ECOLOGY LETTERS

*Ecology Letters*, (2016) **19**: 926–936 doi: 10.1111/ele.12630

LETTER

# The interconnected rhizosphere: High network complexity dominates rhizosphere assemblages

Shengjing Shi,[1,2†] Erin E. Nuccio,[1,3†] Zhou J. Shi,[2] Zhili He,[2] Jizhong Zhou[2,4,5] and Mary K. Firestone[1,4*]

**Abstract**

While interactions between roots and microorganisms have been intensively studied, we know little about interactions among root-associated microbes. We used random matrix theory-based network analysis of 16S rRNA genes to identify bacterial networks associated with wild oat (*Avena fatua*) over two seasons in greenhouse microcosms. Rhizosphere networks were substantially more complex than those in surrounding soils, indicating the rhizosphere has a greater potential for interactions and niche-sharing. Network complexity increased as plants grew, even as diversity decreased, highlighting that community organisation is not captured by univariate diversity. Covariations were predominantly positive (> 80%), suggesting that extensive mutualistic interactions may occur among rhizosphere bacteria; we identified quorum-based signalling as one potential strategy. Putative keystone taxa often had low relative abundances, suggesting low-abundance taxa may significantly contribute to rhizosphere function. Network complexity, a previously undescribed property of the rhizosphere microbiome, appears to be a defining characteristic of this habitat.

**Keywords**

Community ecology, keystone species, microbial ecology, microbial interactions, microbial networks, quorum sensing, random matrix theory, rhizosphere.

*Ecology Letters* (2016) **19**: 926–936

## INTRODUCTION

In many environments, microorganisms coexist in complex arrays in which interactions among members are essential for community assembly and ecosystem function (Fuhrman 2009; Hallam & McCutcheon 2015). In the zone immediately surrounding roots, known as the rhizosphere, plants supply carbon (C) to soil making this millimetre-sized habitat the locus of intense microbial activities and interactions (Philippot *et al*. 2013). Reflecting the ecological importance of the rhizosphere microbiome to nutrient cycling and availability to plants, the literature is rich with studies that have investigated the composition of rhizosphere bacterial assemblages (see reviews by Hinsinger *et al*. (2009) and Philippot *et al*. (2013)). However, few of these studies explore the interactions among members of rhizosphere assemblages, or determine which members share niches within the rhizosphere environment.

Identifying and defining the interactions that occur among soil microorganisms is critical to understanding microbial diversity and function (Hallam & McCutcheon 2015; Ren *et al*. 2015). Network analysis provides a promising start for exploring the organisation and dynamics of microbial interactions and niches (Duran-Pinedo *et al*. 2011; Zhou *et al*. 2011; Faust & Raes 2012). Studies of macrobiological ecological networks have improved our understanding of ecosystem dynamics and species co-evolution (e.g. animal food webs, plant-animal networks, plant-mycorrhizal networks) (Proulx *et al*. 2005; Thompson 2005; Konopka 2009). As microbial community ecology matures, knowledge and ecological theory from macroecology can be extremely useful in providing hypotheses for further testing (Prosser *et al*. 2007). In recent years, microbial network analysis has been used as a tool to explore complex microbial assemblages in environments such as humans (Duran-Pinedo *et al*. 2011; Faust *et al*. 2012), oceans (Steele *et al*. 2011), groundwater (Deng *et al*. 2012, 2016) and soil (Zhou *et al*. 2010, 2011; Barberan *et al*. 2012; Lu *et al*. 2013). These network studies provide perspectives on microbial assemblages beyond those of simple richness and composition, and add a substantial dimension to our understanding of microbial community ecology.

