Cautionary notes on the use of co-occurrence networks in soil ecology

Marta Goberna [a,*], Miguel Verdú [b]

[a] Department of Environment and Agronomy, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA-CSIC), Carretera de La Coruña km 7.5, 28040, Madrid, Spain
[b] Department of Ecology, Centro de Investigaciones sobre Desertificación (CIDE-CSIC), Carretera Moncada-Náquera km 4.5, 46113, Moncada, Valencia, Spain

ABSTRACT

Soil ecology is witnessing exponential growth in the number of studies using co-occurrence network analysis. Researchers reconstruct networks based on the co-occurrence of taxa or genes across soil samples at a wide range of geographic scales - from single aggregates to the whole planet - and taxonomic scopes, some studies targeting specific taxa or guilds to others surveying the whole microbiome as well as micro- and mesofauna. Co-occurrence networks can be very useful to extract simple patterns from complex datasets. Applications include the detection of abiotic and biotic factors that determine community structure, the identification of keystone taxa and their relationship to specific soil functions, and the inference of mechanisms of community assembly. However, networks are more and more often misused and serve as mere graphic tools with no attempt at hypothesis testing. In this perspectives article, we first review the main usage of co-occurrence network analysis in soil ecology during the last decade. We then discuss the applications and caveats of network analysis in soil ecology, leaving apart strictly methodological aspects of network reconstruction, which is beyond the focus of this article. Finally, we include recommendation guidelines – such as the possibility of informing networks with geographic, environmental and/or phylogenetic information – with the hope that this will facilitate network analysis to become a useful tool that helps elucidate meaningful patterns in soil ecology.

1. Where does the use of co-occurrence networks in soil ecology come from?

Co-occurrence is the simultaneous presence of two units in the same place. Co-occurrence matrices have long been a fundamental unit of analysis in community ecology and biogeography to describe the distribution of units (taxa) across multiple locations. The analysis of co-occurrence patterns in ecological communities originally focused on negative co-occurrences, i.e. the so-called checkerboard patterns describing mutually excluding pairs of taxa, and naively assumed that they were indicative of competitive exclusion (Diamond, 1975). Applying appropriate null models to these analyses proved that many co-occurrences did not differ from the random expectation, generating a debate about their validity to reflect community assembly rules (Gotelli, 2000). Later refinements discarding random associations and incorporating abiotic and biotic information improved their predictive capacity (Ulrich et al., 2017). In parallel, network thinking was generalized across many disciplines, including social sciences, computer science, physics, mathematics and biology (Proulx et al., 2005). Graph theory allowed the modelling of all pairwise connections between the units that

co-occur to generate a network, that is, a *collection of units potentially interacting as a system* (Proulx et al., 2005). Networks are usually depicted as a set of nodes, which are the vertices or points representing the units, and edges or links representing pairwise connections between units. Connections represent any type of potential interaction, including gene regulation, disease transmission, metabolic or neural interactions. In ecology, the use of complex networks increased since the 1990′s as a tool to represent the ecological interactions occurring between the species within a community (Ings and Hawes, 2018), leading to an increase in the number of publications of, for instance, plant-pollinator or plant-seed disperser networks (Fig. S1). Ecological interaction networks are based on direct observations of organisms with a well-known ecological role, nodes depicting observed individuals (e.g. bee of species *x* and plant of species *y*) and links reflecting interactions often containing information on the directionality of the interaction (e.g. pollination of species *y* by species *x*). Underlining 'direct observations' and 'organisms with a well-known ecological role' is essential as both are key features of this approach, which are typically not met by soil co-occurrence networks.

Soil ecology has emerged both conceptually and methodologically

---

* Corresponding author.
  E-mail address: marta.goberna@inia.es (M. Goberna).

from sister disciplines, and the use of network analysis is no exception. As in plant or animal ecology, the taxonomic description of soil fauna, and the cultivation or sequencing of microorganisms across multiple samples is translated into co-occurrence matrices, which are more and more frequently analysed based on co-occurrence networks. These networks result from correlations between pairs of taxa (or genes) across multiple soil samples to identify those pairs that are significantly aggregated (i.e. co-present or co-absent) or segregated (i.e. mutually excluding each other). Despite being based on correlative patterns, with all their statistical assumptions and lack of scientific evidence for the real existence or directionality of most potential interactions, co-occurrence networks tend to be interpreted as if they were interaction networks. That is to say, species aggregation is assigned to positive ecological interactions (e.g. cooperation for resource exchange or acquisition) and segregation to negative interactions (e.g. competition for space or resources). This is a flawed approach since several processes can produce the same patterns. For instance, two species may be co-present (or co-absent) across several sites not only because they benefit from (or require) each other's presence, but also if they have similar limitations for dispersal (e.g. neither species has means to reach distant areas), common niche requirements (e.g. both thrive under acidic conditions or are sensitive to saline conditions), or are determined by higher-order interactions (e.g. both are attacked by the same pathogen or predator, or are involved in rock-scissors-paper like competition in which three species can co-exist based on their differential ability to produce or tolerate toxins; Hibbing et al., 2010). For this reason, several authors have called for caution when using co-occurrence networks to describe interactions or community assembly processes (Barner et al., 2018; Blanchet et al., 2020). To prevent soil ecology repeating the same mistakes as other disciplines, we advocate for a reflection on the use of co-occurrence networks. We start by delineating a historical perspective based on a literature review to understand how co-occurrence networks are being used and why they are currently so widespread in soil ecology. We also review the most common applications, as well as the pitfalls that lead to the misuse or misinterpretation of network data. Finally, we provide recommendation guidelines that we hope will help to extract ecologically meaningful patterns from network analysis.

2. Why have co-occurrence networks become so popular in soil ecology?

Reconstruction of co-occurrence networks needs to be based on robust datasets and very large numbers of replicates, frequently resulting in a time-consuming process that requires computational power. Relevant decisions need to be made on how to pre-process abundance data or filter rare taxa, and to select an appropriate tool for network inference, null model, level of resolution of taxonomic assignments or threshold of statistical significance, all of which can have a strong impact on the outcome (Connor et al., 2017; Faust, 2021; Röttjers and Faust, 2018; Weiss et al., 2016). Theoretical background is also needed to interpret the components and emerging patterns of network structure (Cogoul et al., 2019; Kurtz et al., 2015; Li et al., 2016). These considerations have not precluded network analysis to become a popular exploratory tool for the study of soil biological communities. To better understand the elements that spurred network thinking in soil ecology and quantify its spread through time we performed a literature search. We specifically searched for scientific literature containing the terms 'soil', 'co-occurrence' and 'network' in the core collection of the Web of Science in April 2021 and obtained 485 research articles and reviews (Table S1). This exercise confirmed that the number of articles and citations, as well as the proportion of articles using networks with respect to the total number of articles published in the same journals and time period, has increased exponentially during the last decade (Fig. 1). The distribution of these articles across 114 journals, mostly categorized within Microbiology (29%), Environmental sciences (28%), Soil Science (22%) and Ecology (13%), is given in Table S2.
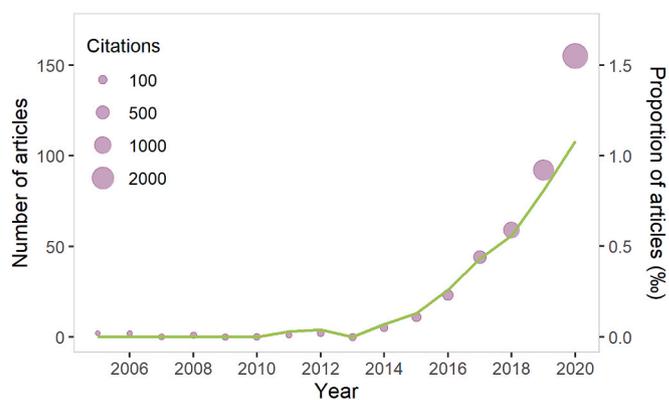

Fig. 1. The number of published articles including the terms 'soil', 'co-occurrence' and 'network' detected in the Web of Science Core Collection has increased exponentially during the last decade (purple bubbles). Bubble size is proportional to the number of citations (detected on 6 April 2021). The proportion of articles researching soil co-occurrence networks with respect to the total number of articles published in the same journals and time period has increased following a similar trend (green line).

