (e.g., bacteria, fungi, protists, and invertebrates) and multiple functions has never been assessed 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 scarce. 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 critical for human well-being and global ecosystem sustainability.

It is also likely that different groups of soil organisms play different roles in maintaining multifunctionality. For instance, larger soil invertebrates (e.g., annelids, tardigrades, arthropods and flatworms) are responsible for processing large amounts of plant and animal litter and 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 (i.e., rates and availabilities) operating at high levels of functioning (relative to their maximum observed levels of functioning; *sensu*[14]). Conversely, the biodiversity of soil microbes (e.g., 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. Thus, we hypothesize that the smallest soil organisms are responsible for bottom up (producers) and top down (consumers) energy transfer via activating nutrients from the soil, and through predation, recirculating energy from larger organisms to smaller ones via 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 including 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 (i.e., taxa supporting 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. In addition, the degree of connectivity (e.g., determined via 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/or across ecological clusters (hub phylotypes), while 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 (e.g., 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 (e.g., 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, to our knowledge, non-existent yet could lend insights into how community structure determines function, and thus is in need of empirical study.

Here, we use a multi-continent observational field study and a controlled microcosm experiment to test the linkages among soil biodiversity and multifunctionality. First, we 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 twelve 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 ~45,000 soil phylotypes (taxa which 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 kinds of richness when speaking in general terms. We also obtained data for a set of eleven ecosystem functions (stocks and processes) influenced by soil organisms, which 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 antibiotic resistance genes (ARG) control (reduced abundance of soil ARGs). Together these measurements represent core ecosystem functions that are both fundamental and quantifiable. In this study, we use four different metrics of richness (the most used, and the simplest metric of biodiversity)[24-25]; the richness (i.e., number of phylotypes or 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 herein were robust to different analytical approaches to quantify multidiversity and multifunctionality.

Results

In soils from globally-distributed ecosystems, we found significant positive relationships between the diversity of single groups of organisms and the multidiversity of all 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 weighted[26] by five groups of ecosystem services (plant productivity, ARG control, pathogen control, nutrient cycling and OM decomposition), so that functions from each ecosystem service contributed equally to multifunctionality (Supplementary Fig. 2)[26]. Similarly, the relationship between soil biodiversity and multifunctionality was maintained when we used an alternative multidiversity index 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 Modeling (SEM; *a priori* model in Supplementary Fig. 4; Supplementary Table 2), as described in ref.[10], suggest the idea 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 plant richness and cover), and soil attributes (soil pH, total organic C and % of clay) (Fig. 2B). The effects of plant diversity on multifunctionality were indirectly driven via changes in soil biodiversity (Fig. 2B). Our model goodness-of-fit was strong, indicating that patterns represent a causal scenario 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 multi-threshold approach[14] (Fig. 3; Supplementary Table 3) and multidimensional functionality[26] (Table S4; 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 twelve soil group studies. For instance, 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 via 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). In addition, soil biodiversity was also fundamental for maintaining the multiple dimensions of ecosystem functioning, mainly represented by plant productivity, OM decomposition, reduced abundance of ARGs (e.g., 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]. Please, note that our study was not explicitly designed to provide a realistic expectation of biodiversity losses (e.g., by soil degradation). In this microcosm, we assessed eight of the eleven key functions presented above, including N and P availability, P mineralization, chitin, sugar and lignin degradation, soil respiration and glucose mineralization, and their relationship to the diversity (richness of soil phylotypes) of microbial communities (fungi and bacteria)[24]. Results from this microcosm study provide independent and experimental verification of a significant and positive link between microbial richness (number of phylotypes of fungi and bacteria) and multifunctionality (Fig. 4; Supplementary Figs. 9-11 and 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 an Non-metric Multi-Dimensional Scaling including the relative abundance of microbial taxa at the phylotypes level) and total abundance (measured via qPCR) 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 on the identity and degree of connectivity of dominant soil phylotypes across globally distributed soil food webs. For instance, the richness of larger soil invertebrates such as tardigrades, annelids (e.g., 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). Conversely, 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 then 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 strongly co-occurring within the soil network (Fig. 5). Conceptually, clusters are likely to 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 available in Supplementary Table 8. As noted above, the number of phylotypes within each ecological cluster changed across soil samples, as not all soil phylotypes occurred in every soil. 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 Fig. 12-13). 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 Fig. 12-13; Supplementary Tables 9-10; Supplementary Table 8 for taxonomic information on these soil phylotypes). 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).

