available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/funeco

Commentary

Plant–fungal symbioses as ecological networks: the need to characterize more than just interaction patterns

ABSTRACT

Next-generation sequencing technologies are providing us with new opportunities to characterize plant–fungal communities in more depth and with better replication than ever before. The application of network concepts and numerical tools to analyze those extensive data sets is also rapidly increasing. Here we show, however, that network-based tools will further advance our understanding of the ecology of plant–fungal symbioses if (1) researchers characterize both the interaction patterns among species, and investigate the likely biotic and abiotic drivers of such interactions (e.g. species' abundance, functional traits, environmental conditions) and (2) researchers make sure that the assumptions made by their network-based numerical tools are met by their data sets.

The increasing accessibility of next-generation sequencing technologies has sparked a new wave of studies that have characterized interaction patterns naturally occurring between plants and their fungal symbionts (e.g. Montesinos-Navarro et al., 2012; Martos et al., 2012). One way to analyze such data sets, is to describe community structure using novel indices/metrics derived from ecological network theory (Bascompte, 2009; Bahram et al., 2014). The advantage of this approach is that it is possible to detect community-level patterns and to evaluate, through the use of null models, their statistical significance and their ecological correlates (Chagnon et al., 2012). Thus, this approach has the potential to shed new light on the processes underpinning the ecological and co-evolutionary dynamics of symbiotic communities (Bascompte, 2009; Ulrich and Gotelli, 2013). However, field sampling schemes and numerical analyses need to be carefully designed in order to maximize the inference that can be extracted from data sets (Heleno et al., 2014). To emphasize this point, we have re-analyzed data describing the interactions found between plants and root-colonizing fungi in an oak-dominated temperate forest in Japan (Toju et al., 2013). This recently published data set provides useful insights on the quantitative nature of plant–fungal interactions in a natural forest setting. Our re-analysis of their data suggests, however, that by strictly characterizing interaction patterns among plant and fungal taxa, the study provides little information about the relative importance of neutral vs. niche-based processes that determine the assembly of plant-fungal communities. Given the growing importance of next-generation sequencing studies in belowground community ecology (Poisot et al., 2013), future network studies may have to invest less effort in characterizing interaction patterns among species to be able to invest more in investigating the biotic and abiotic drivers of community-level patterns.

Toju et al. (2013) sampled a 59 m × 15 m grid comprising 960 soil sampling points. At each point, they collected one root fragment from which DNA was extracted. From these extracts, they identified plant species by amplifying and sequencing chloroplastic DNA. They also used the same DNA extracts to amplify and sequence fungal DNA (using general fungal ITS primers) to determine the fungal taxa composition inside of roots. After thorough bioinformatic filtering of the data set (see Toju et al., 2013), the authors identified a network of 10 plant species interacting with 49 fungal taxa. The aim of the study was to determine the degree of specialization in plant–fungal interactions in a natural forest setting. They calculated the specificity of associations between plants and fungi by computing the d' index, which is an information-derived index (like Shannon diversity, see Blüthgen et al., 2006). This index is bounded between 0 and 1: high values indicate a low diversity of partners (i.e. specialist species). This index was explicitly stated by its developers to be useful in studies focusing on spatial scales that are small enough to avoid situations where the absence of an interaction between two species could be simply ascribed to the absence of overlap in their spatial distribution (Blüthgen et al., 2006). In other words, at large spatial scales, the d' index does not strictly address the issue of partner selection and association specificity, but it is also biased by the neutral effects of species abundances and spatial distributions. Thus, in the study by Toju et al. (2013), to assume that the d' index actually characterizes association specificity, it is necessary to demonstrate that species are homogeneously distributed across the sampling points. To verify this assumption, we plotted the spatial distribution of plant and fungal taxa across the spatial grid sampled by Toju et al. (2013). For many species, there were obvious visual patterns of aggregation (see examples in Fig 1). To test for the significance

of this aggregation, we calculated Besag's *L* function (Besag, 1977), an improved version of Ripley's *K* function, which calculates and compares the frequency with which events occur at small pairwise distances with those predicted from Monte Carlo random simulations. For spatial scales between 1 and10 m, we found that most plant species and about half of the fungal taxa were significantly aggregated (Fig 2). This confirms that species cannot be assumed to be homogeneously distributed across the landscape, and that the *d'* index cannot be interpreted here to strictly infer preferential partner selection. For example, partners that were found to interact more often than predicted by chance may simply have had overlapping spatial distributions arising from stochastic dispersion processes or from their similar responses to environmental gradients (e.g. soil properties).