The two earliest manuscripts in our collection that indeed used network analysis to explore co-occurrence patterns of soil communities appeared almost simultaneously: Boutsis et al. (2011) working with nematodes and Barberán et al. (2012) with microorganisms. Both acknowledged inspiration by Proulx et al. (2005), who wrote a comprehensive review on the study of gene, protein and metabolic networks, and the stability and dynamics of biological communities. This point raises the first fundamental question that needs a cautionary note. While the idea of using networks to explore co-occurrence patterns of soil organisms is very attractive, it implies the extension of network analysis from the inspection of interactions at the individual level, based upon extensive data and experimental evidence (e.g. Kanehisa and Goto, 2000), to the analysis of gene-gene (or taxon-taxon) associations at the community level in a highly heterogeneous habitat where most organisms and the interactions among them and with their environment are virtually unknown. The transference of this technique to soil ecology would seem in principle less problematic for organisms such as nematodes whose functional role is well defined, since genera have coherent ecological strategies (e.g. bacterivorous, fungivorous, phytophagous, nematode predators; Yeates et al., 1993). However, simply by looking at the number of citations that Boutsis et al. (2011) and Barberán et al. (2012) have received (5 *vs*. 822 as of 6 April 2021), one can expect that the use of network analysis is nowadays radically less widespread in soil animal ecology than in soil microbial ecology. To formally analyse this expectation and the main usage of co-occurrence networks in soil ecology during the last decade, we screened ca. 65% of all records found (313 out of 485 records). Due to the uneven temporal distribution of the publications, we checked all articles published in years with less than 50 records (2011–2017), and randomly sampled approximately 60% of the studies published from 2018 onwards. We excluded 54 studies from further analysis based on four criteria: the studied environment (other than soil; 37 records), the method used (other than co-occurrence network analysis; 10), the article type (reviews without experimental data; 2) and the target of network analysis (not biological or exclusively targeting plants; 5). We categorized the remaining 259 articles based on the dominant land use, the spatial scale of study and the biological target (Fig. 2).

Target biological groups included microorganisms (archaea, bacteria, oomycetes, fungi), microfauna (protists, nematodes) and mesofauna (enchytraeids, oribatids, collembola) (Target in Fig. 2). As expected, the most remarkable trend was that a vast majority of all studies focused on the soil microbiome based on high-throughput DNA sequencing data, mainly using phylogenetic (archaeal and/or bacterial 16S rRNA gene or
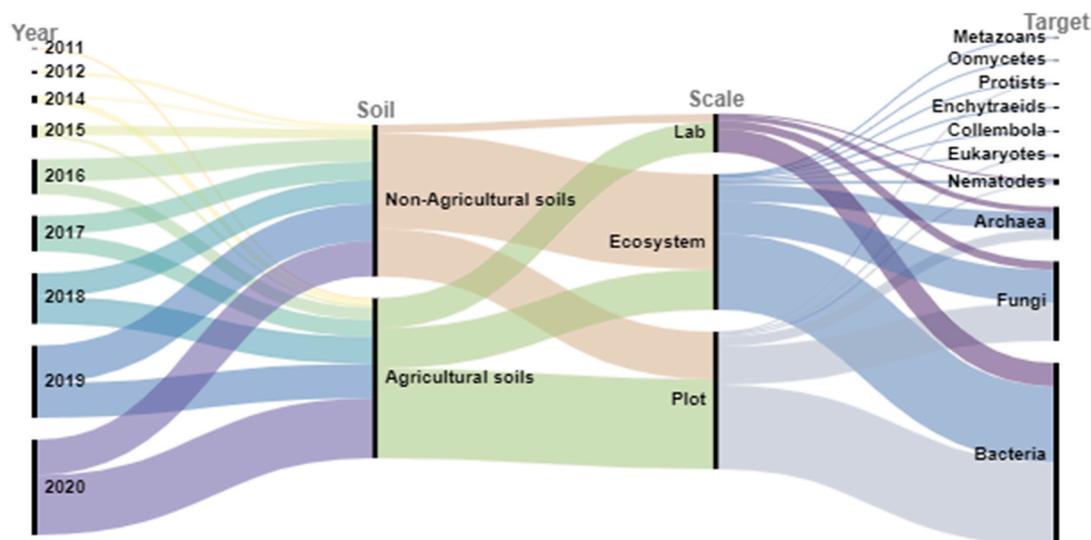
Fig. 2. Land-use, spatial scale and biological target addressed in articles analysing soil co-occurrence networks from 2011 to 2020 (Year). Soil: Half of all surveys were performed in Agricultural soils (48%) and half in Non-agricultural soils (48%), mostly including forest soils and grasslands but also wetlands, floodplains, ornithogenic soils, abandoned mining sites, contaminated soils in refineries, soils in urban greenspaces or cave soils. The remaining 4% analysed Agricultural and Non-agricultural soils simultaneously. Scale: Most studies (48%) were performed at the Plot scale (i.e. single sites, plantations, field experiments, and common gardens), followed by studies at the Ecosystem scale (i.e. landscape, country, continent, or global studies). Only 13% of studies were performed at the Laboratory scale, using Petri dishes or pots in incubation or growth chambers and greenhouses. Target: A vast majority of studies (98.8%) focused on the soil microbiome using high-throughput sequencing, either alone or together with soil micro- or mesofauna. The alluvial diagram was performed with RAWGraphs (Mauri et al., 2017).

the fungal internal transcribed spacer) but also functional molecular markers associated with nutrient cycling (e.g. Jones and Hallin, 2019; Liu and Conrad, 2017; Tu et al., 2016), and metal or antibiotic resistance (e.g. Hu et al., 2016). Among soil microorganisms, bacteria were by far the main group of interest that we detected in our review (Fig. 2). Bacteria were studied either exclusively (51% of all studies) or along with other biological groups (86%). Even if one third of all studies focused on more than one biological group simultaneously, this was predominantly (84 out of 93 papers) to study several microbial groups and seldom to jointly analyse the microbiota, micro- and mesofauna. A specific search of the term "food web" within our collection yielded a single study that combined observations of several groups of soil microorganisms and animals (Creamer et al., 2016). This shows that the inspection of multi-trophic food webs, as has been done in other ecosystems (e.g. Montoya et al., 2015), does not explain the increase in the use of network analysis in soils.

Beyond the historical reasons mentioned above, practical reasons underlie the growing trend in the use of network analysis in soil microbial ecology and environmental microbiology. The most obvious reason is that it is increasingly easy and cheap to sequence soil microbiota utilizing metabarcoding techniques based on well-established molecular markers. The easiness to obtain microbial sequences has presented the opportunity to delve into microbial diversity patterns in soils, helping to understand the decomposer system and incorporating microbial ecology into the general ecological theory (Prosser et al., 2007). Analysing the extremely diverse soil microbial communities, with their multifaceted interactions, requires tools that allow extracting easy-to-analyse patterns from complex datasets. Network analysis not only allows modelling and simplifying the complexity of microbial datasets (Layeghifard et al., 2017), but also provides attractive graphical representations. However, these positive aspects of networks as an analysis tool are somehow turning against the epistemological advance of the discipline. Rather than aiming at hypothesis testing, networks are frequently part of a long list of descriptive analyses of microbiome datasets, and are used more and more often as a mere pictorial tool that adds little information to the scientific discourse. Technical developments have popularized the survey of soil microbiomes using co-occurrence networks among a broad scientific community beyond

microbial ecologists or network specialists. This is creating an avalanche of microbial datasets, whose peculiar nature is often neglected, posing the second major cautionary note (Leite and Kuramae, 2020). High-throughput sequencing microbial data are compositional and sparse: They describe proportions with an irrelevant total since sequencing instruments have a fixed-size capacity (Gloor et al., 2017) and have an excess zeros due to the absence of a wealth of taxa across many samples (Kaul et al., 2017). In addition, soil microbial communities are dominated by a few taxa and have a long tail of low-abundance or rare species (Delgado-Baquerizo et al., 2018). Inherent compositionality, sparsity and rarity of microbiome datasets can lead to the detection of false correlations (Layeghifard et al., 2017; Weiss et al., 2016, 2017). Strategies to transform and analyse this type of data have been extensively reviewed (Gloor et al., 2017; Kaul et al., 2017; Leite and Kuramae, 2020), and need to be taken into account to construct robust microbial networks (Cogoul et al., 2019; Kurtz et al., 2015; Li et al., 2016; Röttjers and Faust, 2018). Beyond data management and network reconstruction, several cautionary notes are needed related to the interpretation of network results.