Finally, we identified those soil phylotypes that were highly connected with other phylotypes within the ecological network[24] (Fig. 5; Supplementary Fig. 15-16; Supplementary Tables 9-10). A total of 76 bacterial phylotypes were classified as hub phylotypes (*sensu*[21]; Supplementary Fig. 15-16; Supplementary Tables 9-10). These phylotypes were highly connected among and/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 9-10). Finally, further statistical analyses suggested that the different soil biodiversity indices explained above (multidiversity, and diversity of taxa within ecological clusters and classified as hub phylotypes) are all important predictors of multifunctionality, and needed to predict multiple ecosystem functions simultaneously (Supplementary Fig. 17).

Discussion

The importance of soil biodiversity as a major driver of multiple ecosystem functions is often assumed[1-6], yet many times 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 a global scale, and experimental evidence linking soil microbial diversity to multifunctionality is scarce. Herein 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 (perennial) plant 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, we provide the most compelling experimental evidence, from a microcosm experiment, that soil microbial diversity is positively associated with multifunctionality, with no evidence of functional redundancy in these relationships. Finally, our work highlights the importance of soil invertebrates, highly connected taxa, and key globally-distributed dominant phylotypes within the soil ecological network for maintaining multiple ecosystem functions simultaneously. 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. Of note, although results of the global survey were consistent with the lab 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 (a) soil abiotic properties and climatic conditions do influence the biodiversity-ecosystem function relations (e.g., Fig. 2B), and (b) the observed relationships among soil biodiversity and functions that occur in nature can be a combination of direct diversity effects offset by co-variance among other ecological factors that can co-vary 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 soil organisms were equally important for maintaining multifunctionality. First, our results indicated that diversity of larger soil invertebrates seem 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 (e.g., earthworms), platyhelminthes (flatworms), and arthropods support a higher 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 microbes, and therefore, controlling the potential rates of multiple ecosystem functions. However, the biodiversity of smaller soil organisms such as bacteria, fungi and protists play 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, while smaller invertebrates are critical for the 'fine-tuning' of multifunctionality (e.g., via nutrient recycling). Moreover, we found multiple potential associations between the biodiversity of soil organisms which might be positively influencing ecosystem multifunctionality. For example, the biodiversity of nematodes and protists were positively associated with bacterial diversity suggesting potential predator-prey associations (Supplementary Table 3), which 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 having a larger number of phylotypes belonging to these three ecological clusters (Supplementary Table 3) also had 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 play an overwhelming role in controlling carbon fluxes in terrestrial ecosystems across the globe[6]. Strikingly, 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). Therefore, these soil phylotypes might not contribute appreciably to multifunctionality. This intriguing result suggests that it is crucial to know the identity of the phylotypes within soil ecological clusters in order 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 ecological clusters are important drivers of ecosystem functioning (Supplementary Tables 9-10).

Finally, our work 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, and 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 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 order Actinomycetales and Rhizobiales, and phyla Verrucomicrobia, which have been previously postulated as 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 and/or function as with those that are combined into multi-metrics; and this is true in both our cross-continent study and the manipulated experiment. Additionally, our results further 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 (e.g., 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 work represents an important step for soil biology and ecosystem ecology. Our collective results suggest that multiple ecosystem functions and services supported by soil biodiversity should not be overlooked, as they likely play key roles for human well-being 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.

Material and 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 (e.g., edaphic characteristics; examples in 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 three 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 instance, mean annual temperature at our sites ranged between -1.8 and 21.6 ºC, and mean annual precipitation between 104 mm and 2,833 mm. Plant cover ranged between 0 and 100 %, pH ranged from 3.19 to 9.45, and soil carbon (C) ranged from 0.3 to 473.6 g C 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 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 below.

Following 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 ºC for molecular analyses. This storage approach is commonly used in global surveys[25,30]. Ten grams of frozen soil sample (from composite soil samples as explained above) were ground using a mortar and liquid N aiming to homogenize soils and obtain a representative sample for sequencing analyses.