Network analyses often reveal non-random co-variation patterns which may reflect community organisation – such as direct interactions (Faust & Raes 2012) or shared guilds or niches (Berry & Widder 2014) – and provide a tool for investigating ecological concepts which are difficult to assess in microbial communities. A study by Duran-Pinedo *et al*. (2011) provides an example of the value of network analysis for identifying metabolic cooperation in microbial systems, where the consistent network co-occurrence of *Tannerella* sp. OT286 and *Prevotella oris* OT311 was further investigated in

[1]*Department of Environmental Science, Policy and Management, University of California, Berkeley, CA 94720, USA*
[2]*Department of Botany and Microbiology, Institute for Environmental Genomics, University of Oklahoma, Norman, OK 73019, USA*
[3]*Nuclear and Chemical Sciences Division, Lawrence Livermore National Laboratory, Livermore, CA 94551, USA*
[4]*Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[5]*State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China*
**Correspondence: E-mail: mkfstone@berkeley.edu*
*† Equal contribution*

co-culture and resulted in the cultivation of the previously uncultivated *Tannerrella*. Highly connected modular structures within networks are thought to represent important ecological units, and have been conceptualised as compartments, guilds, and/or cohesive subgroups (Newman 2006; Olesen *et al*. 2007; Dupont & Olesen 2009). Previous studies have proposed that modules reflect habitat heterogeneity, divergent selection regimes, clusters of phylogenetically closely related species and even the key unit of species co-evolution (Thompson 2005; Olesen *et al*. 2007). In addition, network analysis can identify putative keystone taxa which are critical in maintaining community structure and function (Power *et al*. 1996). As there are currently no other tractable means of identifying keystone microbial taxa in diverse and largely uncultivated soil microbial communities (Zhou *et al*. 2011), network analysis fills a crucial need in microbial community ecology (Berry & Widder 2014).

To identify bacterial assemblages that potentially interact or share niches within rhizosphere soil, we used random matrix theory (RMT) (Deng *et al*. 2012) to construct co-occurrence networks for rhizosphere and bulk soil assemblages throughout the lifespan of *Avena fatua*, a common Mediterranean annual grass. To reduce covariations due to external environmental variability and maximise covariations due to interactions (Berry & Widder 2014), we conducted a highly replicated plant microcosm study with homogenised soil to minimise variability between replicates. Plants were grown in a greenhouse in soil to which *Avena* spp. had naturalised for many decades. Rhizosphere and bulk soils were collected ($n = 16$) at 10 times spanning two seasons of plant growth. High-throughput sequencing of 16S rRNA gene amplicons was used to describe the bacterial assemblages. Bacterial network analysis was used to address four questions: (1) Are rhizosphere bacterial networks significantly different from bulk soil networks in terms of network size and complexity? (2) Do rhizosphere networks change as the bacterial assemblages undergo succession over the lifespan of a plant or between seasons? (3) Are there taxa that play particularly important roles within rhizosphere networks, suggesting they may serve as keystone taxa in rhizosphere communities? (4) Can highly connected groups of nodes (modules) be identified that result from specific interactions? Our work identifies a previously undocumented dimension of the rhizosphere, and offers insight into fundamental properties of these soil habitats.

## MATERIALS AND METHODS

### Experiment description

Rhizosphere and non-rhizosphere (bulk) soil samples were collected from a greenhouse experiment as described in detail in Shi *et al*. (2015). Briefly, *Avena fatua* seedlings were planted in microcosms (one plant per microcosm) and placed in growth chambers at the Environmental Plant Isotope Chamber facility, University of California, Berkeley. Soil was collected beneath *Avena* stands at the Hopland Research and Extension Center (Hopland, CA, USA) during the spring. Microcosms were disassembled for sampling at pre-planted (week 0), seedling (week 3), vegetative (week 6), flowering (week 9) and senescent (week 12) stages for both growing seasons (Fig. S1). After the first growing season, plant shoots from non-harvested microcosms were removed and these microcosms remained un-watered for 3 months to simulate a dry Mediterranean summer. The collection of bulk soil differed between Seasons 1 and 2. During Season 1, bulk soil was collected from root exclusion bags (1 μm mesh). During Season 2, bulk soil was collected after removing live roots with attached rhizosphere soils and contained root debris from the previous season. Only bulk soil was collected prior to planting (at week 0), and for the remaining time points paired rhizosphere (soil firmly attached to roots) and bulk soils were sampled from the same microcosms (Fig. S1). Overall, 288 samples were collected, representing rhizosphere and bulk samples from 16 replicates over 10 time points (8 rhizosphere harvests and 10 bulk soil harvests).