3. How networks can be useful for soil ecologists, if not misused or misinterpreted

The most frequent applications of co-occurrence network analysis in soil ecology that we detected in our review are schematised in Fig. 3. Most articles base the construction of co-occurrence networks on taxon co-occurrence matrices reflecting the abundance (or presence) of taxa across plots (Fig. 3). As an alternative, some studies construct functional gene co-occurrence networks or combine both taxonomic and functional molecular markers (Chen et al., 2016; Hu et al., 2016). Even if the methodological aspects of network reconstruction are beyond the focus of this manuscript, it is worth mentioning that methodology is often poorly described so it is not possible to evaluate the goodness of (or reproduce) network analysis. Just to name some key aspects that caught our attention, many studies base their analyses on a single correlation metrics (e.g. Pearson correlation), which cannot properly deal with biases inherent to microbial datasets and are outperformed by inference tools (Kurtz et al., 2015; Layeghifard et al., 2017). Also, null models are
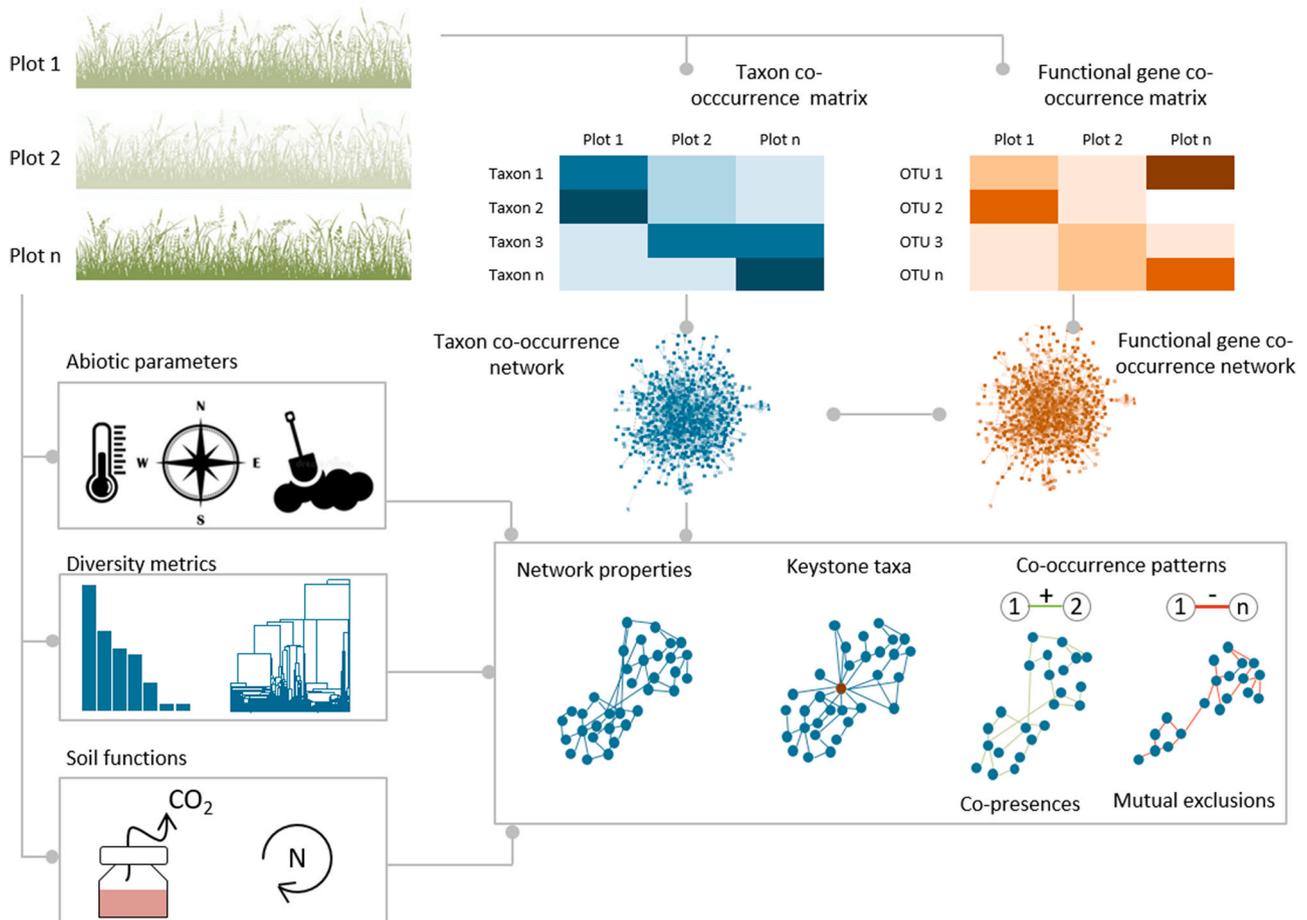
Fig. 3. Overview of frequent applications of co-occurrence networks in soil ecology.

not employed systematically to test whether observed patterns differ from the random expectation (Röttjers et al., 2021), taking us back to the same debate ignited forty years ago for macro-organisms (Connor and Simberloff, 1979; Gotelli, 2000). Other authors do not mention false discovery rate corrections due to multiple testing, posing the question of whether statistically significant associations are meaningful. And, what is probably the most common pitfall, networks are usually performed using too few replicates (as few as three), ignoring that the precision of network inference is highly dependent on sample size (Kurtz et al., 2015). Finally, irrespective of the application, we did not detect any study checking the replicability of soil networks, which is likely not a minor issue considering recent results from other biological systems (Barroso-Bergadà et al., 2021).

The basic descriptive analysis of network results involves the detection of non-random co-occurrence patterns between microbial community members (Barberán et al., 2012; Purahong et al., 2016). These analyses originally served to refute the long-standing view that terrestrial microorganisms tend to be ubiquitous and stochastically distributed based on their high dispersal rates (Horner-Devine et al., 2007; O'Malley, 2008). Regardless of downstream applications, using null network models is essential to discard spurious associations that may confound ecological interpretations of the soil community (Connor et al., 2017). In the next step, co-occurrence networks are typically explored in three ways (Fig. 3), as follows.

1. *Network properties.* Researchers use simple statistics to describe network topological properties, predominantly (but not only) in terms of i. size or number of nodes, ii. connectivity of each node to others, iii. clustering of each node to its neighbours, iv. average or shortest path length between nodes, and v. network organization into modules, which are delimited areas of dense connections that have sparser connections with other areas. Researchers analyse the contribution of environmental factors (i.e. geographic, climatic and soil factors) to network structural properties, so as to reach conclusions on the abiotic determinants of community structure (e.g. Ma et al., 2016). In addition, the description of network structure allows a direct comparison between observed and random networks, or between networks resulting from different environmental conditions or experimental treatments. For instance, network stability can be evaluated through specific biological processes, e.g. litter decomposition (Purahong et al., 2016), or after ecological disturbance (de Vries et al., 2018). Network stability refers to the ability of network topological properties to remain unchanged over time. Changes in the structure of the network can be produced by altering the composition and abundance of taxa co-oscillating in a community and, depending on their functional redundancy, this might concomitantly alter the associated ecosystem functions. High modularity and low connectivity are interpreted as network stabilising properties (de Vries et al., 2018), since they may allow perturbations to stay confined within specific modules preventing their spread through the whole network. Despite the argument is reasonable, we need to bear in mind that co-occurrence networks are based on correlative patterns, and biological conclusions cannot be directly drawn in cases where there is no biological evidence or background information (see Carr et al. (2019) and Weiss et al. (2016) for reviews of the strengths, limitations and recommendations of correlation-based methods in microbial ecology). Once again, network interpretation cannot be directly imported from networks

depicting interactions or flows. As an illustration, taxa shaping a module cannot be assumed to perform similar functions (e.g. Jiao et al., 2016) as in metabolic or computational networks (Newman, 2006), since this assumption disregards that modularity in soils can strongly respond to sources of spatial or abiotic variation (e.g. Jones and Hallin, 2019). Similarly, the spread of a perturbation through a co-occurrence network cannot be thought of as the transmission of an agent of disease or a metabolic flow in interaction or metabolic networks. Such an interpretation would require informing co-occurrence networks with additional functional layers. In epidemiology, virus transmission (functional layer) can be explained based on the network of contacts (co-occurrence layer). A parallelism in soil ecology can be drawn, for instance, with the superimposition of the patterns of co-existence of antibiotic resistance genes (ARGs; functional layer) and 16S rRNA genes (co-occurrence layer) that some authors use to reach conclusions on potential ARG hosts and horizontal gene transfer events. Such an application can be useful as a tool to generate hypotheses that need further experimental testing (e.g. Chen et al., 2016; Hu et al., 2016).