To test whether spatial co-occurrence patterns could predict the interactions observed between plants and fungi, we constructed a null model that allocated interactions in the network based on the co-occurrence patterns of plant and fungal taxa. First, we calculated a pairwise co-occurrence index under the form of a z-score. Briefly, for each plant–fungal pair, we compared the total number of co-occurrences observed in the field to a null distribution that was obtained by shuffling the spatial distribution of the fungus. We thus ended up calculating: $z_c = (C_{obs} - \overline{C_{null}})/SD_{C_{null}}$

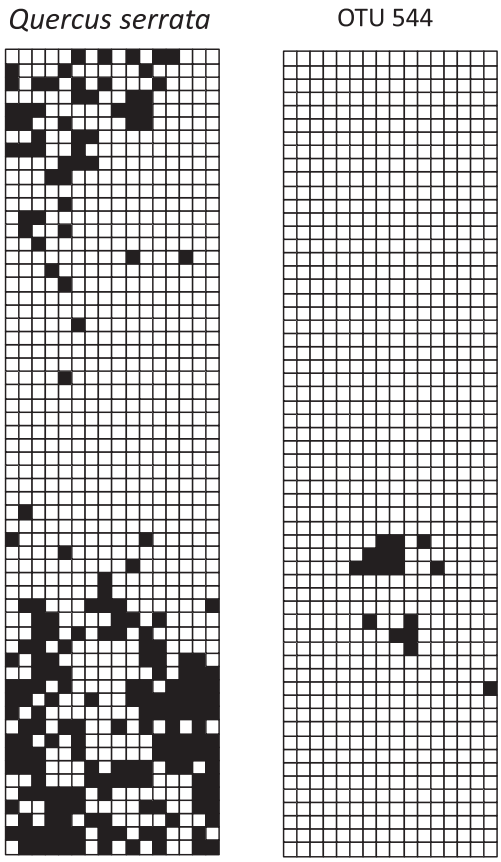

Fig 1 – Examples of species that had clearly and significantly aggregated spatial distributions (i.e. the plant species *Quercus serrata*, and the fungal OTU 544). The matrices represent species' occurrences across each spatial sampling units (i.e. cells) in a binary way (occurrences = filled cells).

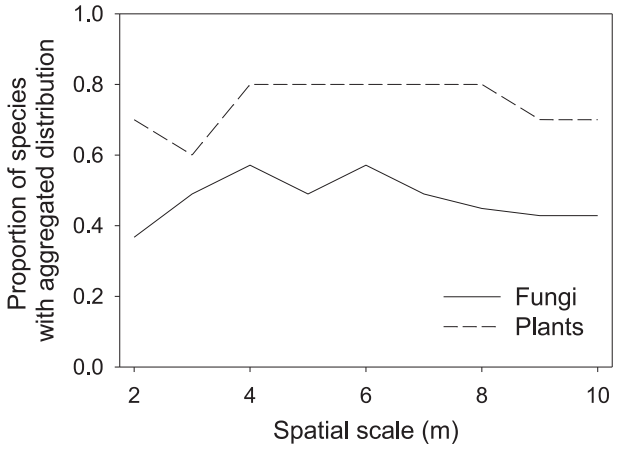

Fig 2 – Proportion of plant and fungal species in the data set having a significantly aggregated spatial distribution (test using Besag's *L* function and Monte Carlo randomizations) at spatial scales ranging from 1 to 10 m.

where z_c is the co-occurrence index, C_{obs} is the total number of co-occurrences between the plant and fungal taxa in the field, $\overline{C_{null}}$ is the mean number of co-occurrences from 1 000 simulations, and $SD_{C_{null}}$ is the standard deviation around $\overline{C_{null}}$ (z-scores shown in SI.2, Table S1). We then built 1 000 random networks, allocating interactions using the z-scores calculated above as probabilities (i.e. high z-score = higher probability of interacting in simulated networks). It should be noted that the z-scores can be negative if there are less co-occurrences than expected by chance between two species. To allow using them as probabilities in our simulations, we transformed the values by bounding them between 0 and 1, using the function *decostand* as implemented in the *R* package *vegan* (Oksanen et al., 2013). While assembling our random networks, we constrained interaction probabilities according to two important network attributes: (1) the total number of interactions (i.e. connectance) in the network, and (2) the total number of interactions per fungal taxon (*R* codes and results provided in SI.1 and SI.2 Table S2, respectively). Controlling for connectance is a routine procedure when simulating random interaction networks, because connectance is highly correlated to many network metrics (e.g. Almeida-Neto et al., 2008; Blüthgen et al., 2008). Controlling for the total number of interactions per fungal taxon was also important because the large number of taxa with very few interactions in the original data set would have artificially inflated the number of empty columns in the simulated random networks. Interactions that were found in more than half (i.e. >500) of our null network simulations were then assumed to be predictable by spatial co-occurrence patterns. As a result, we found that 257 of the 274 interactions present in the original data set (i.e. 94 %) could be predicted by spatial co-occurrence patterns (SI.2, Table S3).

The close relationship between spatial distributions and the observed interaction patterns was not surprising, given the nature of the data set: at each sampling point, fungal DNA was sequenced from roots of a single plant species. Thus, if a fungus co-occurred at a given sampling point with a given plant species, it was necessarily because it was found interacting with that plant (i.e. sequenced from its roots). In other