Soil Properties

Soil properties were determined using standardized protocols[25]. pH was measured in all the soil samples with a pH meter, in a 1: 2.5 mass: volume soil and water suspension. Soil total organic carbon was determined as described in ref.[25]. Texture (% of clay) was determined on a composite sample from each site according to ref.[31]. pH, carbon (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 via amplicon sequencing using the Illumina MiSeq platform. Soil DNA was extracted using the Powersoil® DNA Isolation Kit (MoBio Laboratories, Carlsbad, CA, USA) 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. Bioinformatic processing was performed using a combination of QIIME[20], USEARCH[34] and UNOISE3[35]. Sequences were clustered into soil phylotypes (aka zOTUs) using a 100% identity level. Annotation of the representative sequences of zOTU was performed against the Greengenes (16S

gene) and PR2 (18S gene) databases[20,36]. Before we calculate the richness of soil organisms (explained below), the zOTU abundance tables were rarefied at 5,000 (bacteria via 16S rRNA gene), 2,000 (fungi via 18S rRNA gene), 800 (protists via 18S rRNA gene), and 300 (invertebrates via 18S rRNA gene) sequences per sample, respectively, to ensure 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 samples passed our rarefaction cut-off. We obtained information for 81/83 plots. This information was used for the downstream analyses. The approach used here is expected to provide similar results to that one using Operational Taxonomic Units[37]. 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 vs. 18,000 sequences/sample), fungi (rarefied at 2,000 vs. 10,000 sequences/sample), protists (rarefied at 800 vs. 4,000 sequences/sample), and invertebrates (rarefied at 300 vs. 1,800 sequences/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.

In this study, we used richness (i.e., number of soil phylotypes) as our metric of soil biodiversity. Richness is the most used, and simplest metric of biodiversity. Prior to 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/plot) was averaged. Using these averaged zOTU tables, we then calculated the richness of the twelve 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 allowed us to obtain site-level estimates of the total number of phylotypes within each 50m × 50m plot. Even so, we would like to highlight the potential limitation of sequencing approaches for quantifying the biodiversity of soil invertebrates. Thus, clarify that the 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 only used high 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 to Shannon diversity in all cases (Pearson $r = 0.80\text{-}0.95$; $P < 0.001$). Moreover, the richness of soil bacteria, fungi, protists and invertebrates calculated at the plot scale (from averaged zOTU tables) was highly correlated to the richness of soil organisms calculated as the average of five soil replicates (Pearson $r = 0.88\text{-}0.93$; $P < 0.001$). These analyses suggest that the choice of diversity metric should 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), organic matter 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 net primary production; NPP) and pathogen (reduced relative abundance of fungal plant pathogens in soils), and ARG control (reduced abundance of antibiotic resistance genes in soils). In all soil samples, N (ammonium and nitrate) and P availability were obtained from K_2SO_4 and bicarbonate extracts, respectively using colorimetric assays as explained in ref.[39]. The measure of available P used here (Olsen P) was significantly positively correlated with other commonly used measure of soil P (resin-P) (Spearman $\rho = 0.64$; $P < 0.001$), suggesting the choice of available P cannot 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 by fluorometry as described in ref.[40]. In addition, 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 antibiotic resistance genes (ARGs) encoding resistance to all the major categories of antibiotics was obtained using the high throughput quantitative PCR (HT-qRCR) explained in ref.[42] from soil samples. 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 are 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 (explained above) and were inferred by parsing the soil phylotypes with FUNguild[38]. We only used 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 x -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 1 g dry soil samples were placed in 20-ml glass vials at 50% of the water-holding capacity, sealed with a rubber septum and pre-incubated for one week at 28ºC 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, 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, Tewksbury, Massachusetts, US) 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 subjected to the same procedure adding water without glucose; this sample set was used for measuring soil respiration rates. Soils were then incubated for 16 days at 28ºC in the dark. After incubation, 4 ml of headspace gas from each vial were transferred to pre-evacuated glass vials (Labco Limited, Lampeter, Wales, UK) 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 plant net primary productivity (NPP) during 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 (MODIS) aboard NASA's Terra satellites at 250-m resolution. The NDVI index during sampling dates was highly correlated to monthly averages for this variable between 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, which explains the slight methodological differences between these two studies, and allows us to test relationships between soil diversity and function independently of the data used to assess the global patterns. This microcosm experiment further allowed us to account for any effects of community composition and abundance of fungi and bacteria in our conclusions.