2. *Keystone taxa.* Another frequent way of exploring co-occurrence networks involves the classification of taxa (nodes) based on topological properties, such as the degree of node connectivity with other nodes within its own module or in other modules. Such structural classifications are interpreted in terms of the ecological role of each taxon within the community, e.g. as having central or peripheral positions in the maintenance of community structure (Zhou et al., 2011). The most frequent application in this context is the identification of keystone taxa, i.e. species with a supposedly crucial role in community structure and function irrespective of their abundance (Banerjee et al., 2018). Keystone taxa are frequently determined based on their centrality in the network, meaning that they are highly connected nodes whose disappearance compromises network structure and stability (Martín González et al., 2010; Berry and Widder, 2014). Other methods rather than centrality have been also used, but are similarly based on topological properties (Layeghifard et al., 2017). However, diagnosing a keystone species just on the basis of its topological role within the network is not useful unless we are able to identify its major ecological role in the ecosystem. Some authors explore how keystone taxa are affected by specific conditions or treatments, for instance, agricultural management (Li et al., 2017; Hartman et al., 2018) or soil contamination (Jiao et al., 2016). Given their theoretically key functional roles, researchers also address how the abundance of keystone taxa correlates to specific soil functions, such as the rates of organic matter decomposition (Banerjee et al., 2016). The term "keystone species" was originally coined to designate species whose influence on the associated assemblage were disproportionately greater than would be expected from its abundance. However, the term was later distorted to include an esoteric list of meanings that could make the term useless (Paine, 1995). The fact that a taxon is connected to many other taxa in the network could be just a consequence of its abundance and therefore, strictly speaking, should not be considered a keynote species. As for other applications, it should be noted that abundance-weighted estimations or targeted follow-up experiments should be necessary to identify keystone species, as has been done in food webs (see Libralato et al., 2006 and references herein).

3. *Co-occurrence patterns.* Finally, many studies explore the spatial or temporal patterns of co-occurrence between pairs of taxa. These can be analysed in terms of taxa aggregation across samples, i.e. pairwise co-presences or co-absences based on positive correlations, and segregation, i.e. mutual exclusions based on negative correlations (Dini-Andreote et al., 2014). Co-occurrence patterns are typically explored based on taxonomic assignments in order to analyse whether organisms belonging to the same taxon (e.g. genus, phylum) tend to co-occur or exclude each other more often than expected by chance (e.g. Hartman et al., 2018; Nielsen et al., 2014). Other

authors assign ecological or functional categories to the taxa, e.g. life strategies of generalists versus specialists based on their degree of site occupancy (Barberán et al., 2012), to reach further conclusions on the ecological mechanisms behind community structure. Finally, the patterns of co-occurrence are used to infer community assembly processes, i.e. the prevalence of stochastic or niche-based processes, such as abiotic filtering and (positive or negative) ecological interactions (e.g. Dini-Andreote et al., 2014; Nielsen et al., 2014). The inference of assembly mechanisms from diversity patterns is prone to pitfalls based on simplified (or false) assumptions in community ecology (Münkemüller et al., 2020) that also apply to the interpretation of soil networks.

The most frequent mistake is to equate the patterns of co-presence and mutual exclusion to ecological interactions with no other consideration, despite the repeated reports against this practice (Barner et al., 2018; Freilich et al., 2018; Blanchet et al., 2020). In soils, authors interpret co-occurrences as interactions even when working at coarse spatial scales, which are dominant in the literature according to our review (Fig. 2, Scale): Half of all studies were performed at the plot scale (i.e. single sites, plantations, field experiments, and common gardens; 48% of all articles) or even coarser scales encompassing one or multiple ecosystems (i.e. landscape, country, continent, hemisphere or planet; 39%). We specifically searched for a temporal increase in the tendency to work at broad scales, which we expected but could not confirm. Instead, we found that the scale of analysis was associated to land use, broader scales prevailing for the survey of non-agricultural soils (Fig. 2, Soil).

The interpretation of ecological interactions is meaningful when the spatial scale of analysis matches the scale of interactions (Blanchet et al., 2020). Spatial scaling generally affects the interpretation of assembly mechanisms from diversity patterns (Münkemüller et al., 2020), but this issue is particularly serious when considering microscopic organisms living in soils. Soils are composed of minerals and organics spanning a full spectrum of materials in terms of size and chemical nature that aggregate into a multi-dimensional matrix. This solid matrix configures a labyrinthine pore space, where resources are scattered and air and water in continuous competition. The highly heterogeneous nature and dynamism of soil structure strongly determines ecological interactions (Erktan et al., 2020). Under these conditions, microbial communities are spatially structured at very fine scales, even within soil aggregates (Szoboszlay and Tebbe, 2021). In addition, interactions between microbial cells in soils are expected to occur at the range of micrometres as in synthetic communities (Co et al., 2020). However, we typically sample soils at the millimetre to centimetre scale, and then mix this complex habitat prior to DNA extraction necessarily obscuring real microbial patterns (Armitage and Jones, 2019). Thus, even when working at fine scales, interpretations need to be cautious. It should be noted that the scale of spatial resolution does not equally limit the detection of positive and negative interactions, as predicted by mathematical models (Araújo and Rozenfeld, 2014). It is generally easier to discern co-occurrences due to positive interactions regardless of the scale (Araújo and Rozenfeld, 2014), and indeed co-occurrence networks frequently reveal larger proportions of aggregated than segregated pairs (Freilich et al., 2018; Goberna et al., 2019). Freilich et al. (2018) suggested that positive interactions involving habitat engineers, which expand the niche for beneficiary species, might leave the most detectable signals. Serving to illustrate this idea, microorganisms involved in biofilm formation, such as Cyanobacteria, Planctomycetes or Chloroflexi show positive co-occurrences in soil networks (Goberna et al., 2019). On the opposite extreme, competitive interactions are very sensitive to spatial scales, and difficult to detect at coarser scales of analysis (Araújo and Rozenfeld, 2014). In silico simulations show that co-occurrence networks can detect competition (Pérez-Valera et al., 2017), with different success depending on the inference tool (Weiss et al., 2016). Still, there are

both mathematical and biological limitations to detect the signature of competition (e.g. Dallas and Melbourne, 2019) and this should be kept in mind when assigning patterns to processes. Considering the most common pitfalls that we have detected in the literature we provide the following recommendation guidelines.

4. Recommendation guidelines for the use of co-occurrence networks in soil ecology

Main recommendations are listed in Box 1 and developed below.

4.1. Work with robust datasets

Compositionality, rarity and sparsity are idiosyncratic features of microbiome datasets. Strategies to transform and analyse this type of data have been reviewed and should be considered (Cogoul et al., 2019; Gloor et al., 2017; Kaul et al., 2017; Kurtz et al., 2015; Leite and Kuramae, 2020; Weiss et al., 2017). Several authors have compared available methods for network reconstruction to address how they deal with such biases (Cogoul et al., 2019; Layeghifard et al., 2017; Röttjers and Faust, 2018; Weiss et al., 2016).

4.2. Perform good network reconstructions

There are excellent reviews available on the various specialized methods for network inference, their strengths and limitations, as well as best practices (Berry and Widder, 2014; Faust, 2021; Faust and Raes, 2012; Layeghifard et al., 2017; Röttjers and Faust, 2018; Weiss et al., 2016). Network specialists keep on developing new tools (e.g. Röttjers et al., 2021; Tackmann et al., 2019), so end-users need to stay updated on technical advances. Regardless of the method of choice, the performance of network inference increases with sample size (Kurtz et al., 2015). Thus, it is key to avoid low sample sizes to construct robust networks. Using in silico microbial datasets, Kurtz et al. (2015) calculated that over 1300 samples are needed to achieve an almost perfect performance, which is unrealistic for most soil studies but helps understanding the magnitude of the problem of using low sample sizes. Network analyses may correctly detect (true positives, TP) or fail to detect (false positives, FP) correlations that actually occur. Similarly, the absence of correlations may be correctly detected (true negatives, TN) or incorrectly predicted (false negatives, FN). Network specificity, calculated as TN/(TN + FP) in simulated networks, is reasonably good using a minimum of 25 sites (Berry and Widder, 2014). Network performance in terms of the proportion of correctly detected correlations is also acceptable with a minimum of 30 sites (Brisson et al., 2019). Thus, a sample size of 25–30 sites can be considered reasonable. Still, network sensitivity, calculated as TP/(TP + FN), improves using up to 100 sites, and it is lost faster than other metrics of performance as environmental heterogeneity increases (Berry and Widder, 2014). Finally, the main methodological aspects of network reconstruction (data generation, normalization and filtering, network inference tool, null model, multiple test corrections, etc.) need to be systematically described, and raw data need to be findable and accessible, to ensure reproducibility.

4.3. Check network replicability

Community ecologists usually calculate metrics like richness or diversity in replicated sampling units to test for a particular hypothesis. However, this is not a common procedure when co-occurrence network metrics are used. Co-occurrence networks should be replicated, i.e. multiple networks calculated based on real replicates, to take into account the variability due to environmental heterogeneity and sampling biases. We conveniently ignore the replicability of soil co-occurrence networks, but it is expected to be low as in other networks like those of fungi in plant leaves (Barroso-Bergadà et al., 2021). Future studies need to incorporate this aspect.

