



HAL
open science

Towards Unraveling Macroecological Patterns in Rhizosphere Microbiomes

Caroline Brunel, Robin Pouteau, Wayne Dawson, Michael Pester, Kelly
Ramirez, Mark van Kleunen

► **To cite this version:**

Caroline Brunel, Robin Pouteau, Wayne Dawson, Michael Pester, Kelly Ramirez, et al.. Towards Unraveling Macroecological Patterns in Rhizosphere Microbiomes. *Trends in Plant Science*, 2020, 25 (10), pp.1017-1029. 10.1016/j.tplants.2020.04.015 . hal-02985872

HAL Id: hal-02985872

<https://hal.inrae.fr/hal-02985872>

Submitted on 2 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Review

Towards Unraveling Macroecological Patterns in Rhizosphere Microbiomes

Caroline Brunel,^{1,2,*} Robin Pouteau,^{1,3} Wayne Dawson,⁴ Michael Pester,^{5,6} Kelly S. Ramirez,^{7,8} and Mark van Kleunen^{1,9}

It is generally accepted that plants locally influence the composition and activity of their rhizosphere microbiome, and that rhizosphere community assembly further involves a hierarchy of constraints with varying strengths across spatial and temporal scales. However, our knowledge of rhizosphere microbiomes is largely based on single-location and time-point studies. Consequently, it remains difficult to predict patterns at large landscape scales, and we lack a clear understanding of how the rhizosphere microbiome forms and is maintained by drivers beyond the influence of the plant. By synthesizing recent literature and collating data on rhizosphere microbiomes, we point out the opportunities and challenges offered by advances in molecular biology, bioinformatics, and data availability. Specifically, we highlight the use of exact sequence variants, coupled with existing and newly generated data to decipher the rules of rhizosphere community assembly across large spatial and taxonomic scales.

Unearthing the Macroecology of Rhizosphere Soil Microbes

Recent advances in sequencing technologies have expanded our ability to study plant-associated microbial communities. This has transformed our perception of the interactions between the plant and its **microbiome** (see [Glossary](#)) [1], which are now increasingly regarded jointly as **holobionts** [2–5]. The **rhizosphere** (i.e., the interface of plant roots and soils [6]) hosts diverse communities of microorganisms that are crucial to the plants they associate with. The rhizosphere microbiome can supply plants with nutrients [7], and protect plants against pathogens [8]. Furthermore, microbiomes can stimulate plant growth by producing phytohormones [9,10], and improve plant resistance and tolerance to abiotic stressors. Recent research suggests that rhizosphere microbiomes can alter plant phenology (e.g., flowering time) [11], modify morphological and size-related traits (e.g., shoot and root length and biomass, and number of secondary roots and leaves) [12], have a major role in plant community dynamics [13–15], and mediate plant responses to global change [16,17]. While much progress has been made, we are still far from understanding the mechanisms that control rhizosphere microbiome assembly and maintain community structure and composition.

It has long been hypothesized that the **macroecological** patterns (i.e., ecological patterns across large spatial scales) of the rhizosphere microbiome will relate to the macroecological patterns of plants [18]. However, few studies have found clear relationships between plant and rhizosphere diversity [19,20]. This is likely because there is a disconnect between the plant and the microbial scales ([Box 1](#)). For other (non-rhizosphere) microbiomes, large-scale sampling campaigns are leading the way by generating standardized raw data and metadata (e.g., the Earth Microbiome Project [21] or the Human Microbiome Project [22]). For macro-organisms, macroecological patterns are frequently identified by collating existing and newly generated data from individual studies [e.g., based on the Global Biodiversity Information Facility (GBIF)ⁱ; Map of Life]. However, to our knowledge, the same has not been done for rhizosphere

Highlights

The processes shaping rhizosphere microbial communities are currently unclear because of both a lack of knowledge about biogeographical patterns and the disconnection between plant and microbial scales.

Sequence databases have now collected a sufficient amount of data covering a range of biomes and plant taxa to allow synthesis across studies. Recently, new bioinformatic methods have been developed that allow us to overcome former spatial and taxonomic limitations.

Understanding the processes that shape rhizosphere microbial communities will provide important insights into plant ecology and evolutionary biology, and can enable us to manage microbial and plant ecosystem services.

¹Zhejiang Provincial Key Laboratory of Plant Evolutionary Ecology and Conservation, Taizhou University, 318000 Taizhou, China

²IRD, IPME, 911 Avenue Agropolis, BP 64501, 34394 Montpellier, France

³AMAP, IRD, CNRS, CIRAD, INRA, Université de Montpellier, 34398 Montpellier Cedex 05, France

⁴Department of Biosciences, Durham University, Durham, DH1 3LE, UK

⁵Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures, Inhoffenstr. 7B, D-38124 Braunschweig, Germany

⁶Technical University of Braunschweig, Institute for Microbiology, Spielmannstrasse 7, 38106 Braunschweig, Germany

⁷Netherlands Institute of Ecology, Droevendaalsesteeg 10, 6708, PB, Wageningen, The Netherlands

⁸University of Texas at El Paso, Department of Biological Sciences, 500 W University Ave, El Paso, TX 79968, USA



microbiomes. Synthesizing rhizosphere data represents a fundamental way to connect below-ground diversity with aboveground diversity. Yet, to successfully synthesize rhizosphere microbiome data and better predict macroecological patterns in the rhizosphere microbiome, the following challenges must be addressed: (i) obtain sufficient geographical and plant-host taxonomic coverage; and (ii) quantify between-study heterogeneity in sequence data. In this review, we highlight how new bioinformatic tools coupled with meta-analyses offer an opportunity to better understand the rules of rhizosphere microbiome assembly.

⁹Department of Biology, University of Konstanz, 78464 Konstanz, Germany

*Correspondence:
caroline.brunel@ird.fr (C. Brunel).

Rhizosphere Microbiomes Research Remains Incomplete

To reveal what currently restricts our understanding of rhizosphere microbiome assembly and identify steps to move the field forward, we summarize here the most dominant research themes on rhizosphere microbiomes in 2018. A literature search using the topic queries 'rhizosphere microbiome' and 'rhizosphere microbial communities' in Web of Science revealed a total 662 articles published in 2018, of which 382 were suitable for analysis (see the supplemental information online). From these 382 studies, we extracted the main topics, and information on their geographical, temporal, and taxonomic scales, the studied microbial taxa, the origins of the samples, and the sequencing analytic approaches used (Figure 1). Overall, scientists considering rhizosphere microbiomes were predominantly interested in the following four topics: (i) the rhizosphere environment (exudates, autotrophic effects of microbial communities, or exploratory

Box 1. The Macroecology of Rhizosphere Microbiomes

At larger spatial scales, rhizosphere microbiome assembly results partly from the metabolic synchronization of microbial substrate-utilization traits and plant-root exudation [82]. From this postulate, macroecological patterns in rhizosphere microbiomes are expected to result from the biogeography of both free-living soil microbes (determined by geography, dispersal, environment, and biotic interactions [83]) and the host plants (e.g., species, developmental stage, and symbiosis establishment [84], and functional traits). Currently, a large body of evidence challenges the 'everything is everywhere, but the environment selects' hypothesis [85], claiming that microorganisms are not cosmopolitan [86], and implying that, in addition to the environmental drivers, dispersal limitation also controls the assembly of root-associated microorganisms [87]. Therefore, quantifying the dispersal ability of microbes and vectors of dispersal is the first step in understanding the macroecology of rhizosphere microbiomes.