Soil sampling was carried out in March 2014 in two locations from Eastern Australia (Microcosm A: NSW 33.9867° S, 145.7115° E; and Microcosm B: NSW, 33.7035° S, 148.2612° E) with contrasting precipitation regimes –an important environmental factor which often lead 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 because of their contrasting precipitation regimes: 400 (site A) and 657 mm (site B). Clay %, total soil organic C, and pH (estimated as explained above) were 32 and 37%, 1.7 and 1.8% and 6.0 and 5.6 for soils for sites A and B, respectively.

Microcosm preparation

Soil samples from each site were sieved to < 2mm 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, Sydney. Gamma radiation was used as it is known to cause minimal change 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 phosphate buffered saline (PBS). The mixture was vortexed on 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 x 5 replicates x 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 in ref.[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. Thus, we 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, Carlsbad, CA, USA) as per the manufacturer's instructions. In order 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 on a CFX-96 thermocycler (Bio-Rad, USA) as described in ref.[48]. Standard curves were generated using ten-fold serial dilutions of plasmids containing the correct insert of each gene. The diversity of soil bacteria and fungi was measured via amplicon sequencing using 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. Bioinformatic and rarefaction analyses were done as explained above for the cross-biome study. Note that not all samples passed our rarefaction cut-off. We obtained information for 17/20 microcosms for soil A, and in 19/20 microcosms for soil B. We calculated the richness of bacteria and fungi in each soil replicate from rarefied zOTU (zero-radius OTUs) 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 functions but soil respiration and glucose mineralization were measured as explained 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 in a glass jar (12 cm depth, 75 cm diameter, Ball, USA), 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, 30 and 60 min after sealing. Soil gas flux for CO_2 was measured in an Agilent-7890a gas chromatograph (Agilent Technologies, Wilmington, DE, USA). 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 multi-threshold 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 between 0 and 1 (rawDiversity-min(rawDiversity)/(max(rawDiversity) - min(rawDiversity)), the ecosystem functions evaluated, and then averaged. In the case of the global survey, prior to this analysis, we averaged the soil variables observed in the five replicates (five composite samples/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 use 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 multi-threshold approach[14], aims to evaluate 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 PCA (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 eleven functions and the twelve 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 (OLS) linear regressions between soil multidiversity (standardized averaged of the diversity of twelve soil organisms) and single soil organisms with multifunctionality, multidimensional functioning (axes of a PCA analysis including eleven functions) and the number of functions > 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 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 (portion of variation in multifunctionality not explained by sampling date) with multidiversity.

Structural Equation Modelling

We used structural equation modeling (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 % of clay) attributes simultaneously (See a priori model in Supplementary Fig. 4; Supplementary Table 2). Mean annual temperature (MAT) and Aridity Index (AI = precipitation / evapotranspiration) were obtained from WorldClim derived data (http://www.worldclim.org) at 1 km resolution. Aridity was calculated as the inverse of the Aridity Index (-1 × AI). A useful characteristic of SEM for our purposes lies on its utility for partitioning the effects that a variable may have on another, and for estimating the strengths of these multiple effects. Unlike regression or ANOVA, SEM offers the ability to separate multiple pathways of influence and view them as parts of a system, and thus is useful for investigating the complex relationships among predictors commonly found in natural ecosystems[10]. All variables were included as independent observable variables. The diversity of twelve soil organisms was included as a composite variable in our SEM, 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 fashion, and that this relationship was well described by a second-order polynomial. In order to introduce polynomial relationships into our model, we calculated the square of aridity and introduced it into our model

using a composite variable approach described above. SEM models were conducted with the software AMOS 20 (IBM SPSS Inc, Chicago, IL, USA).

Correlation networks

To identify ecological clusters of strongly associated soil taxa including unique soil phylotypes, a correlation network, i.e., 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. Such filtering, is aimed to reduce potential spurious correlation from the rare taxa. We used the definition of dominant phylotype explained in ref.[20] to apply an additional constraint to ensure we identified dominant phylotypes. While many bacterial taxa are globally distributed[20], this is unlikely to be the case for eukaryotic organisms. Because of this, here we applied a >25% ubiquity threshold. 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 1782 dominant soil phylotypes strongly co-occurring with each other. These soil phylotypes were dominated by 1674 bacteria, 53 fungi, 77 protists, and 5 nematodes.