4.4. Beware of taxonomic delimitation

Most soil co-occurrence networks are reconstructed with nodes representing operational taxonomic units (OTUs) arbitrarily defined on the basis of sequence similarity and, logically, network parameters and co-occurrence patterns are widely dependent on the threshold used to delimitate these OTUs (Cardinale et al., 2015; Hemprich-Bennett et al., 2021). On the other side, collapsing OTUs at high taxonomic levels (i.e., orders) could lead to inconsistent conclusions. For example, Williams

---

Box 1
Recommendation guidelines for the use co-occurrence networks in soil ecology.

1 WORK WITH ROBUST DATA
Make informed decisions on abundance data generation, pre-processing and filtering depending on the nature of your dataset

2 PERFORM GOOD NETWORK RECONSTRUCTIONS
Check the excellent reviews on the methods for network inference, which are continuously updated

3 CHECK NETWORK REPLICABILITY
Soil ecology needs to address network reproducibility across replicated sampling units

4 BEWARE OF TAXONOMIC DELIMITATION
Comparing networks with OTUs delimited at different cut-offs needs attention

5 RECALL THAT DATA INTERPRETATION IS LINKED TO THE SCALE OF THE STUDY
Caution if working at coarse scales to extract biological conclusions

6 DO NOT INTERPRET CO-OCCURRENCE NETWORKS AS INTERACTION NETWORKS
Mathematics and experimentation have shown that it does not work

7 ADD ENVIRONMENTAL INFORMATION TO NETWORK ANALYSIS
Geographic, climatic and soil factors are main determinants of network structure and emerging patterns

8 INFORM YOUR NETWORKS WITH BIOLOGICAL BACKGROUND KNOWLEDGE
Make the best out of your phylogenetic and functional markers, perform experimental validation, and collect trait data if you intend reaching mechanistic conclusions

et al. (2014) have shown that the most connected nodes in a soil co-occurrence network were Bacillales, Actinomycetales, and Clostridiales when the network was constructed with correlations >0.75 but Thermoleophilales, Desulfovibrionales, and Sphingobacteriales when the correlation cut-off was relaxed to 0.5. This issue can be especially problematic when a network containing taxa form different ecosystems are constructed (Williams et al., 2014). Assigning amplicon sequence variants (ASVs), which represent actual sequences and are not collapsed into OTUs, may be a way to circumvent this issue.

4.5. Recall that data interpretation is linked to the scale of the study

Networks constructed at coarse scales could be informative of the abiotic processes determining the geographical distribution of taxa, but not of ecological interactions. While abiotic filtering can result from processes occurring at broad spatial scales, biotic interactions necessarily occur at the scales where individuals coexist (Weiher and Keddy, 1995). Supporting this assertion, mathematical models have demonstrated that checkerboard patterns of co-occurrence generated by negative interactions can only be discerned at fine spatial scales (Araújo and Rozenfeld, 2014). Furthermore, microbial community composition can change at the scale of single soil aggregates (Szoboszlay and Tebbe, 2021). Therefore, even at fine spatial scales, co-occurrence cannot be uncritically assigned to interactions. The following recommendations try to build in that direction.

4.6. Do not interpret co-occurrence networks as ecological interaction networks

Interpretation of soil ecological networks has been inherited to a significant extent from sister disciplines, which frequently use interaction networks (i.e. direct recording of ecological interactions) rather than co-occurrence networks (i.e. correlation patterns across multiple assemblages). However, as stated above, co-occurrence patterns cannot be directly equated to ecological interactions (Barner et al., 2018; Blanchet et al., 2020). Co-occurrence metrics are based on symmetric correlation coefficients whereas most of the ecological interactions are asymmetric. Interaction asymmetry in ecological networks is a result of abundance differences, with rare taxa tending to be asymmetrically affected by abundant partners (Vázquez et al., 2007). Strong abundance differences across taxa exist in soil microbial communities (Delgado-Baquerizo et al., 2018) and therefore, interaction asymmetries are also expected. There is also experimental evidence in intertidal ecosystems supporting that co-occurrence networks, even when they are extremely well resolved with a spatially intensive sampling, have a poor correspondence with the known interaction network (Freilich et al., 2018). Additional information beyond the sign of the correlation is required to identify ecological interactions from network links (Carr et al., 2019). We recompile some options below.

4.7. Add environmental information to network analysis

Several approaches are available to take into account abiotic factors when interpreting networks, including options to incorporate them in network inference (e.g. Faust et al., 2015; Duarte Ritter et al., 2021), among others (Faust, 2021). In addition, geostatistical modelling can be combined with co-occurrence networks, as Jones and Hallin (2019) did to find the (geographic and edaphic) sources of variation of network topological properties. We also find useful the conceptual framework that Blois et al. (2014) proposed to discern the influence of geographic and environmental factors in co-occurrence patterns. We applied this framework to soil microbiota and provide a customary R code in Goberna et al. (2019). Using that framework to understand the assembly of soil bacterial communities in drylands, we detected that a majority of co-presence and mutual exclusion patterns can be explained based on abiotic niche preferences. It is important to underline that not only

co-presences can be the result of niche overlap, but also mutual exclusions mostly result from niche segregation (Goberna et al., 2019). Thus, having a good knowledge of the study system and gathering a large collection of geographic, climatic and soil physical and chemical parameters is essential to interpret network results. In addition, technical limitations to quantify this type of parameters at the micrometre scale need to be overridden to adequately characterise the microbial habitat. Theoretically, the variation that cannot be explained by the environment could be due to ecological interactions, and there are algorithms available to specifically subtract the habitat effects (Brisson et al., 2019). On the other hand, simulation exercises show that networks are able to capture the co-occurrence patterns generated by interactions (Pérez-Valera et al., 2017; Weiss et al., 2016). Still, detecting the signature of interactions can be challenging, particularly under significant abiotic filtering (Berry and Widder, 2014).

4.8. Inform your networks with biological background knowledge

Performing ad-hoc experiments is the most convenient option to validate network results. For instance, Cardinale et al. (2015) used FISH-CLSM (Fluorescence in situ hybridization−confocal laser scanning microscopy) to validate the interactions detected with co-occurrence networks in lettuce roots. Kaupper et al. (2021) coupled SIP (stable isotope probing) using ^{13}C–CH$_4$ to co-occurrence analysis based on the ^{13}C-enriched 16S rRNA gene fraction to evaluate the impact of methanotrophic activity on network structure during desiccation-rewetting. These approaches can be used to address specific questions or confirm certain interactions. But we need to admit that, currently, there is no straightforward option to validate the array of co-occurrences detected through networks when analysing soil microbiomes. Is then time to stop constructing soil microbial networks? We sincerely believe that there are intermediate options between uncritically assuming the results of a statistical exercise and abandoning the approach. Gathering data that facilitate the interoperability and integration of datasets and developing more accurate statistical methods will help to predict the structure of species interaction networks (Strydom et al., 2021).

In addition, while new techniques are devised, we propose that more information can be obtained from the phylogenetic and functional markers used for network reconstruction. Phylogenetic information, such as the evolutionary distance between co-occurring or mutually excluding pairs, can help confirm whether network results fit theoretical assumptions (Goberna et al., 2019). This option is highly dependent on phylogenetic reconstructions, which should be technically robust (i.e. replicated, matching well-resolved phylogenies, etc.), particularly if based on a single phylogenetic marker. In addition, background ecological knowledge on the organisms that bear specific functional markers can be used to pose conceptual frameworks, as e.g. to infer the community assembly mechanisms of denitrifiers carrying either *nirS* or *nirK* genes (Goberna et al., 2021). Other branches of ecology, in an attempt to identify the biological mechanisms underlying interaction networks have moved, after a hectic period of algorithmic craziness, towards trait-based approaches (Blüthgen, 2010; Dormann et al., 2009; Rafferty and Ives, 2013). In the meantime, classical taxonomic studies in microbiology are being drastically reduced, producing a concomitant loss of phenotypic information (i.e. diagnostic traits) that would be needed to explain microbial network patterns (Chan et al., 2012). Generating ad-hoc information for microbial traits that confer tolerance to abiotic conditions or competitive abilities (Goberna et al., 2014) could be key to assign community assembly processes in future studies using co-occurrence networks.