### Microbial community analysis by MiSeq sequencing of 16S rRNA gene amplicons

Soil microbial DNA was extracted, amplified and barcoded with primer set F515 and R806 (Caporaso *et al*. 2012), and sequenced on an MiSeq (Illumina, San Diego, CA, USA) at the Institute for Environmental Genomics, University of Oklahoma. Sequencing and bioinformatics methods were described in our previous paper (Shi *et al*. 2015) and Supplemental Information. Briefly, sequences were processed with using an in-house pipeline at the University of Oklahoma, where 288 samples were rarified to a depth of 11 914 sequences per sample (quality score $\geq 20$ and length between 251 and 256 bp without ambiguous bases).

### Network construction and analysis

Networks were constructed for rhizosphere and bulk soil communities based on OTU relative abundances at each time point, yielding a total of 18 networks. Covariations were measured across 16 biological replicates to create each network. Only OTUs detected in 10 out of 16 replicate samples were used for network construction. RMT was used to automatically identify the appropriate similarity threshold (*St*) prior to network construction; *St* defines the minimal strength of the connections between each pair of nodes (Zhou *et al*. 2010, 2011) (see Supplemental Information for details and comparison to other methods). Global network properties were characterised according to Deng *et al*. (2012). All analyses were performed using the Molecular Ecological Network Analyses (MENA) Pipeline (http://ieg2.ou.edu/MENA/) (Deng *et al*. 2012) and networks were graphed using Cytoscape 2.8.2 (Shannon *et al*. 2003) and gephi 0.8.2-beta (Bastian *et al*. 2009).

### Detection of modules and identification of node roles

We characterised network modularity for each network created in this study. A module is a group of nodes (i.e. OTUs) that are highly connected within the group with few

connections outside the group (Newman 2006). In this study, modules were detected using the greedy modularity optimisation method (Deng *et al*. 2012) (see Supplemental Information for details). Modularity ($M$) is an index measuring the extent to which a network is divided into modules, and we used $M > 0.4$ as the threshold to define modular structures (Newman 2006). The connectivity of each node was determined based on its within-module connectivity ($Zi$) and among-module connectivity ($Pi$) (Guimera & Amaral 2005), which were then used to classify the nodes based on the topological roles they play in the network. Node topologies are organised into four categories: module hubs (highly connected nodes within modules, $Zi > 2.5$), network hubs (highly connected nodes within entire network, $Zi > 2.5$ and $Pi > 0.62$), connectors (nodes that connect modules, $Pi > 0.62$) and peripherals (nodes connected in modules with few outside connections, $Zi < 2.5$ and $Pi < 0.62$) (Olesen *et al*. 2007; Zhou *et al*. 2010; Deng *et al*. 2012).

Previous work in our group had identified quorum sensing (QS) as a potential communication strategy in the *Avena fatua* rhizosphere, where QS organisms were isolated from a highly similar experimental system (DeAngelis *et al*. 2008). In addition to growing the same plant host at the same greenhouse in almost identical microcosm units, the DeAngelis study was conducted using soil collected from the same field station. To investigate the QS potential of taxa in modules, we used BLAST to identify taxa that were > 97% similar to these QS isolates (DeAngelis *et al*. 2008).

### Statistical analyses

All statistical analyses were conducted using the 'stats' package in R version 3.2.2 (Team 2015). We conducted a two-way analysis of covariance (ANCOVA) to test if the number of nodes and links significantly differed between the rhizosphere and bulk soil over time, and if these changes significantly differed between seasons (R: aov). A significant interaction term between sample type (rhizosphere vs. bulk) and time would indicate that the slopes of the rhizosphere and bulk samples were significantly different and therefore followed different trajectories over time. ANCOVA was not used to adjust the means of the factors. Model simplification was accomplished by removing the least significant terms in a stepwise manner until removing a parameter significantly altered the model (R: anova). Linear regression analyses characterised the slopes for rhizosphere and bulk soils during each season and determine if they were significantly different from zero (R: lm).