However, from the perspective of a microbe, a whole plant may already be approaching a macroecological spatial scale. Consider Figure 1. Assembly of a plant community at a location might be determined by dispersal ability of species in the regional pool (light blue); abiotic (orange), and biotic filters (green) should then act in concert to shape plant and rhizosphere microbial communities [88]. However, biotic filters can operate at the spatial scale of the plant or root scale to shape plant and microbial communities, respectively. While filters shaping plant communities at larger spatial scales influence the pool of species from which microbial communities are constructed, abiotic and biotic filters constrain the species pool further at the root-system scale (Figure 1, root system box). If we zoom in to the scale of individual roots, we would expect rhizosphere microbial community composition to be shaped by root characteristics such as root thickness, root hairs, root-associated fungi, priority effects resulting from other microbes, and exudate compounds (Figure 1, 'Root Characteristics'; in green).

To fully understand patterns of microbiome diversity, composition and functions, we need to measure and study rhizosphere microbiomes at the scale of individual roots and within root systems. For example, we do not know how much variation there is in microbial communities and their functioning among different roots within a single root system. We need to learn more about how this variation relates to the age of roots, root depth, and neighboring plants (Figure 1). We do not know how much variation in rhizosphere microbiomes there is at this scale between different plants, different plant populations, and different regions. Moreover, we also do not fully understand how or why rhizosphere microbiomes vary in composition temporally, and what the effects of this variation might be for plant–soil interactions and their consequences over time. Finally, we lack information on whether microbial diversity in the rhizosphere is structured any differently from bulk soils at different spatial and temporal scales. While evidence runs counter to the 'everything is everywhere, and the environment selects' idea, we may not be testing this idea at scales most relevant to microbes. For example, the 'regional' species pool from which rhizosphere communities are assembled might largely comprise all the microbes occurring in a single root system and surrounding bulk soil, and in the root systems of neighboring plants. Associations between plant and soil microbial communities might also be organized at the scale of 'root communities', determined by the identity and distribution of the roots of species in a volume of soil [19]. We argue that, to fully understand the macroecology of rhizosphere microbiomes, we need to refocus our lens of enquiry to consider the factors governing microbiome assembly at finer spatial scales and how they relate to the larger scale patterns of diversity and composition more typical of macroecological studies.

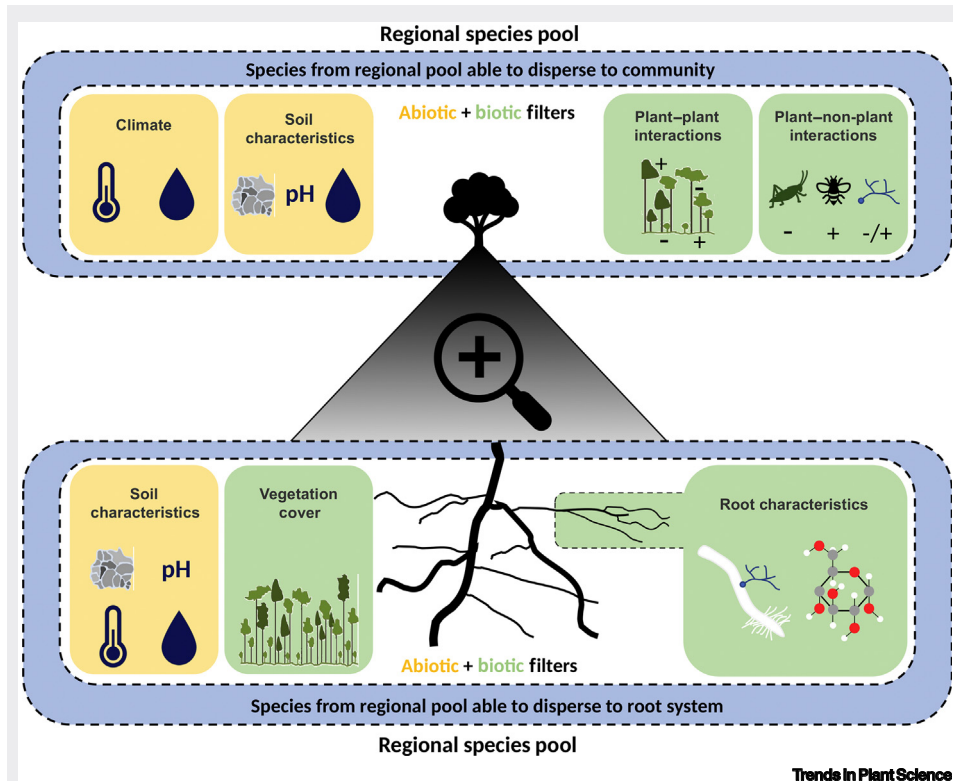


Figure 1. Hypothetical Filters Affecting Assembly of Plant Communities and Rhizosphere Microbiomes in Plant Root Systems. Abiotic filters shaping communities include temperature, precipitation (climate), soil structure, temperature, moisture, and pH (soil characteristics). Biotic filters shaping plant communities include plant–plant interactions (competition [–] and facilitation [+]), and plant–non-plant interactions (herbivory, pollination, and fungal pathogens/mutualists). Biotic filters shaping rhizosphere microbial communities include vegetation cover and root characteristics of focal and neighboring plants (root thickness, root hairs, root-associated fungi, and exudate compounds).

studies); (ii) the relation to the host plant; (iii) external biotic factors (i.e., other than the holobiont, e.g., pathogens, microbiomes of other plants, bulk-soil microbiome or mesofauna); and (iv) the effects of abiotic factors (anthropogenic or geoclimatic parameters).

Community assembly is assumed to work as a hierarchy of nested constraints (i.e., dispersal constraints and abiotic together with biotic filters) of potentially varying strengths acting at different spatial scales [23,24]. Despite covering a variety of environmental filters, our synthesis found that current research exhibits a pronounced specialization and compartmentalization of studied topics, ~43% of the studies focused on only one of the four topics (i.e., the rhizosphere environment, the relation to the host plant, external biotic factors or abiotic factors; Figure 1). Furthermore, there was a limitation in ‘scales’ studied. Approximately 78% of the reviewed literature focused on a single plant species, 69% had sampled only at a single time point, and 60% did not consider spatial scales (Figure 1). The detection of **low-abundance taxa** is extremely limited when focusing on very small scales of study [25], which represents a major issue because these taxa have a leading role in microbial community structure and functions [25–28] and in the way that microbial communities interact with plants [29]. At the same time, few studies examined the assembly of rhizosphere microbiomes at the plant scale. For example, ~16% of the studies (i.e., 61 out of 382) described variation across plant traits, developmental stages, or generations (Figure 1). Approximately 23% (i.e., 87 out of 382) focused on the effect of plant identity

Glossary

Amplicon sequence variant (ASV): a higher-resolution analog of the traditional OTU. ASVs are inferred regardless of other sequences or reference databases by a process that sufficiently controls errors so that sequences can be resolved exactly, down to the level of single-nucleotide differences. Consequently, ASVs combine advantages of simple merging between independently processed data sets and accurate measurement of diversity with applicability to communities lacking deep coverage in reference databases [73].

Holobiont: an integrated unit including a host and its associated microorganisms, which must be considered as a whole to understand evolutionary and ecological processes [2,4].