Most, if not all, studies depicting soil co-occurrence networks include organisms involved in multiple interaction types which are currently explored as unipartite networks. This means that organisms belonging to different biological and/or trophic groups are combined in a single community matrix which is used to compute a single network. This approach may be valid to generate hypotheses (if all technical demands

are met), but soil ecology needs to move a step forward to build real interaction networks and start disentangling the soil food webs (Morriën, 2016). Attempts in this direction will require tracking the fluxes of matter as has been done for instance with fine-scale SIP experiments targeting decomposers (López-Mondéjar et al., 2018). In other ecosystems, combining such experiments with direct observations (e.g. predation) and gut content analysis have allowed building interaction networks including several trophic levels (Montoya et al., 2015). Such a complexity can be captured in the future using multilayer networks, each layer representing different interaction types, sampling patches or times (Pilosof et al., 2017).

5. Concluding perspectives

The tendency to work with co-occurrence networks will likely keep on growing in the coming years, since the detection of patterns in complex soil networks in the era of big (genomic) data could be substantially improved with the development of machine learning algorithms (Pichler et al., 2020; Xun et al., 2021). However, this trend might be disconnected from the ecology of communities if biological interpretations are unclear and incautious, as has been before with ecological interaction networks (Blüthgen, 2010). Soil networks need to be properly reconstructed, based on robust datasets, null models and inference tools, and their replicability addressed. Collective efforts to provide soil networks with a biological context and experimentally validate interactions at appropriate scales are needed to truly understand and portray the complex architecture of soil biodiversity.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi. org/10.1016/j.soilbio.2021.108534.

References

Araújo, M.B., Rozenfeld, A., 2014. The geographic scaling of biotic interactions. Ecography 37, 406–415. https://doi.org/10.1111/j.1600-0587.2013.00643.x.

Armitage, D.W., Jones, S.E., 2019. How sample heterogeneity can obscure the signal of microbial interactions. The ISME Journal 13, 2639–2646. https://doi.org/10.1038/s41396-019-0463-3.

Banerjee, S., Kirkby, C.A., Schmutter, D., Bissett, A., Kirkegaard, J.A., Richardson, A.E., 2016. Network analysis reveals functional redundancy and keystone taxa amongst bacterial and fungal communities during organic matter decomposition in an arable soil. Soil Biology and Biochemistry 97, 188–198. https://doi.org/10.1016/j.soilbio.2016.03.017.

Banerjee, S., Schlaeppi, K., van der Heijden, M.G.A., 2018. Keystone taxa as drivers of microbiome structure and functioning. Nature Reviews Microbiology 16, 567–576. https://doi.org/10.1038/s41579-018-0024-1.

Barberán, A., Bates, S.T., Casamayor, E.O., Fierer, N., 2012. Using network analysis to explore co-occurrence patterns in soil microbial communities. The ISME Journal 6, 343–351. https://doi.org/10.1038/ismej.2011.119.

Barner, A.K., Coblentz, K.E., Hacker, S.D., Menge, B.A., 2018. Fundamental contradictions among observational and experimental estimates of non-trophic species interactions. Ecology 99, 557–566. https://doi.org/10.1002/ecy.2133.

Barroso-Bergadà, D., Pauvert, C., Vallance, J., Delière, L., Bohan, D.A., Buée, M., Vacher, C., 2021. Microbial networks inferred from environmental DNA data for biomonitoring ecosystem change: strengths and pitfalls. Molecular Ecology Resources 21, 762–780. https://doi.org/10.1111/1755-0998.13302.

Berry, D., Widder, S., 2014. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. Frontiers in Microbiology 5, 1–14. https://doi.org/10.3389/fmicb.2014.00219.

Blanchet, F.G., Cazelles, K., Gravel, D., 2020. Co-occurrence is not evidence of ecological interactions. Ecology Letters 23, 1050–1063. https://doi.org/10.1111/ele.13525.

Blois, J.L., Gotelli, N.J., Behrensmeyer, A.K., Faith, J.T., Lyons, S.K., Williams, J.W., Amatangelo, K.L., Bercovici, A., Du, A., Eronen, J.T., Graves, G.R., Jud, N., Labandeira, C., Looy, C.V., Mcgill, B., Patterson, D., Potts, R., Riddle, B., Terry, R., Tóth, A., Villaseñor, A., Wing, S., 2014. A framework for evaluating the influence of climate, dispersal limitation, and biotic interactions using fossil pollen associations across the late Quaternary. Ecography 37, 1095–1108. https://doi.org/10.1111/ecog.00779.

Blüthgen, N., 2010. Why network analysis is often disconnected from community ecology: a critique and an ecologist's guide. Basic and Applied Ecology 11, 185–195. https://doi.org/10.1016/j.baae.2010.01.001.

Boutsis, G., Stamou, G.P., Argyropoulou, M.D., 2011. Short term effects of soil disinfection with metham sodium and organic alternatives on nematode communities. Community Ecology 12, 161–170. https://doi.org/10.1556/ComEc.12.2011.2.3.

Brisson, V., Schmidt, J., Northen, T.R., Vogel, J.P., Gaudin, A., 2019. A new method to correct for habitat filtering in microbial correlation networks. Frontiers in Microbiology 10, 1–10. https://doi.org/10.3389/fmicb.2019.00585.

Cardinale, M., Grube, M., Erlacher, A., Quehenberger, J., Berg, G., 2015. Bacterial networks and co-occurrence relationships in the lettuce root microbiota. Environmental Microbiology 17, 239–252. https://doi.org/10.1111/1462-2920.12686.

Carr, A., Diener, C., Baliga, N.S., Gibbons, S.M., 2019. Use and abuse of correlation analyses in microbial ecology. The ISME Journal 13, 2647–2655. https://doi.org/10.1038/s41396-019-0459-z.

Chan, J.Z.M., Halachev, M.R., Loman, N.J., Constantinidou, C., Pallen, M.J., 2012. Defining bacterial species in the genomic era: insights from the genus Acinetobacter. BMC Microbiology 12. https://doi.org/10.1186/1471-2180-12-302.

Chen, Q., An, X., Li, H., Su, J., Ma, Y., Zhu, Y.-G., 2016. Long-term field application of sewage sludge increases the abundance of antibiotic resistance genes in soil. Environment International 92 (93), 1–10. https://doi.org/10.1016/j.envint.2016.03.026.

Co, A.D., Vliet, S. Van, Kiviet, D.J., Schlegel, S., Ackermann, M., 2020. Short-range interactions govern the dynamics and functions of microbial communities. Nature Ecology & Evolution 4, 366–375. https://doi.org/10.1038/s41559-019-1080-2.

Cogoul, A., Bailly, X., Vourc'h, G., Gasqui, P., 2019. Rarity of microbial species : in search of reliable associations. PLoS One 14, e0200458. https://doi.org/10.1371/journal.pone.0200458.

Connor, E.F., Simberloff, D., 1979. The assembly of species communities: chance or competition? Ecology 60, 1132–1140.

Connor, N., Barberan, A., Clauset, A., 2017. Using null models to infer microbial co-occurrence networks. PLoS One 12. https://doi.org/10.1371/journal.pone.0176751.

Creamer, R.E., Hannula, S.E., Van Leeuwen, J.P., Stone, D., Rutgers, M., Schmelz, R.M., de Ruiter, P.C., Hendriksen, N.B., Bolger, T., Bouffaud, M.L., Buee, M., Carvalho, F., Costa, D., Dirilgen, T., Francisco, R., Griffiths, B.S., Griffiths, R., Martin, F., da Silva, P., Mendes, S., Morais, P.V., Pereira, C., Philippot, L., Plassart, P., Redecker, D., Roembke, J., Sousa, J.P., Wouterse, M., Lemanceau, P., 2016. Ecological network analysis reveals the inter-connection between soil biodiversity and ecosystem function as affected by land use across Europe. Applied Soil Ecology 97, 112–124. https://doi.org/10.1016/j.apsoil.2015.08.006.

Dallas, T., Melbourne, B.A., 2019. When can competition and dispersal lead to checkerboard distributions? Journal of Animal Ecology 88, 269–276. https://doi.org/10.1111/1365-2656.12913.

de Vries, F.T., Griffiths, R.I., Bailey, M., Craig, H., Girlanda, M., Gweon, H.S., Hallin, S., Kaisermann, A., Keith, A.M., Kretzschmar, M., Lemanceau, P., Lumini, E., Mason, K. E., Oliver, A., Ostle, N., Prosser, J.I., Thion, C., Thomson, B., Bardgett, R.D., 2018. Soil bacterial networks are less stable under drought than fungal networks. Nature Communications 9, 3033. https://doi.org/10.1038/s41467-018-05516-7.

Delgado-Baquerizo, M., Oliverio, A.M., Brewer, T.E., Benavent-gonzález, A., Eldridge, D. J., Bardgett, R.D., Maestre, F.T., Singh, B.K., Fierer, N., 2018. A global atlas of the dominant bacteria found in soil. Science 325, 320–325.