Pearson's product moment correlations were used to determine if increasing network size and connectivity were significantly correlated with decreasing diversity previously observed in this experimental system (R: cor.test) (Shi *et al*. 2015). The following univariate diversity metrics were correlated with the number of nodes or links in each network: phylogenetic diversity, Shannon's diversity (H), richness (S) and evenness (J). If data were nonlinear and monotonic (linear correlation $P$ value > 0.05 and data points only increasing or decreasing), we performed a Spearman's rank correlation with determine if the two variables were correlated (R: cor.test).

## RESULTS

### Characteristics of constructed networks

To identify potential microbe-microbe interactions and niche-sharing in rhizosphere and neighbouring bulk soil, we constructed 18 bacterial co-occurrence networks during bacterial succession over two growing seasons of *Avena fatua* (Fig. 1). Similarity threshold ($St$) values imposed ranged from 0.79 to 0.83 (Tables S1 and S2). All the networks obtained exhibited scale-free characteristics, as indicated by $R^2$ of power law ranging from 0.74 to 0.99, and were significantly different from random networks generated using identical numbers of nodes and links (Tables S1 and S2). These metrics indicate that the network structures were non-random and unlikely due to chance.

### Distinct networks in rhizosphere and bulk soils

The rhizosphere networks differed profoundly from the bulk soil networks and these differences increased over time (Fig. 1). Multiple network topological metrics consistently showed that microbial co-occurrence patterns in the rhizosphere and bulk soil were markedly different. Rhizosphere assemblages formed larger networks with more nodes than the bulk soil networks (Fig. 2a, Tables S1 and S2), and these differences were significant over time, as indicated by a significant interaction between time and sample type during two-way ANCOVA modelling ($F_{1,14} = 45.8$, $P < 0.0001$; see Table S3 for full F table). Rhizosphere networks were larger than bulk soil networks despite the fact that rhizosphere assemblages contained fewer OTUs than bulk soil assemblages (Shi *et al*. 2015). The bulk soil networks contained approximately $464 \pm 9.4$ nodes, and the network size remained consistent across all time points (Fig. 2a). In contrast, the number of nodes in rhizosphere networks increased steadily over time during both seasons (Season 1: $r^2 = 0.94$, $P = 0.0039$; Season 2: $r^2 = 0.86$, $P = 0.015$) (Fig. 2a). There was also a significant interaction between season and sample type ($F = 5.14$, $P = 0.040$), where rhizosphere networks appear larger in the second season (Fig. 2b).

Not only were rhizosphere networks larger than bulk soil networks, but the rhizosphere networks were also more connected and complex. Rhizosphere networks contained more connections (links) between nodes than bulk soil networks (Fig. 2b), which increased the density of connections in the rhizosphere and created more intricate network patterns (Fig. 1). Bulk soil networks, on the other hand, often only had one or two links, resulting in isolated networks (Fig. 1). The connectivity of the rhizosphere and bulk soil networks significantly differed over time, as indicated by the significant interaction between time and sample type during two-way ANCOVA analysis ($F_{1,16} = 53.6$, $P < 0.0001$; see Table S4 for full F table). Rhizosphere networks significantly increased in connectivity over both seasons (Season 1: $r^2 = 0.89$, $P = 0.01$; Season 2: $r^2 = 0.99$, $P = 0.00047$), while bulk soil networks again remained static over time (Fig. 2b). The increased complexity of the rhizosphere networks was reflected by the increased average degree (i.e. average links per node in the network) (Tables S1 and S2), as well as the shorter harmonic geodesic distances (HD) (Deng *et al*. 2012). The differences in