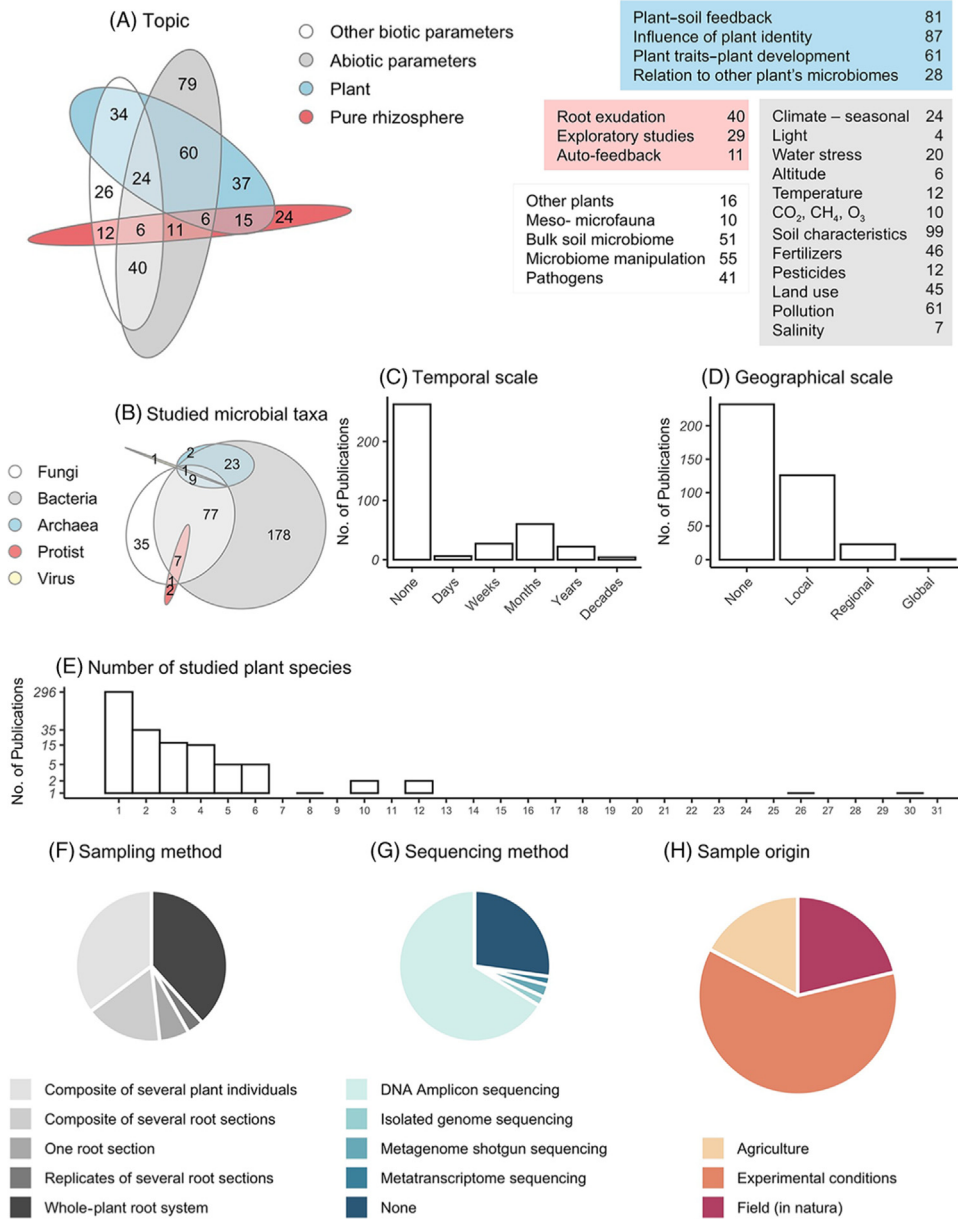
Low-abundance microbial taxa: unlike the few microbial species that are very abundant and account for most of the biomass and carbon cycling, most microbial species have an individual relative abundance of <0.1% and are defined as the ‘rare biosphere’ [75]. Their contribution to community dynamics is still unclear. Owing to their low relative abundance, traditional molecular techniques [clone libraries and fingerprinting methods, such as denaturing gradient gel electrophoresis (DGGE), and terminal restriction fragment length polymorphism (T-RFLP)] could not retrieve them [76,77].

Macroecology: the study of relationships between organisms and their environment at large spatial scales to characterize and explain patterns of abundance, distribution, and diversity.

Microbiome: the set of microorganisms present at a given site, with consideration often given to their genomes and the host or surrounding habitat [78].

Operational taxonomic unit (OTU): sequence clusters that constitute the standard unit of marker-gene analysis. There are two main methods of defining molecular OTUs. Closed-reference methods involve searching for a sequence in a reference database that is sufficiently similar to reads and subsequent recruitment of reads into a corresponding OTU. *De novo* methods involve grouping of reads grouped into OTUs as a function of their pairwise sequence similarities [73].

Orthologs: genes found in different species that evolved from a common ancestral gene, and usually have retained the same function. Ortholog



identification is a critical process for reliable prediction of gene function in metagenomes.

Rhizosphere: this term, coined by Lorenz Hiltner in 1904 [79], refers to the narrow zone of soil around plant roots, which is under the chemical, physical, and biological influence of root exudates. The size of the rhizosphere (affected by soil type, plant species, and phenological growth stage) ranges from <1 mm to several centimeters around the root system [80]. The rhizosphere microbiome differs from but interacts with the endosphere microbiome (i.e., microorganisms living inside the roots), the rhizoplane microbiome (i.e., the biofilm surrounding the roots), and the bulk soil microbiome (i.e., microorganisms that do not directly interact with the roots, but may indirectly benefit from metabolites derived from the rhizosphere soil microbiota).

Figure 1. How Did Scientists Consider the Rhizosphere Microbiomes in 2018? Overview of the 382 articles, published in 2018, that responded to a 'rhizosphere microbiome' and 'rhizosphere microbial communities' topic query in Web of Science (query done in March 2020, see details in the supplemental information online). (A) Venn diagram representing the topics and subtopics raised in each article. The number of studies considering each of them are indicated in the boxes with matching colors (each publication can account for several topics and subtopics). (B) Venn diagram representing the repartitioning of studied microbial taxa. (C) Temporal scales ('None' stands for a single time point). (D) Geographical scales ('None' refers to experimental conditions). (E) Numbers of studied plant species. (F) Rhizosphere soil-sampling method. (G) Sequencing approaches used to characterize rhizosphere communities ('None' refers to methods that do not rely on sequencing, e.g., the description of catabolic profiles or phospholipid-derived fatty acids measurements). (H) Sample origins.

(i.e., species, varieties, or genotypes), and only 4% (16 out of 382) included the plant community scale (i.e., considered the influence of surrounding plants; [Figure 1](#)). Collectively, this suggests that, for a better understanding of the rhizosphere microbiome, research could benefit from more studies that consider a geographical gradient or temporal approach. Specifically, experimental designs and analytic tools could be deployed to resolve variation not only between plant individuals or composite samples, as is commonly done, but also within root systems, ([Figure 1](#)).

In addition to scale, microbial taxonomic and functional diversity must be more fully assessed. Metabarcoding has proven to be a valuable tool and has been widely used for characterizing the microbial diversity in different environments [30]. Specifically, we found that sequencing approaches were used in ~73% of the cases, and amplicon sequencing of taxonomic or functional marker genes was by far the most popular approach (used in ~66% of the cases; [Figure 1](#)). Furthermore, experiments were designed to improve knowledge of gene and metabolite expression profiles involved in plant–microbe interactions by using pairwise interactions under controlled conditions: most studies used semiartificial experimental set-ups: ~61% of the rhizosphere microbiome studies were conducted under controlled conditions, whereas 39% were described as being conducted *in situ* ([Figure 1](#)). While these experiments are valuable, they are admittedly difficult to scale up to natural ecosystems [31]. Furthermore, research on rhizosphere microbiomes was often limited to a phylum-level perspective and bacteria had received special attention (~88% of the studies analyzed bacterial communities), while complex interactions among fungi, bacteria, archaea and the plant host remained poorly studied ([Figure 1](#)). The heterogeneity in sampling methods ([Figure 1](#)) also highlights the vague definition of 'rhizosphere' that can complicate comparisons across studies. Additionally, a standardized protocol for rhizosphere soil collection could be developed for each life form (e.g., tree, shrub, or herbaceous) to reduce disparity of results. Such standardization steps are especially important if different processes dominate at different nested scales within rhizosphere microbiomes. This overview underlines the need for a more holistic approach to studying rhizosphere microbiomes considering large spatial, temporal, and taxonomic dimensions.