Diamond, J.M., 1975. Assembly of species communities. In: Cody, M.L., Diamond, J.M. (Eds.), Ecology and Evolution of Communities. Harvard University Press, Cambridge, USA, pp. 342–444.

Dini-Andreote, F., Michele, de C.P. e S., Triado-Margarit, X., Casamayor, E.O., van Elsas, J.D., Salles, J.F., 2014. Dynamics of bacterial community succession in a salt marsh chronosequence: evidences for temporal niche partitioning. The ISME Journal 8, 1989–2001. https://doi.org/10.1038/ismej.2014.54.

Dormann, C.F., Frund, J., Bluthgen, N., Gruber, B., 2009. Indices, graphs and null models: analyzing bipartite ecological networks. The Open Ecology Journal 2, 7–24. https://doi.org/10.2174/1874213000902010007.

Duarte Ritter, C., Forster, D., Azevedo, J.A.R., Antonelli, A., Nilsson, R.H., Trujillo, M.E., Dunthorn, M., 2021. Assessing biotic and abiotic interactions of microorganisms in amazonia through Co-occurrence networks and DNA metabarcoding. Microbial Ecology. https://doi.org/10.1007/s00248-021-01719-6.

Erktan, A., Or, D., Scheu, S., 2020. The physical structure of soil : determinant and consequence of trophic interactions. Soil Biology and Biochemistry 148, 107876. https://doi.org/10.1016/j.soilbio.2020.107876.

Faust, K., 2021. Open challenges for microbial network construction and analysis. The ISME Journal. https://doi.org/10.1038/s41396-021-01027-4.

Faust, K., Lima-Mendez, G., Lerat, J.-S., Sathirapongsasuti, J.F., Knight, R., Huttenhower, C., Lenaerts, T., Raes, J., 2015. Cross-biome comparison of microbial association networks. Frontiers in Microbiology 6. https://doi.org/10.3389/fmicb.2015.01200.

Faust, K., Raes, J., 2012. Microbial interactions: from networks to models. Nature Reviews Microbiology 10, 538–550. https://doi.org/10.1038/nrmicro2832.

Freilich, M.A., Wieters, E., Broitman, B.R., Marquet, P.A., Navarrete, S.A., 2018. Species co-occurrence networks: can they reveal trophic and non-trophic interactions in ecological communities? Ecology 99, 690–699. https://doi.org/10.1002/ecy.2142.

Gloor, G.B., Macklaim, J.M., Pawlowsky-Glahn, V., Egozcue, J.J., 2017. Microbiome datasets are compositional: and this is not optional. Frontiers in Microbiology 8, 1–6. https://doi.org/10.3389/fmicb.2017.02224.

Goberna, M., Donat, S., Pérez-Valera, E., Hallin, S., Verdú, M., 2021. *Nir* gene-based co-occurrence patterns reveal assembly mechanisms of soil denitrifiers in response to fire. Environmental Microbiology 23, 239–251. https://doi.org/10.1111/1462-2920.15298.

Goberna, M., Montesinos-Navarro, A., Valiente-Banuet, A., Colin, Y., Gómez-Fernández, A., Donat, S., Navarro-Cano, J.A., Verdú, M., 2019. Incorporating phylogenetic metrics to microbial co-occurrence networks based on amplicon sequences to discern community assembly processes. Molecular Ecology Resources 19, 1552–1564. https://doi.org/10.1111/1755-0998.13079.

Goberna, M., Navarro-Cano, J.A., Valiente-Banuet, A., García, C., Verdú, M., 2014. Abiotic stress tolerance and competition-related traits underlie phylogenetic clustering in soil bacterial communities. Ecology Letters 17, 1191–1201. https://doi.org/10.1111/ele.12341.

Gotelli, N.J., 2000. Null model analysis of species co-occurrence patterns. Ecology 81, 2606–2621. https://doi.org/10.1890/0012-9658(2000)081[2606:NMAOSC]2.0.CO;2.

Hartman, K., van der Heijden, M.G.A., Wittwer, R.A., Banerjee, S., Walser, J.-C., Schlaeppi, K., 2018. Cropping practices manipulate abundance patterns of root and soil microbiome members paving the way to smart farming. Microbiome 6, 14. https://doi.org/10.1186/s40168-017-0389-9.

Hemprich-Bennett, D.R., Oliveira, H.F.M., Le Comber, S.C., Rossiter, S.J., Clare, E.L., 2021. Assessing the impact of taxon resolution on network structure. Ecology 102. https://doi.org/10.1002/ecy.3256.

Hibbing, M.E., Fuqua, C., Parsek, M.R., Peterson, S.B., 2010. Bacterial competition: surviving and thriving in the microbial jungle. Nature Reviews Microbiology 8, 15–25. https://doi.org/10.1038/nrmicro2259.Bacterial.

Horner-Devine, M.C., Silver, J.M., Leibold, M.A., Bohannan, B.J.M., Colwell, R.K., Fuhrman, J.A., Green, J.L., Kuske, C.R., Martiny, J.B.H., Muyzer, G., Øvreås, L., Reysenbach, A.L., Smith, V.H., 2007. A comparison of taxon co-occurrence patterns for macro- and microorganisms. Ecology 88, 1345–1353. https://doi.org/10.1890/06-0286.

Hu, H.-W., Wang, J.-T., Li, J., Li, J.-J., Ma, Y.-B., Chen, D., He, J.-Z., 2016. Field-based evidence for copper contamination induced changes of antibiotic resistance in agricultural soils. Environmental Microbiology 18, 3896–3909. https://doi.org/10.1111/1462-2920.13370.

Ings, T.C., Hawes, J., 2018. The history of ecological networks. In: Dáttilo, W., Rico-Gray, V. (Eds.), Ecological Networks in the Tropics. Springer, pp. 15–28.

Jiao, S., Liu, Z., Lin, Y., Yang, J., Chen, W., Wei, G., 2016. Bacterial communities in oil contaminated soils: biogeography and co-occurrence patterns. Soil Biology and Biochemistry 98, 64–73. https://doi.org/10.1016/j.soilbio.2016.04.005.

Jones, C.M., Hallin, S., 2019. Geospatial variation in co-occurrence networks of nitrifying microbial guilds. Molecular Ecology 28, 293–306. https://doi.org/10.1111/mec.14893.

Kanehisa, M., Goto, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Research 28, 17–30. https://doi.org/10.1093/nar/28.1.27.

Kaul, A., Mandal, S., Davidov, O., Peddada, S.D., 2017. Analysis of microbiome data in the presence of excess zeros. Frontiers in Microbiology 8, 1–10. https://doi.org/10.3389/fmicb.2017.02114.

Kaupper, T., Mendes, L.W., Lee, H.J., Mo, Y., Poehlein, A., Jia, Z., Horn, M.A., Ho, A., 2021. When the going gets tough: emergence of a complex methane-driven interaction network during recovery from desiccation-rewetting. Soil Biology and Biochemistry 153, 108109. https://doi.org/10.1016/j.soilbio.2020.108109.

Kurtz, Z.D., Müller, C.L., Miraldi, E.R., Littman, D.R., Blaser, M.J., Bonneau, R.A., 2015. Sparse and compositionally robust inference of microbial ecological networks. PLoS Computational Biology 11, 1–25. https://doi.org/10.1371/journal.pcbi.1004226.

Layeghifard, M., Hwang, D.M., Guttman, D.S., 2017. Disentangling interactions in the microbiome: a network perspective. Trends in Microbiology 25, 217–228. https://doi.org/10.1016/j.tim.2016.11.008.

Leite, F.A., Kuramae, E.E., 2020. You must choose , but choose wisely : model-based approaches for microbial community analysis. Soil Biology and Biochemistry 151, 108042. https://doi.org/10.1016/j.soilbio.2020.108042.

Li, C., Lim, K.M.K., Chng, K.R., Nagarajan, N., 2016. Predicting microbial interactions through computational approaches. Methods 102, 12–19. https://doi.org/10.1016/j.ymeth.2016.02.019.

Li, F., Chen, L., Zhang, J., Yin, J., Huang, S., 2017. Bacterial community structure after long-term organic and inorganic fertilization reveals important associations between soil nutrients and specific taxa involved in nutrient transformations. Frontiers in Microbiology 8. https://doi.org/10.3389/fmicb.2017.00187.

Libralato, S., Christensen, V., Pauly, D., 2006. A method for identifying keystone species in food web models. Ecological Modelling 195, 153–171.

Liu, P., Conrad, R., 2017. Syntrophobacteraceae-affiliated species are major propionate-degrading sulfate reducers in paddy soil. Environmental Microbiology 19, 1669–1686. https://doi.org/10.1111/1462-2920.13698.