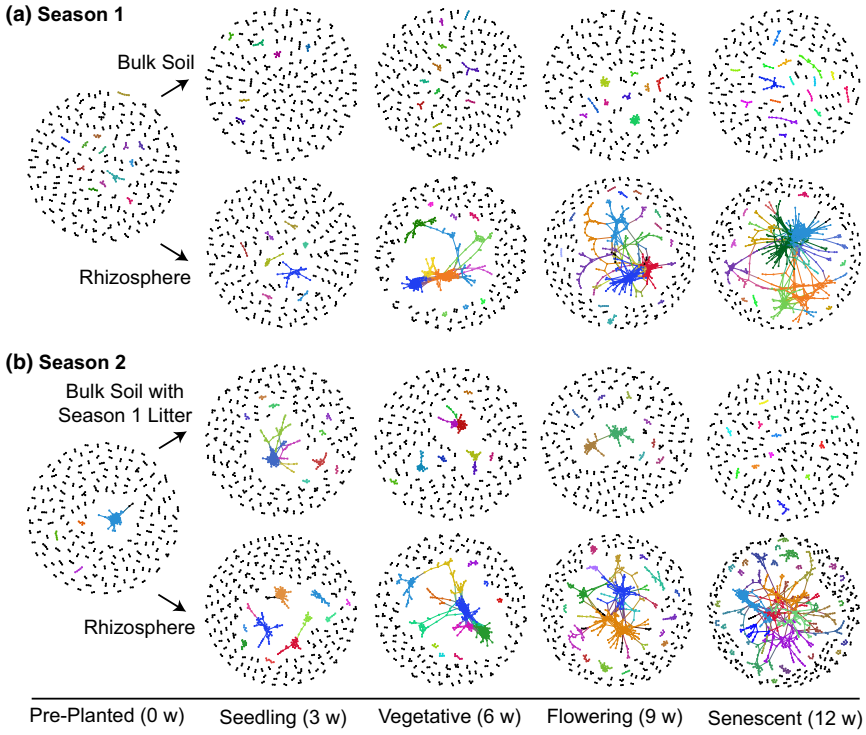


**Figure 1** Succession of rhizosphere and bulk soil networks over two seasons at different stages of vegetative growth (0, 3, 6, 9, 12 weeks). Networks represent random matrix theory co-occurrence models derived from 16 biological replicates at each time point, where nodes represent OTUs, and links between the nodes indicate significant correlation. Modules are randomly coloured at each time point, and nodes in modules with less than 5 nodes are coloured black. During the first season (a) bulk soil was harvested from root exclusion bags, while during the second season (b) bulk soil includes the root detritus from the previous season.

network connectivity between rhizosphere and bulk soil were reproducible between seasons as ANCOVA modelling could not statistically distinguish the patterns between Season 1 and 2; therefore, season was not ultimately included in the full model (Table S4). The slope of network connectivity over time for bulk soil was not statistically different from zero (Fig. 2b). Collectively the above results indicate that the rhizosphere network gradually became more complex over time as the plant grew, but the bulk soil network remained relatively static over time.

Increased network size and connectivity in the rhizosphere was accompanied by decreased bacterial diversity according to multiple univariate diversity metrics (Figs 3 and S2). Phylogenetic diversity and Shannon's diversity (H) were significantly inversely correlated with both network size and connectivity during both seasons ($r$ and $r_s < -0.88$, $P$ range: 0.0072–0.049). Richness and evenness were also significantly (or marginally significantly) inversely correlated with both network size and connectivity during both seasons ($r < -0.83$, p range: 0.0011–0.079) (Fig. S2). Rhizosphere networks became larger and more connected as the overall rhizosphere community became less rich and less even.