Can We Transcend Taxonomic and Geographic Limitations?

When a large-scale sampling campaign is not feasible, the compilation of existing data with sufficient geographical and plant-host taxonomic coverage is a viable alternative. Data sharing among scientists has been encouraged by public data repositories as well as the demands from journals for making sequence data publicly available as a condition for publication. For example, the International Nucleotide Sequence Database Collaboration (which includes the DNA Data Bank of Japan, the European Nucleotide Archive, and the National Center for Biotechnology Information; hereafter NCBI) solicits the banking of complete sequence data and metadata [32,33]. These databases now provide access to thousands of sequence data sets from rhizosphere samples. The metadata associated with the rhizosphere metagenomes (in NCBI, referring mainly, but not only, to amplicon-gene sequencing) are essential for finding, retrieving, and reusing data stored in online repositories. The metadata include environmental data (e.g., details about the host plant or the geographical location of sampling) and methodological data (e.g., the applied protocol). Collectively, these data provide great opportunities to address questions of community assembly and co-existence rules of rhizosphere microbiomes, and to connect belowground diversity to aboveground diversity. To get an impression of the amount and quality of data available, we used the NCBI BioSample database [34], which stores all sequence metadata submitted to NCBI. We obtained information on the publication date, plant host, and geographical coordinates of the collection locations of the root and rhizosphere microbiomes [i.e., deriving from the taxonomic id 'txid939928' entitled 'rhizosphere metagenome' ([Figure 2](#)) 38 146 metadata records in August

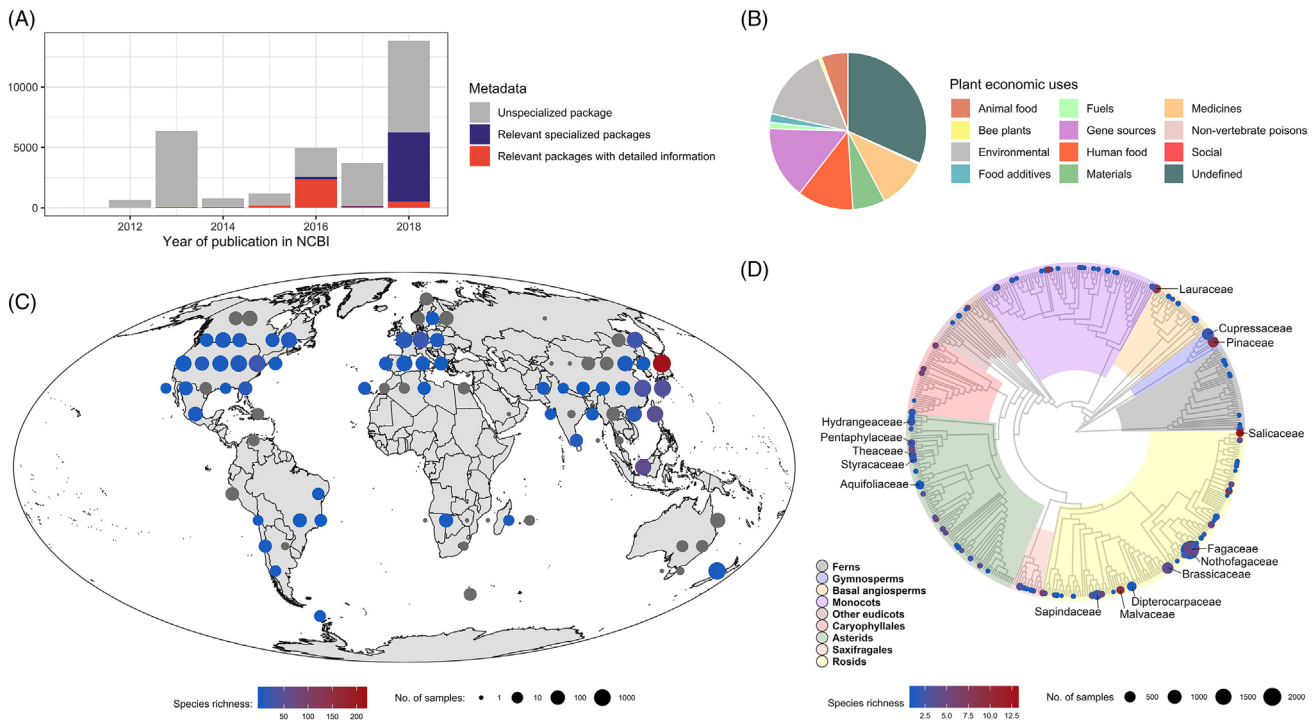
**Trends in Plant Science**

Figure 2. The Rhizosphere Microbiome Collection. Out of the 38 146 rhizosphere microbiomes collected worldwide (data stored in the NCBI Biosample repository with the taxonomy ID: 'txid939928' recorded before August 2019), only metadata of 10 670 (~28%) were recorded in the dedicated MIMS/MIMARKS plant-associated and host-associated packages, 20 460 (~54%) were geolocalized, 7845 (~20%) were assigned to host plant species (264 different plant species), and only 4465 (~12%) fulfilled all three criteria. (A) Temporal trend of data publication ($N = 38\ 146$) in the generic and nonspecialized packages (in gray), in specialized package (i.e., packages allowing a relevant description of associated metadata, in blue) and in specialized packages with information about the host species and the localization (in red). (B) Plant economic use (according to the GRIN database [81]) of the 264 different plant species. (C) World map showing collection locations of the 20 460 geolocalized rhizosphere microbiome samples. (D) Phylogenetic representation of all vascular plant families (including angiosperms, gymnosperms, and ferns) and associated rhizosphere microbiomes ($N = 7845$, all family names are given in Figure S1 in the supplemental information online). In (C) and (D), different colors represent the sample species richness, point size represents the number of samples, and gray points stand for data without information about the host plant.

2019]. NCBI BioSample promotes the use of the MIMS and MIMARKS host- or plant-associated packages to implement standardization and description of sequence-based data and to provide some control over metadata submission [35]. Nevertheless, we found that rhizosphere sequence metadata were not always registered in the corresponding specialized packages, because ~72% were recorded in nondedicated packages, which do not provide a relevant description of the samples. Most submitters used the generic package, which has no specific requirements.

Thousands of rhizosphere microbiome studies are now available spanning a range of terrestrial ecosystems (Figure 2). We found a clear exponential increase in the number of 'rhizosphere metagenome'-associated samples published each year, and data came from all continents, which provides opportunities for global analyses (Figure 2). Moreover, according to the metadata, these samples were associated with 264 different host-plant species belonging to 134 families (Figure 2). The majority of these plant species (~72%) have economic uses, with plants used for environmental (mainly ornamental) purposes (~14%), as genetic resources (~13%), as human food (~10%) and medicines (~9%) being well represented (Figure 2). Taken together, these numbers suggest that the recent exponential accumulation of sequence data and their associated metadata provide useful information for meta-analyses. However, out of the 38 146 rhizosphere samples recorded, ~46% were not geographically localized (and some were

incorrectly localized, e.g., mixed up latitude and longitude or used indeterminate coordinate systems) and 79% were not associated with a standardized plant species name (i.e., according to the Plant List[†]). Thus, there is still much scope for improvements in data reporting.