We used a correlation cut-off of Spearman $\rho > 0.65$, $P < 0.001$, which is largely used in the current literature, and comparable across studies[18], to generate statistically robust correlations and control the false positive rate as much as possible. This cut-off, which is largely used in the microbial literature[18], is expected to 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 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 (based on relative abundance) where 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 with 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[50] and brainGraph[51]. We then computed the richness of soil organisms within each ecological cluster, and that of highly connected soil taxa (classified as hubs; Fig. 2 in ref.[21]) across 81 globally-distributed locations.

We also estimated the richness of dominant taxa within ecological clusters, and 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 overlap, allowing us to match (>97% similarity) representative sequences for bacterial soil phylotypes found in both databases; and (2) based on global survey, bacterial taxa accounted for 94% of all taxa included in our correlation network (based on our global survey), and was the only group of organisms including 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 was 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 (qPCR 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.

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 U S A.* **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 U S A.* **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. *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. *Pro Natl Acad Sci USA* **111**, 308-313 (2014).
9. Bradford M.A. (2014). Discontinuity in the responses of ecosystem processes and multifunctionality to altered soil community composition. *Pro Natl Acad Sci U S A* **111**, 14478-14483.
10. Delgado-Baquerizo, M. et al., Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nat Comm* **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 U S A* **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. *The 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 Comm* **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 U S A.* **104,** 19891-19896 (2007).
23. Bascompte J. Stouffer D.B. The assembly and disassembly of ecological networks. *Philos Trans R Soc Lond B Biol Sci.* **364,** 1781–1787 (2009).
24. Gotelli N.J., Colwell R.K. Estimating species richness. Frontiers in measuring biodiversity. Oxford University Press, New York (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. *The 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, Marcel. Keystone taxa as drivers of microbiome structure and functioning. *Nat Rev Microb* **16**. 1. 10.1038/s41579-018-0024-1 (2018).
30. Bahram M, Hildebrand F, Forslund SK Anderson JL, Soudzilovskaia NA, Bodegom PM et al. Structure and function of the global topsoil microbiome. Nature 560: 233-237 (2018).
31. Kettler TA et al., Simplified Method for Soil Particle-Size Determination to Accompany Soil-Quality Analyses. *Soil Sci Soc Am J* **65,** 849 (2001).
32. Fierer N. et al., Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc Nat Acad Sci U S A* **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 Roy Soc B: Biol Sci* **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 Meth* **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 SI, Martiny JBH. Broadscale Ecological Patterns Are Robust to Use of Exact Sequence Variants versus Operational Taxonomic Units. mSphere. Doi. 10.1128/mSphere.00148-18 (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., (2013) High-throughput fluorometric measurement of potential soil extracellular enzyme activities. *J Vis Exper* **81,** e50961.
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 Micro* **69,** 3593-3599 (2013).
42. Hu H-W. et al., Diversity of herbaceous plants and bacterial communities regulates soil resistome across forest biomes. *Envir Micro* **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 Envir 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 Envir Contamin Toxic* **54,** 668–675 (1995).
49. Fierer, F. et al., Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. *Appl Envir Micro* **71,** 4117-4120 (2005).
50. Csárdi G. igraph, Network Analysis and Visualization v. 1.2.2. R package (2018).
51. Watson C.G. brainGraph, Graph Theory Analysis of Brain MRI Data v. 2.2.0. R package (2018).
52. Delgado-Baquerizo et al. Data from: Multiple elements of soil biodiversity drive ecosystem functions across biomes. Figshare digital repository DOI: 10.6084/m9.figshare.9976556 (2019).

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 702057. This research is supported by the Australian Research Council projects (DP170104634; DP190103714). We would like to thank Noah Fierer, Matt Gebert, Jessica Henley, Victoria Ochoa, Fernando T. Maestre and Beatriz Gozalo for their help with lab analyses. We also want to thank Osvaldo Sala, Christina Siebe, Courtney Currier, Matthew A. Bowker, Vicki Parry, Hans Lambers, Peter Vitousek, Victor M. Peña-Ramírez, Lynn Riedel, Julie Larson, Katy Waechter, Wendy Williams, Steve Williams, Benjamin Sulman, David Buckner and Brian Anacker for their help with soil sampling in Colorado, Hawaii, Iceland, New Mexico, Arizona, Mexico, and Australia, and to the City of Boulder Open Space and Mountain Parks for allowing us to conduct these samplings. We also thank Concha Cano-Díaz for her advice with regard to R analyses. Moreover, we would like to thank Samantha K. Travers for her help with mapping.