### Modularity in rhizosphere communities

To identify assemblages that potentially interact or share niches within rhizosphere soil, we focused on representative networks from five time points: the rhizosphere soils for weeks 6 and 12 from Seasons 1 and 2, and the pre-planted soils from season 2 (Fig. 4). We focused on modules with at least five nodes, and visualised the phylogeny for major modules with at least 10 nodes. The modules from the season 1 pre-planted soils were small (< 10 nodes) and not included in this analysis. Networks from all five time points contained modules with modularity ($M$) values > 0.73 (Table S1). Overall, taxa tended to co-occur (positive correlations, red lines) rather than co-exclude (negative correlations, blue lines); positive correlations accounted for 82–94% of the potential interactions observed at each time point (Fig. 4). Similar to the overall network structure, rhizosphere modules became larger and more connected as the plant grew, such that week 12 had the largest number of modules in each season (Fig. 4, Table S5).

The composition of modules differed within each network and changed over time (Fig. 4). During the vegetative phase in the first season (week 6), Betaproteobacteria dominated three of the large modules, and primarily co-occurred with either Acidobacteria or Actinobacteria. By 12 weeks, Betaproteobacteria had become minor components of the modules. Alphaproteobacteria were prominent members of the modules at all time points, particularly during the senescent phase of season 1 (week 12), where they co-occurred primarily with Actinobacteria and Bacteroidetes. After drying for 3 months to simulate a Mediterranean summer (Season 2, week 0), only one large module was present, in which Acidobacteria was the dominant phylum, co-occurring with Deltaproteobacteria, Betaproteobacteria and Actinobacteria. Actinobacteria were

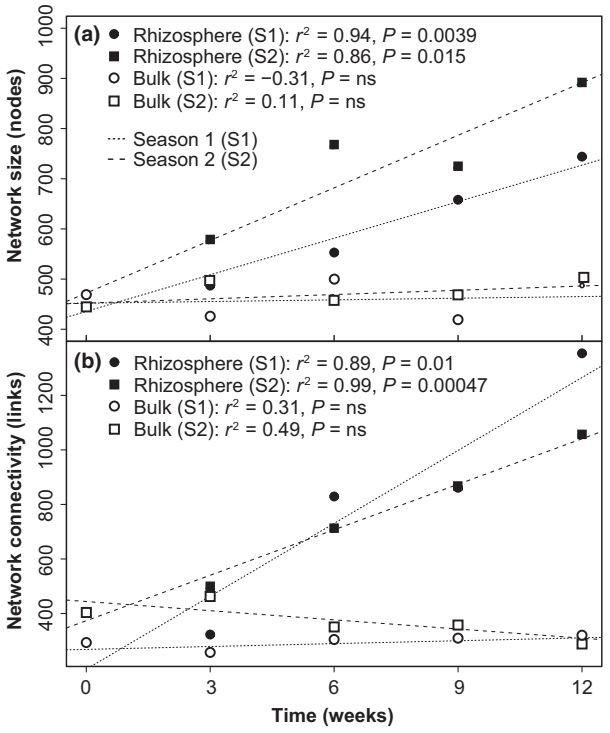


**Figure 2** Progression of (a) network size and (b) network connectivity over time for rhizosphere and bulk soils during two seasons. Lines indicate linear regression analyses for rhizosphere (solid symbols) and bulk soils (hollow symbols) for each season (S1: season 1, dotted lines, circles; S2: season 2, dashed lines, squares). "ns" indicates non-significance with a $P$ value > 0.01.

also prominent or dominant members of the modules throughout all time points. At the final sampling, Acidobacteria, Actinobacteria and Alphaproteobacteria co-occurred mostly evenly in many modules. Intriguingly, taxa from Verrucomicrobia, a largely uncharacterised phylum, are members of many modules during both the vegetative and senescent stages.

### Module hubs and connectors as putative keystone taxa

To assess possible topological roles of taxa in the networks, we classified nodes into four categories based on their within-module connectivity (*Zi*) and among-module connectivity (*Pi*) values (Deng *et al.* 2012): peripherals, connectors, module hubs and network hubs (Fig. 5, see methods for definitions). The majority of the nodes in each network were peripherals with most of their links inside their modules (Fig. 5). No module hubs or connectors were identified in the season 2 pre-planted network (S2-W0) (Fig. 4). In contrast, multiple nodes (ranging from 3 to 10 per time point) were classified as module hubs in the rhizosphere networks (Figs 3 and 4). The 24 module hubs identified originated from a variety of taxonomic groups; 10 belonged to Proteobacteria (4 from α-Proteobacteria), 5 belonged to Acidobacteria and others to Actinobacteria, Armatimonadetes, Bacteriodetes, Cyanobacteria, Gemmatimonadetes and Verrucomicrobia (Figs 3 and 4; see Table S6 for detailed taxonomy). Connectors were also detected in three of the four rhizosphere networks (no connectors detected in Season 2 week 6). Five out of seven connectors in these rhizosphere networks were Proteobacteria