The metadata describing sampling location, associated plants, and other important details, such as nucleotide extraction and sequencing protocols, were also often not included with the published supplemental data sets. This problem is also common to other fields of research. A recent meta-analysis of 11.4 million metadata of samples used in biomedical experiments revealed multiple anomalies and highlighted the need for a more robust approach to reporting metadata [36]. Inclusion of metadata is crucial for finding suitable data sets for meta-analyses, will help to ensure interoperability of the metadata, and ultimately will allow us to begin to understand how the host plant (the environment of the rhizosphere) shapes macroecological patterns in rhizosphere microbiomes. Thus, it is vital that public repositories implement stricter control of submitted data for standardization, and that authors add sufficient details about sample origin and nucleotide sequences, and carefully use formal language when submitting metadata. Transcending geographical and taxonomic boundaries already appears feasible by using shared data, but increased awareness of the importance of metadata would multiply the strength of this approach.

Improvements and Limitations of Sequence-Based Data

Describing microbial functions is required to understand what connects a microbe to its habitat, particularly in the rhizosphere, where plant exudates and microbial substrate preferences constitute an important driver of microbial community assembly [37]. Furthermore, linking information from plant and microbial data sources can help to answer macroecological questions (Box 2).

Box 2. Plant Reference Databases

For plant ecology, the collation of local data sets from independent studies into global databases has revealed global patterns and provided insights into the underlying ecological processes [89]. Plants have a critical role in biogeochemical cycling [90], the supply of food and goods [91], and ecosystem resilience [92]. Similarly, our understanding of rhizosphere microbiomes and their contribution to ecosystem services [93] would be greatly improved from coordinated efforts that are global in scope. It is known that the rhizosphere microbiome has a fundamental role in numerous processes related to plant and ecosystem functioning. For example, rhizosphere processes are quantitatively important drivers of soil organic matter decomposition and nutrient release at the ecosystem scale, with consequences for global C cycling and vegetation feedbacks to global change [94–97].

We have largely described rhizosphere data from the perspective of the microbe, yet, to date, there are no global efforts to link rhizosphere data to global plant or vegetation data. If we consider plants and their associated microbiomes as holobionts, we need to combine data on plant ecology and physiology with rhizosphere microbiome data to understand rhizosphere microbiome composition and function. This would be valuable for studies involving a range of plant taxa, but such studies are still scarce. We need to link plant databases to rhizosphere microbiome studies, allowing us to relate plant identity to microbial community composition to address specific macroecological questions (see Outstanding Questions). Databases describing plant species functional traits [98], phylogeny [99], and native and exotic ranges [100] are available. Data processing has been facilitated by the development of bioinformatic tools allowing, for example, plant-name assignment [101], collection of plant occurrences through the GBIF database, and placement of species in phylogenetic trees [102]. Such tools and databases have increased our understanding of many macroecological patterns and ecosystem-level processes involving plants. The amount of available data describing plant root trait is increasing [103], and plant data sets could be used for studying plant–microbe interactions. Recently, forest-plot data and known associations of specific tree species with ectomycorrhizal and arbuscular mycorrhizal fungi were used to make the first global maps of the distribution of these mycorrhizal groups [13]. The patterns revealed demonstrate the critical role of root symbionts in shaping the distribution of tree species. Thus, we call for deeper levels of collaboration between plant and microbial ecologists, so that mutually informative linkages between plant and microbial data and macroecological studies can be forged. To begin with, microbiome-focused studies should better record the taxonomy and locations of the host plant(s) and describe plant traits while providing details on the sequencing methods. Gathering all these data in a dedicated database, or developing existing database packages for associated metadata will be crucial for improving knowledge on the macroecology of rhizosphere microbiomes. At the same time, it will also promote scientists' awareness of the importance of data sharing. Ideally, in the future, plant ecologists may be able to extract rhizosphere microbiome DNA from study species and locations for sequencing as a standard practice, as sequencing costs continue to fall.

There are currently three main methods that use microbial sequence information: taxonomic classification with rRNA-gene and internal transcribed spacer (ITS) sequences, microbial meta-omics, and complete genome sequencing. No one method is without bias, yet all provide critical information for understanding the ecology of rhizosphere microbiomes.

As for all microbiome sequence data sets, taxonomic classification of rhizosphere microbiomes relies on the assignment of sequences to taxonomic and functional units via reference-sequence databases (recently presented and compared in [38]). The amount and reliability of data in these databases is improving due to contributions from researchers worldwide and the colossal efforts of global initiatives such as the Earth Microbiome Projectⁱⁱⁱ and to the ongoing efforts of database managers to carefully curate and align high quality sequences. Thus, these reference databases cover more and more microbial taxa and offer great opportunities for global scale analyses on, for example, microbial diversity. Nevertheless, despite the improved methods and data availability, microbial diversity may be underestimated when conducting 16S- and 18S-rRNA gene or ITS analyses, if the assignment relies on incomplete reference gene databases [39]. In 2011, ~100 000 fungal species were described, but the extent of global fungal diversity was estimated at 800 000–5 100 000 species [40]. Thus, only 2–13% of all fungal taxa have been described so far. In 2016, Locey and Lennon [41] estimated that the Earth may host 10^{12} microbial taxa (encompassing Bacteria, Archaea, and microscopic fungi). By contrast, Overmann and colleagues determined that only ~14 000 species of Bacteria and Archaea, probably representing 0.001–0.1% of the estimated global species number, have so far been cultivated and formally described [42]. The entirety of the Earth Microbiome Project has cataloged fewer than 10^7 taxa, of which only one third have been detected at least twice. The latter could indicate that many microbial taxa are indeed low-abundance taxa [21]. Efforts to understand this so-called rare biosphere are now ongoing [43].

Coverage of the bacterial and archaeal 16S-rRNA-gene sequence space will continue to improve as metagenome-assembled genomes reveal previously unknown genes that had not been detected by universal primers [44]. This trend will also generally hold true for the 18S-rRNA-gene of eukaryotic microorganisms. For now, taxonomic delineation of microorganisms remains difficult because minor differences in 16S- or 18S-rRNA gene sequences may not necessarily translate into different traits of microorganisms. Furthermore, observed differences in taxonomic marker-gene sequences might also originate from sequencing errors. New denoising algorithms that correct sequencing errors and determine real biological sequences (i.e., exact **amplicon sequence variants, ASVs**) have recently been developed, but the few hundred bases generated by Illumina sequencing often provide insufficient taxonomic resolution. For a reliable taxonomic assignment across all taxonomic levels, 1300 base pairs are currently recommended [44]. With emerging single-molecule sequencing technologies (e.g., Oxford Nanopores and PacBio), taxonomic assignments will improve in the near future. However, even if most of the microbial taxa have been discovered by now based on their 16S-rRNA gene [44], the fact that most of them are not cultured [45] means that we still have limited information about their physiology and ecology [42,46].

Beyond taxonomic classifications, the use of functional groups of microorganisms should be prioritized to help improve the field of microbial meta-omics [47]. However, this requires consistent ways of sorting phylotypes into ecologically meaningful categories. For example, a recently published global atlas of the dominant bacteria found in soil sorted them into ecological groups, albeit at a coarse resolution [48]. Moreover, databases that classify microbial taxa according to their lifestyles or guilds are growing. For example, the FunGuild database [49] classifies fungal taxa as having a symbiotroph, pathotroph, or saprotroph lifestyle, while the FunGene database

provides a repository for the most common functional marker genes used in microbial ecology [50]. Additionally, the *nifH* (encoding the dinitrogenase reductase subunit of nitrogenase) database facilitates analyses of the evolution and ecology of nitrogen-fixing organisms [51], and the PhytoPath database records data on plant pathogen species [52]. This offers opportunities to assess macroecological questions on functional diversity of rhizosphere microbiomes.