S.A. and F.D.A. are funded by FONDECYT 1170995, IAI-CRN 3005, PFB-23 (from CONICYT) and P05-002 (from Millennium Scientific Initiative). N.A.C. acknowledge support from Churchill College, University of Cambridge. M.W. acknowledges support from the Wilderness State Park, MI for access to sample soil and conduct ecosystem survey. BKS also acknowledge research award from the Humboldt Foundation. J-Z He acknowledge supported from the Australia Research Council (project DP170103628). A.G. acknowledges support from the Spanish Ministry (project CGL2017-88124-R). F.B. thanks the Spanish Ministry and FEDER funds for the CICYT project AGL2017-85755-R, the CSIC project 201740I008 and funds from "Fundación Séneca" from Murcia Province (19896/GERM/15). P.T. thanks Kyali for her help with lab analyses. S.R. was supported by the U.S. Geological Survey Ecosystems Mission Area. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. S.N. was funded by the Austrian Science Fund (grant Y801-B16).

Author contribution: 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 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 done by M.D-B., A.G., L.G-V., P.T., C.T., J-Z.H, H-W.H and F.B. Bioinformatic analyses were done by M.D-B and J-T.W. Statistical modeling and network analyses were done by M.D-B. The first draft of the paper was written by M.D-B., further drafts by M.D-B., P.B.R., D.J.E. and B.K.S., and all authors contributed to those subsequent drafts.

Competing interests: We declare no competing interest.

Data Availability: Soil biodiversity and functional data from the global field survey and the microcosm experiment are publicly available in Figshare[52].

Figure legends

Figure 1. Linear relationships between the biodiversity of selected groups of soil organisms (number of species, richness) and multidiversity (standardized between 0 and 1) with multifunctionality (n = 81). P-values (Pearson regressions) as follow: *P < 0.05; **P < 0.01.

Figure 2. Links between soil biodiversity and ecosystem multifunctionality in a global field survey. Panel A represents the fitted linear relationships between 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) with average multifunctionality (Pearson regressions; $P \leq 0.05$; n = 81). Panel B represents a fitted Structural Equation Model aiming to identify the direct relationship between the combined biodiversity of twelve groups of soil organisms and averaging ecosystem multifunctionality (EMF) (n = 81). We grouped the different categories of predictors (climate, soil properties, plants and spatial influence) in 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 from the biodiversity of twelve selected soil taxa. 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 Structural Equation Modelling) are $^{**}P < 0.01$ and $^{*}P \leq 0.05$. MAT (mean annual temperature). Information on BOX A-C and direct effects for other SEM arrows can be found in Supplementary Table 2. Information on our *a priori* model can be found in Supplementary Fig. 4 and Supplementary Table 2. Panel C includes 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).

Figure 3. Relationship between the biodiversity of selected groups of soil taxa (number of phylotypes) and of a composite metric of their joint diversity (multidiversity; standardized between 0 and 1) with multi-threshold functioning in a global field survey (n = 81). Fitted linear regressions between the diversity of single groups of soil organisms and the number of functions over multiple thresholds. Different colors represent different thresholds of functioning. P-values (Pearson regressions) as follow: *P < 0.05; **P < 0.01.

Figure 4. Linkages between soil biodiversity and ecosystem multifunctionality in a microcosm study. Panels show 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. Different colors represent different dilutions from our dilution-to-extinction approach (D0-D6; n = 5). P-values (Pearson regressions) as follow: **P < 0.01.

Figure 5. Linkages between the soil biodiversity within ecological networks and multifunctionality. Panels show the linear relationships between 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 with averaging multifunctionality (n = 81). Microcosms A and B were conducted in two different soils from Eastern Australia. Different

colors represent different dilutions from our dilution-to-extinction approach (D0-D6; n = 5). P-values (Pearson regressions) as follow: **P < 0.01.