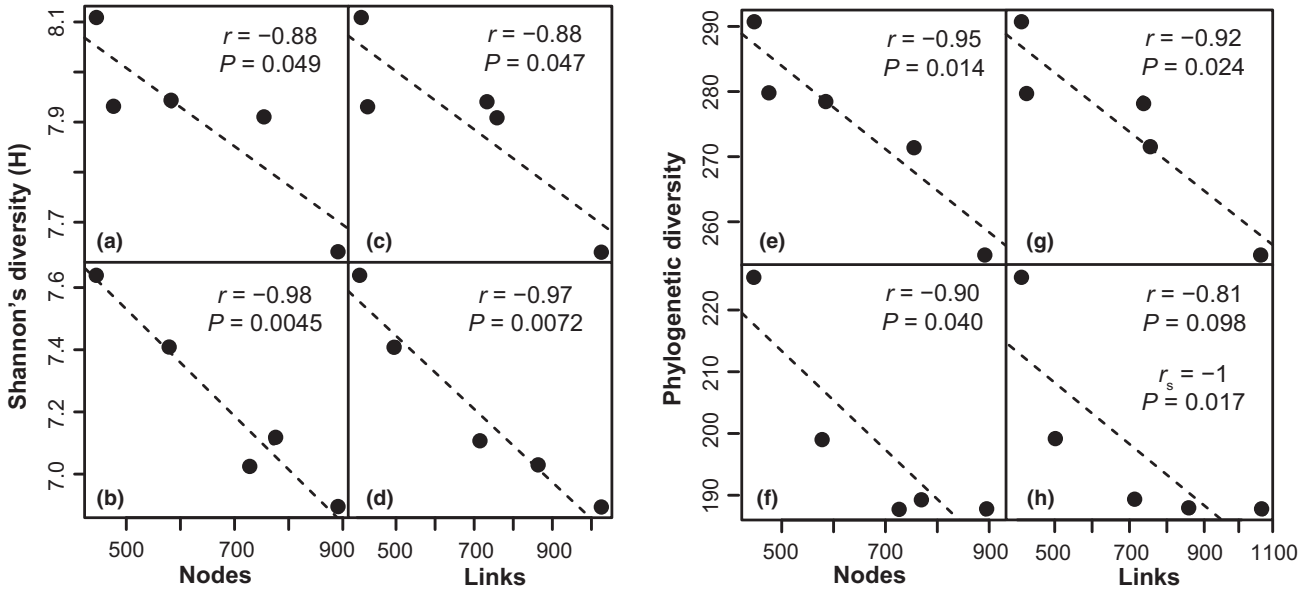


**Figure 3** Univariate diversity metrics (Shannon's diversity, phylogenetic diversity) are inversely correlated with increasing rhizosphere network complexity in terms of network size (nodes) and network connectivity (links), and is reproducible for two seasons of plant growth (Season 1: top row – a, c, e, g; Season 2: bottom row – b, d, f, h). *r*-values represent Pearson's product-moment correlation coefficients. The correlation between links and phylogenetic diversity in season 2 (h) was nonlinear and monotonic, so the correlation was assessed using a Spearman's rank correlation ($r_s$).