At the same time, sequencing of complete microbial genomes is being successfully and more regularly used to identify the metabolic potential of microbes through gene identification [53]. This provides an alternative way to reveal functional groups of microbes, and can help to reduce the overall complexity of rhizobiomes, because orthologous genes can be assigned and functions of uncultivated microbes might be inferred. Some initiatives already give interesting perspectives in this regard. For example, OrthoMCL [54], KEGG [55], and related packages [56,57] allow **ortholog** prediction among multiple species and raise the possibility of assigning a microbial lifestyle through a genome signature [58,59]. Besides being limited by incomplete genomes and/or the quality of gene annotation, the predictions can be biased because the presence of genes in a metagenome does not necessarily mean that they are in fact expressed. To reduce such biases, assigning functions should ideally be complemented by following *in situ* expression and/or translation of assigned genes using metatranscriptomic [60] or metaproteomic [61] data, and the predicted function of a gene should be tested, such as with enzymatic assays. Despite the current limitations, new approaches related to deeper taxonomic assignment and functional description of microorganisms are promising and open new avenues for a better understanding of the macroecological patterns in rhizosphere microbiomes.

Methods to Synthesize across Studies

Heterogeneity of data describing microbial communities due to differences in methodology represents a significant obstacle for pooling data across studies. Before the recent advances in molecular biology, meta-analytical efforts focused on measuring the diversity or abundance of major microbial taxa derived from a variety of analytical methods. For example, Hendershot *et al.* [62] performed a global meta-analysis on belowground diversity (including taxonomic and phylogenetic indices) and abundance (including microbial biomass, phospholipid-derived fatty acids abundance, and fungal colonization rate) of 325 soil communities across 20 studies along temperature and soil-pH gradients. They addressed inconsistencies among studies by using meta-regression to explore relationships between the microbiome and abiotic factors both within and among studies, but they did not find any such relationships. Possibly, no patterns were detectable due to a mismatch between the scale of community sampling and the scale at which the data on the environmental variables are available.

Due to the emergence and rapid growth of high-throughput sequencing studies, it has now become possible to describe the taxonomic (approximated by **operational taxonomic units, OTUs**) and functional group composition of microbial communities with near-complete coverage. Therefore, the merging of numerous meta-barcoding studies now represents a good opportunity to detect macroecological patterns and infer their drivers. However, comparing microbiome-sequence data from independent studies remains challenging due to the lack of standard protocols. Potential biases include the variety of methods used to collect samples, to extract DNA, the use of different primers, PCR bias, the different sequencing platforms and sequencing depths, the clustering methods, and the taxonomic classification pipelines [63–65]. Each step of the workflow, from sampling to data set standardization, can produce variations that might blur biological patterns.

Despite these challenges, it is possible to compare microbiomes across multiple data sets, if one uses a reference population and standardized bioinformatic protocols. In a seminal study, Lozupone *et al.* [66] successfully compared human microbiomes using raw data derived from 12 independent data sets. The authors used sequences from studies that targeted different non-overlapping regions of the 16S-rRNA gene, and then related them to near full-length sequences from a reference database. This protocol allowed for comparison of sequences generated from different regions of the 16S-rRNA gene, but was in the meantime limited by the partial detection of taxa associated with primer affinity [65]. They showed that variation in the composition of microbiomes from different studies was consistently larger than the variation caused by differences in methods among the studies, and they demonstrated that cross-study comparisons of human microbiomes are possible, at least when the variable of interest has a large effect size. Subsequent studies have then applied comparable methodologies with the analysis of mammal gut microbiomes [67,68], and initiated the establishment of international research consortia to collect data and set up databases [69]. Such studies and initiatives are still missing for rhizosphere microbiomes.

Recently, Ramirez *et al.* [25] combined data from multiple independent studies of soils from around the globe, which they then analyzed with a machine-learning technique to assess macroecological patterns in soil bacterial communities. They used the metadata of the studies to identify technical biases and estimated with *in silico* analyses the biases of primer pairs used to produce the data. Then, they disentangled technical and environmental effects from taxonomic effects using random-forest techniques. This way, they were able to jointly analyze microbial community data from 30 individual studies comprising 1998 soil microbiome samples from 21 countries. They found that a subset of taxa with low relative abundances that occurred in most soil samples were more important for structuring soil communities compared with abundant taxa. These low-abundance taxa were also better predictors of community structure compared with environmental factors, which were often confounded across studies. Similar patterns on a smaller scale had previously been observed but required more controlled experimental conditions [29].

Recently, bioinformaticians have developed ASV-based methods that provide the finest-scale taxonomic partitioning of microbial phylotypes possible and allow simpler merging of sequence data across studies [70–73]. Generated phylotypes can be rapidly and efficiently compared across independently processed datasets, as long as the underlying sequence data has been derived from the same genetic locus (amplified with the same primers). The first meta-analysis using this methodology was recently published by Rocca *et al.*, combining 606 microbiomes that had been exposed to various environmental conditions sampled within different habitats [74]. The authors selected nine independent studies on 16S-rRNA-gene amplicon sequencing with publicly available raw 16S-rRNA-gene sequencing data from the Illumina MiSeq V4 hyper-variable region (reprocessed into ASVs). Their results contrast with the earlier mentioned results of Hendershot [62], which could not identify global patterns related to temperature using other analytic methods. Thus, the study by Rocca *et al.* [74] proved the efficiency of combining sequence data using ASVs and generated a more detailed understanding of how specific environmental parameters influence microbiome responses. Applied to the rhizosphere, the use of ASVs coupled with meta-analyses across large-scale studies would allow us to infer assembly mechanisms of plant-associated microbial communities.

Concluding Remarks

Recent advances in molecular biology are helping to unravel macroecological patterns in rhizosphere microbiomes. However, current work remains highly specialized and lacks the coverage of large temporal, geographical, and taxonomic scales. Deciphering the rules of rhizosphere

Outstanding Questions

Deciphering the macroecology of rhizosphere microbiomes could help us to answer the following fundamental and applied questions:

What are the evolutionary and environmental drivers of rhizosphere microbial community composition?

How do different functional groups of microorganisms differ in their biogeographical distributions?

What are the mechanisms within and across plant hosts that mediate microbe–microbe and host–microbe interactions?

Which are the major microbial taxa in control of plant performance (e.g., pathogens, plant disease suppressors, or phytohormone producers)?

What is the role of rhizosphere microbiomes in large-scale patterns of plant invasion, and how can the rhizosphere microbiome be effectively included in risk assessments of invasive non-native species?

How will rhizosphere microbiomes in different parts of the world respond to climate change (e.g., CO₂ increase and warming), and what will be the consequences for individual plant species, plant communities, and biogeochemical processes (e.g., greenhouse gas emissions or modification of carbon and nitrogen cycles)?

community assembly requires not only scaling up to detect macroecological patterns in the traditional sense, but also scaling down to understand spatial structure of rhizosphere microbiomes within root systems. In this review, we considered that the emerging new bioinformatic tools, specifically the use of exact sequence variants, coupled with existing and newly generated data, could be used to fill the knowledge gaps. This information will enable researchers to better explain and predict community and ecosystem responses to environmental change (see [Outstanding Questions](#)). To pursue this path, we encourage microbial and plant ecologists to collaborate in: (i) building links between databases; (ii) designing sampling campaigns guided by approaches used in macroecology and community ecology; and (iii) designing a standard protocol for recording of microbiome metadata and sampling of rhizosphere soil and DNA. These efforts will go a long way to unifying microbial ecology with macroecology and to advance our understanding of rhizosphere microbiomes.

Acknowledgments

This review was initiated by C.B. and greatly benefited from the work of all co-authors. C.B. acknowledges financial support from the research project of Taizhou University No. 2018YQ001 (China). W.D. acknowledges funding from the Royal Society Research Grant No. RG160281 (United Kingdom).