López-Mondéjar, R., Brabcová, V., Štursová, M., Davidová, A., Jansa, J., Cajthaml, T., Baldrian, P., 2018. Decomposer food web in a deciduous forest shows high share of generalist microorganisms and importance of microbial biomass recycling. The ISME Journal 12, 1768–1778. https://doi.org/10.1038/s41396-018-0084-2.

Ma, B., Wang, H., Dsouza, M., Lou, J., He, Y., Dai, Z., Brookes, P.C., Xu, J., Gilbert, J.A., 2016. Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. The ISME Journal 10, 1891–1901. https://doi.org/10.1038/ismej.2015.261.

Martín González, A.M., Dalsgaard, B., Olesen, J.M., 2010. Centrality measures and the importance of generalist species in pollination networks. Ecological Complexity 7, 36–43. https://doi.org/10.1016/j.ecocom.2009.03.008.

Mauri, M., Elli, T., Caviglia, G., Uboldi, G., Azzi, M., 2017. RAWGraphs: a visualisation platform to create open outputs. In: Proceedings of the 12th Biannual Conference on Italian SIGCHI. ACM, New York, p. 28. https://doi.org/10.1145/3125571.3125585, 1–28:5.

Montoya, D., Yallop, M.L., Memmott, J., 2015. Functional group diversity increases with modularity in complex food webs. Nature Communications 6, 1–9. https://doi.org/10.1038/ncomms8379.

Morriën, E., 2016. Understanding soil food web dynamics, how close do we get? Soil Biology and Biochemistry 102, 10–13. https://doi.org/10.1016/j.soilbio.2016.06.022.

Münkemüller, T., Gallien, L., Pollock, L.J., Barros, C., Carboni, M., Chalmandrier, L., Mazel, F., Mokany, K., Roquet, C., Smyčka, J., Talluto, M.V., Thuiller, W., 2020. Dos and don'ts when inferring assembly rules from diversity patterns. Global Ecology and Biogeography 29, 1212–1229. https://doi.org/10.1111/geb.13098.

Newman, M.E.J., 2006. Modularity and community structure in networks. Proceedings of the National Academy of Sciences of the United States of America 103, 8577–8582. https://doi.org/10.1073/pnas.0601602103.

Nielsen, S., Minchin, T., Kimber, S., van Zwieten, L., Gilbert, J., Munroe, P., Joseph, S., Thomas, T., 2014. Comparative analysis of the microbial communities in agricultural soil amended with enhanced biochars or traditional fertilisers. Agriculture, Ecosystems & Environment 191, 73–82. https://doi.org/10.1016/j.agee.2014.04.006.

O'Malley, M.A., 2008. "Everything is everywhere: but the environment selects": ubiquitous distribution and ecological determinism in microbial biogeography. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences 39, 314–325. https://doi.org/10.1016/j.shpsc.2008.06.005.

Paine, R.T., 1995. A conversation on refining the concept of keystone species. Conservation Biology 9, 962–964.

Pérez-Valera, E., Goberna, M., Faust, K., Raes, J., García, C., Verdú, M., 2017. Fire modifies the phylogenetic structure of soil bacterial co-occurrence networks. Environmental Microbiology 19, 317–327. https://doi.org/10.1111/1462-2920.13609.

Pichler, M., Boreux, V., Klein, A.M., Schleuning, M., Hartig, F., 2020. Machine learning algorithms to infer trait-matching and predict species interactions in ecological networks. Methods in Ecology and Evolution 11, 281–293. https://doi.org/10.1111/2041-210X.13329.

Prosser, J.I., Bohannan, B.J.M., Curtis, T.P., Ellis, R.J., Firestone, M.K., Freckleton, R.P., Green, J.L., Green, L.E., Killham, K., Lennon, J.J., Osborn, A.M., Solan, M., Gast, C.J. Van Der, Young, J.P.W., 2007. The role of ecological theory in microbial ecology. Nature Reviews Microbiology 5, 384–392.

Proulx, S.R., Promislow, D.E.L., Phillips, P.C., 2005. Network thinking in ecology and evolution. Trends in Ecology & Evolution 20, 345–353. https://doi.org/10.1016/j.tree.2005.04.004.

Purahong, W., Wubet, T., Lentendu, G., Schloter, M., Pecyna, M.J., Kapturska, D., Hofrichter, M., Krüger, D., Buscot, F., 2016. Life in leaf litter: novel insights into community dynamics of bacteria and fungi during litter decomposition. Molecular Ecology 25, 4059–4074. https://doi.org/10.1111/mec.13739.

Rafferty, N.E., Ives, A.R., 2013. Phylogenetic trait-based analyses of ecological networks. Ecology 94, 2321–2333. https://doi.org/10.1890/12-1948.1.

Röttjers, L., Faust, K., 2018. From hairballs to hypotheses–biological insights from microbial networks. FEMS Microbiology Reviews 42, 761–780. https://doi.org/10.1093/femsre/fuy030.

Röttjers, L., Vandeputte, D., Raes, J., Faust, K., 2021. Null-model-based network comparison reveals core associations. ISME Communications 1–8. https://doi.org/10.1038/s43705-021-00036-w.

Strydom, T., Catchen, M.D., Banville, F., Caron, D., 2021. A Roadmap toward Predicting Species Interaction Networks (Across Space and Time) EcoEvoRxiv, pp. 1–34. https://doi.org/10.32942/osf.io/eu7k3.

Szoboszlay, M., Tebbe, C.C., 2021. Hidden heterogeneity and co-occurrence networks of soil prokaryotic communities revealed at the scale of individual soil aggregates. MicrobiologyOpen 10, 1–16. https://doi.org/10.1002/mbo3.1144.

Tackmann, J., Matias Rodrigues, J.F., von Mering, C., 2019. Rapid inference of direct interactions in large-scale ecological networks from heterogeneous microbial sequencing data. Cell Systems 9, 286–296. https://doi.org/10.1016/j.cels.2019.08.002 e8.

Tu, Q., Zhou, X., He, Z., Xue, K., Wu, L., Reich, P., Hobbie, S., Zhou, J., 2016. The diversity and Co-occurrence patterns of N-2-Fixing communities in a CO2-enriched grassland ecosystem. Microbial Ecology 71, 604–615. https://doi.org/10.1007/s00248-015-0659-7.

Ulrich, W., Jabot, F., Gotelli, N.J., 2017. Competitive interactions change the pattern of species co-occurrences under neutral dispersal. Oikos 126, 91–100. https://doi.org/10.1111/oik.03392.

Vázquez, D.P., Melián, C.J., Williams, N.M., Blüthgen, N., Krasnov, B.R., Poulin, R., 2007. Species abundance and asymmetric interaction strength in ecological networks. Oikos 116, 1120–1127. https://doi.org/10.1111/j.0030-1299.2007.15828.x.

Weiher, E., Keddy, P.A., 1995. Assembly rules , null models , and trait dispersion : new questions from old patterns. Oikos 74, 159–164.

Weiss, S., Van Treuren, W., Lozupone, C., Faust, K., Friedman, J., Deng, Y., Xia, L.C., Xu, Z.Z., Ursell, L., Alm, E.J., Birmingham, A., Cram, J.A., Fuhrman, J.A., Raes, J., Sun, F., Zhou, J., Knight, R., 2016. Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. The ISME Journal 10, 1669–1681. https://doi.org/10.1038/ismej.2015.235.

Weiss, S., Xu, Z.Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J.R., Vázquez-Baeza, Y., Birmingham, A., Hyde, E.R., Knight, R., 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome 5, 1–18. https://doi.org/10.1186/s40168-017-0237-y.

Williams, R.J., Howe, A., Hofmockel, K.S., 2014. Demonstrating microbial co-occurrence pattern analyses within and between ecosystems. Frontiers in Microbiology 5, 1–10. https://doi.org/10.3389/fmicb.2014.00358.

Xun, W., Liu, Y., Li, W., Ren, Y., Xiong, W., Xu, Z., Zhang, N., Miao, Y., Shen, Q., Zhang, R., 2021. Specialized metabolic functions of keystone taxa sustain soil microbiome stability. Microbiome 9, 35. https://doi.org/10.1186/s40168-020-00985-9.

Yeates, G.W., Bongers, T., De Goede, R.G.M., Freckman, D.W., Georgieva, S.S., 1993. Feeding habits in soil nematode families and genera-an outline for soil ecologists. Journal of Nematology 25, 315–331.

Zhou, J., Deng, Y., Luo, F., He, Z., Yanga, Y., 2011. Phylogenetic molecular ecological network of soil microbial communities in response to elevated CO2. mBio 2, 1–8. https://doi.org/10.1128/mBio.00122-11.