Resources

ⁱwww.gbif.org

ⁱⁱwww.theplantlist.org

ⁱⁱⁱwww.earthmicrobiome.org

Supplemental information

Supplemental information associated with this article can be found online at <https://doi.org/10.1016/j.tplants.2020.04.015>.

References

- McFall-Ngai, M. *et al.* (2013) Animals in a bacterial world, a new imperative for the life sciences. *PNAS* 110, 3229–3236
- Moran, N.A. and Sloan, D.B. (2015) The hologenome concept: helpful or hollow? *PLoS Biol.* 13, e1002311
- Zilber-Rosenberg, I. and Rosenberg, E. (2008) Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiol. Rev.* 32, 723–735
- Bordenstein, S.R. and Theis, K.R. (2015) Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLoS Biol.* 13, e1002226
- Douglas, A.E. and Werren, J.H. (2016) Holes in the hologenome: why host-microbe symbioses are not holobionts. *mBio* 7, e02099-15
- Philippot, L. *et al.* (2013) Going back to the roots: the microbial ecology of the rhizosphere. *Nat. Rev. Microbiol.* 11, 789–799
- Hunter, P. (2016) Plant microbiomes and sustainable agriculture: deciphering the plant microbiome and its role in nutrient supply and plant immunity has great potential to reduce the use of fertilizers and biocides in agriculture. *EMBO Rep.* 17, 1696–1699
- Chiava, M. *et al.* (2018) Not only priming: soil microbiota may protect tomato from root pathogens. *Plant Signal. Behav.* 13, e1464855
- Egamberdieva, D. *et al.* (2017) Phytohormones and beneficial microbes: essential components for plants to balance stress and fitness. *Front. Microbiol.* 8, 2104
- Spaepen, S. (2015) Plant hormones produced by microbes. In *Principles of Plant-Microbe Interactions* (Lugtenberg, B., ed.), pp. 247–256, Springer International Publishing
- Lu, T. *et al.* (2018) Rhizosphere microorganisms can influence the timing of plant flowering. *Microbiome* 6, 231
- Senthil Kumar, C.M. *et al.* (2018) Multifarious plant growth promotion by an entomopathogenic fungus *Lecanicillium psalliotae*. *Microbiol. Res.* 207, 153–160
- Steidinger, B.S. *et al.* (2019) Climatic controls of decomposition drive the global biogeography of forest-tree symbioses. *Nature* 569, 404
- Delavaux, C.S. *et al.* (2019) Mycorrhizal fungi influence global plant biogeography. *Nat. Ecol. Evol.* 3, 424
- Crawford, K.M. *et al.* (2019) When and where plant-soil feedback may promote plant coexistence: a meta-analysis. *Ecol. Lett.* 22, 1274–1284
- Terrer, C. *et al.* (2016) Mycorrhizal association as a primary control of the CO₂ fertilization effect. *Science* 353, 72–74
- Bragazza, L. *et al.* (2013) Biogeochemical plant-soil microbe feedback in response to climate warming in peatlands. *Nat. Clim. Chang.* 3, 273–277
- Berg, G. *et al.* (2017) Plant microbial diversity is suggested as the key to future biocontrol and health trends. *FEMS Microbiol. Ecol.* 93, fix050
- Leff, J.W. *et al.* (2018) Predicting the structure of soil communities from plant community taxonomy, phylogeny, and traits. *ISME J.* 12, 1794–1805
- Ramirez, K.S. *et al.* (2019) Range-expansion effects on the belowground plant microbiome. *Nat. Ecol. Evol.* 3, 604–611
- Gilbert, J.A. *et al.* (2014) The Earth Microbiome project: successes and aspirations. *BMC Biol.* 12, 69
- Lloyd-Price, J. *et al.* (2017) Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* 550, 61–66
- Scherrer, D. *et al.* (2019) Disentangling the processes driving plant assemblages in mountain grasslands across spatial scales and environmental gradients. *J. Ecol.* 107, 265–278
- Trojelsgaard, K. and Olesen, J.M. (2016) Ecological networks in motion: micro- and macroscopic variability across scales. *Funct. Ecol.* 30, 1926–1935
- Ramirez, K.S. *et al.* (2018) Detecting macroecological patterns in bacterial communities across independent studies of global soils. *Nat. Microbiol.* 3, 189
- Jousset, A. *et al.* (2017) Where less may be more: how the rare biosphere pulls ecosystems strings. *ISME J.* 11, 853–862

27. Wei, X. *et al.* (2019) Rare taxa of alkaline phosphomonoesterase-harboring microorganisms mediate soil phosphorus mineralization. *Soil Biol. Biochem.* 131, 62–70
28. Cerro-Galvez, E. *et al.* (2019) Microbial responses to anthropogenic dissolved organic carbon in the Arctic and Antarctic coastal seawaters. *Environ. Microbiol.* 21, 1466–1481
29. Dawson, W. *et al.* (2017) A small number of low-abundance bacteria dominate plant species-specific responses during rhizosphere colonization. *Front. Microbiol.* 8, 975
30. Quince, C. *et al.* (2017) Shotgun metagenomics, from sampling to analysis. *Nat. Biotechnol.* 35, 833–844
31. Sergaki, C. *et al.* (2018) Challenges and approaches in microbiome research: from fundamental to applied. *Front. Plant Sci.* 9, 1205
32. Blaxter, M. *et al.* (2016) Reminder to deposit DNA sequences. *Science* 352, 780–780
33. Salzberg, S.L. (2016) Databases: reminder to deposit DNA sequences. *Nature* 533, 179
34. Barrett, T. *et al.* (2012) BioProject and BioSample databases at NCBI: facilitating capture and organization of metadata. *Nucleic Acids Res.* 40, D57–D63
35. Yilmaz, P. *et al.* (2011) Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIXS) specifications. *Nat. Biotechnol.* 29, 415–420
36. Gonçalves, R.S. and Musen, M.A. (2019) The variable quality of metadata about biological samples used in biomedical experiments. *Sci. Data* 6, 190021
37. Zhainina, K. *et al.* (2018) Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nat. Microbiol.* 3, 470
38. Balvočiūtė, M. and Huson, D.H. (2017) SILVA, RDP, Greengenes, NCBI and OTT — how do these taxonomies compare? *BMC Genomics* 18, 114
39. Choi, J. *et al.* (2017) Strategies to improve reference databases for soil microbiomes. *ISME J.* 11, 829
40. Blackwell, M. (2011) The fungi: 1, 2, 3 ... 5.1 million species? *Am. J. Bot.* 98, 426–438
41. Locey, K.J. and Lennon, J.T. (2016) Scaling laws predict global microbial diversity. *Proc. Natl. Acad. Sci. U. S. A.* 113, 5970–5975
42. Overmann, J. *et al.* (2019) Relevance of phenotypic information for the taxonomy of not-yet-cultured microorganisms. *Syst. Appl. Microbiol.* 42, 22–29
43. Lynch, M.D.J. and Neufeld, J.D. (2015) Ecology and exploration of the rare biosphere. *Nat. Rev. Microbiol.* 13, 217–229
44. Yarza, P. *et al.* (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat. Rev. Microbiol.* 12, 635–645
45. Konstantinidis, K.T. *et al.* (2017) Uncultivated microbes in need of their own taxonomy. *ISME J.* 11, 2399–2406
46. Nesme, J. *et al.* (2016) Back to the future of soil metagenomics. *Front. Microbiol.* 7, 73
47. Fierer, N. (2017) Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat. Rev. Microbiol.* 15, 579–590
48. Delgado-Baquerizo, M. *et al.* (2018) A global atlas of the dominant bacteria found in soil. *Science* 359, 320–325
49. Nguyen, N.H. *et al.* (2016) FUNGuild: an open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol.* 20, 241–248
50. Fish, J.A. *et al.* (2013) FunGene: the functional gene pipeline and repository. *Front. Microbiol.* 4, 291
51. Gaby, J.C. and Buckley, D.H. (2014) A comprehensive aligned nifH gene database: a multipurpose tool for studies of nitrogen-fixing bacteria. *Database (Oxford)* 2014, bau001
52. Pedro, H. *et al.* (2016) PhytoPath: an integrative resource for plant pathogen genomics. *Nucleic Acids Res.* 44, D688–D693
53. Sasse, J. *et al.* (2018) Feed your friends: do plant exudates shape the root microbiome? *Trends Plant Sci.* 23, 25–41
54. Li, L. *et al.* (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13, 2178–2189
55. Kanehisa, M. *et al.* (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45, D353–D361
56. Althauer, K.P. *et al.* (2015) Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics* 31, 2882–2884
57. Huang, D.W. *et al.* (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57
58. Dutta, C. and Paul, S. (2012) Microbial lifestyle and genome signatures. *Curr. Genomics* 13, 153–162
59. Hacquard, S. *et al.* (2016) Survival trade-offs in plant roots during colonization by closely related beneficial and pathogenic fungi. *Nat. Commun.* 7, 11362
60. Hausmann, B. *et al.* (2019) Long-term transcriptional activity at zero growth of a cosmopolitan rare biosphere member. *mBio* 10, e02189-18
61. Zecchin, S. *et al.* (2018) Rice paddy Nitrospirae carry and express genes related to sulfate respiration: proposal of the new genus 'Candidatus sulfobium'. *Appl. Environ. Microbiol.* 84, e02224-17
62. Hendershot, J.N. *et al.* (2017) Consistently inconsistent drivers of microbial diversity and abundance at macroecological scales. *Ecology* 98, 1757–1763
63. Walters, W. *et al.* (2015) Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems* 1, e00009-15
64. Klindworth, A. *et al.* (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* 41, e1
65. Tremblay, J. *et al.* (2015) Primer and platform effects on 16S rRNA tag sequencing. *Front. Microbiol.* 6, 771
66. Lozupone, C.A. *et al.* (2013) Meta-analyses of studies of the human microbiota. *Genome Res.* 23, 1704–1714
67. Holman, D.B. *et al.* (2017) Meta-analysis to define a core microbiota in the swine gut. *mSystems* 2, e00004-17
68. Duvallet, C. *et al.* (2017) Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat. Commun.* 8, 1784
69. Wang, J. *et al.* (2018) Meta-analysis of human genome-microbiome association studies: the MiBioGen consortium initiative. *Microbiome* 6, 101
70. Callahan, B.J. *et al.* (2016) DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583
71. Amir, A. *et al.* (2017) Deblur rapidly resolves single-nucleotide community sequence patterns. *mSystems* 2, e00191-16
72. Edgar, R.C. (2018) Updating the 97% identity threshold for 16S ribosomal RNA OTUs. *Bioinformatics* 34, 2371–2375
73. Callahan, B.J. *et al.* (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* 11, 2639
74. Rocca, J.D. *et al.* (2019) The Microbiome Stress Project: toward a global meta-analysis of environmental stressors and their effects on microbial communities. *Front. Microbiol.* 9, 3272
75. Sogin, M.L. *et al.* (2006) Microbial diversity in the deep sea and the underexplored 'rare biosphere'. *PNAS* 103, 12115–12120
76. Pedrós-Alió, C. (2012) The rare bacterial biosphere. *Annu. Rev. Mar. Sci.* 4, 449–466
77. Shade, A. *et al.* (2014) Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. *mBio* 5, e01371-14
78. Marchesi, J.R. and Ravel, J. (2015) The vocabulary of microbiome research: a proposal. *Microbiome* 3, 31
79. Hartmann, A. *et al.* (2008) Lorenz Hiltner, a pioneer in rhizosphere microbial ecology and soil bacteriology research. *Plant Soil* 312, 7–14
80. Sharifi, R. and Ryu, C.-M. (2017) Chatting with a tiny below-ground member of the holobiome: communication between plants and growth-promoting rhizobacteria. *Adv. Bot. Res.* 82, 135–160
81. Wiersma, J.H. *et al.* (2016) *World Economic Plants: A Standard Reference* (2nd edn), CRC Press
82. Beattie, G.A. (2018) Metabolic coupling on roots. *Nat. Microbiol.* 3, 396
83. Prosser, J.I. *et al.* (2007) The role of ecological theory in microbial ecology. *Nat. Rev. Microbiol.* 5, 384–392
84. Uroz, S. *et al.* (2019) Plant symbionts are engineers of the plant-associated microbiome. *Trends Plant Sci.* 24, 905–916
85. Baas-Becking, L.G.M. (1934) *Geobiologie; of Inleiding tot de Milieukunde*, WP Van Stockum & Zoon NV

86. van der Gast, C.J. (2015) Microbial biogeography: the end of the ubiquitous dispersal hypothesis? *Environ. Microbiol.* 17, 544–546
87. Peay, K.G. and Bruns, T.D. (2014) Spore dispersal of basidiomycete fungi at the landscape scale is driven by stochastic and deterministic processes and generates variability in plant–fungal interactions. *New Phytol.* 204, 180–191
88. Cadotte, M.W. and Tucker, C.M. (2017) Should environmental filtering be abandoned? *Trends Ecol. Evol.* 32, 429–437
89. Franklin, J. *et al.* (2017) Big data for forecasting the impacts of global change on plant communities. *Glob. Ecol. Biogeogr.* 26, 6–17
90. Wieder, W.R. *et al.* (2013) Global soil carbon projections are improved by modelling microbial processes. *Nat. Clim. Chang.* 3, 909–912
91. Mercado-Blanco, J. *et al.* (2018) Belowground microbiota and the health of tree crops. *Front. Microbiol.* 9, 1006
92. Scheffer, M. *et al.* (2001) Catastrophic shifts in ecosystems. *Nature* 413, 591
93. Saccá, M.L. *et al.* (2017) Ecosystem services provided by soil microorganisms. In *Soil Biological Communities and Ecosystem Resilience* (Lukac, M. *et al.*, eds), pp. 9–24, Springer
94. Kivlin, S.N. *et al.* (2013) Fungal symbionts alter plant responses to global change. *Am. J. Bot.* 100, 1445–1457
95. Finzi, A.C. *et al.* (2015) Rhizosphere processes are quantitatively important components of terrestrial carbon and nutrient cycles. *Glob. Chang. Biol.* 21, 2082–2094
96. Dijkstra, F.A. *et al.* (2013) Rhizosphere priming: a nutrient perspective. *Front. Microbiol.* 4, 216
97. Philippot, L. *et al.* (2009) Biochemical cycling in the rhizosphere having an impact on global change. *Plant Soil* 321, 61–81
98. Kattge, J. *et al.* (2020) TRY plant trait database – enhanced coverage and open access. *Glob. Chang. Biol.* 26, 119–188
99. Durka, W. and Michalski, S.G. (2012) Daphne: a dated phylogeny of a large European flora for phylogenetically informed ecological analyses. *Ecology* 93, 2297–2297
100. van Kleunen, M. *et al.* (2019) The Global Naturalized Alien Flora (GloNAF) database. *Ecology* 100, e02542
101. Cayuela, L. *et al.* (2012) taxonstand: an R package for species names standardisation in vegetation databases. *Methods Ecol. Evol.* 3, 1078–1083
102. The Angiosperm Phylogeny Group (2009) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* 161, 105–121
103. Ma, Z. *et al.* (2018) Evolutionary history resolves global organization of root functional traits. *Nature* 555, 94–97