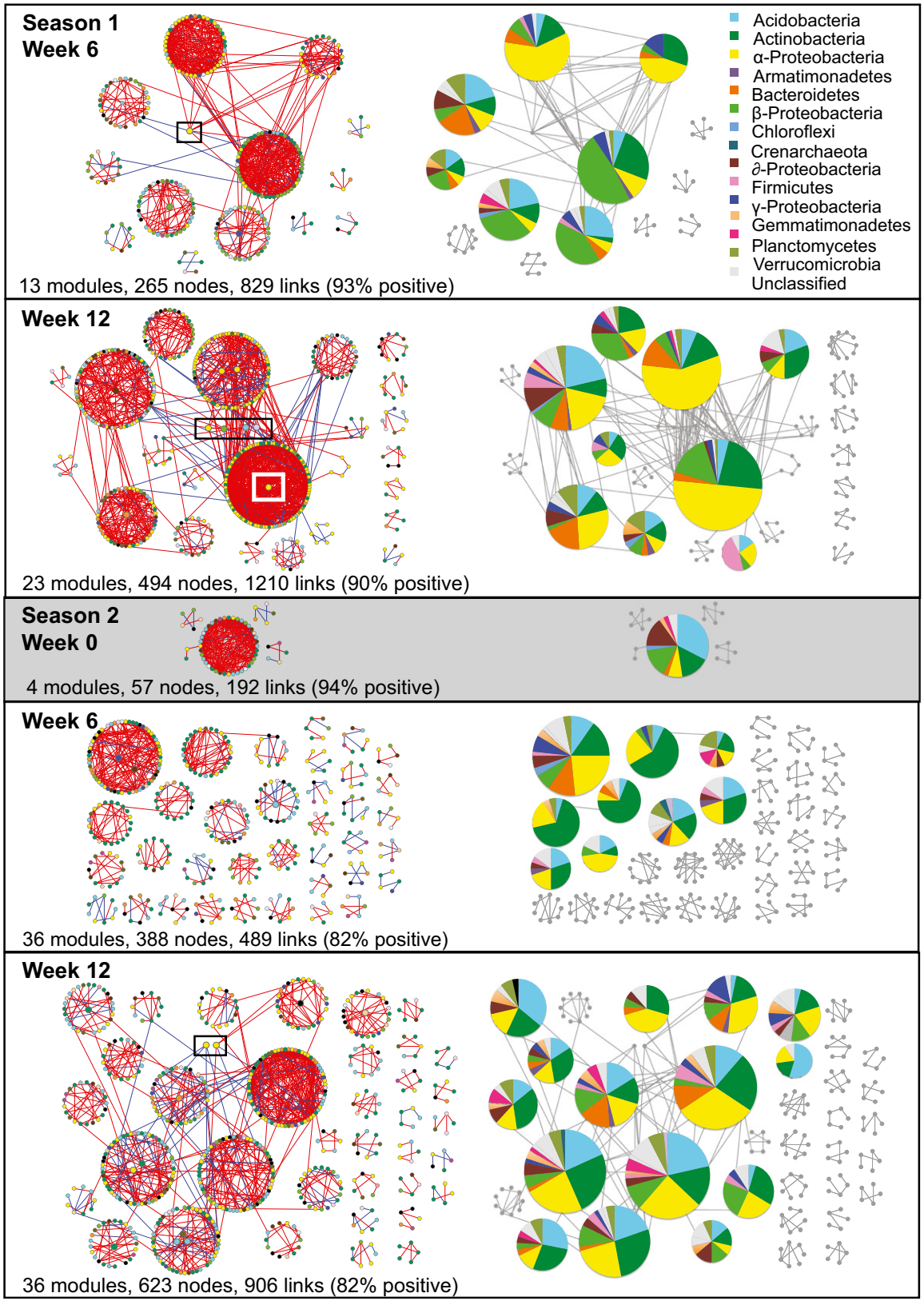


**Figure 4** Highly connected modules within rhizosphere and bulk soil networks at the vegetative and senescent phases of Season 1 and 2 (Weeks 6, 12), as well as the pre-planted soil from the beginning of Season 2 (Week 0). Colours of nodes indicate different major phyla (sub-phylum for Proteobacteria); pie charts represent the composition of modules with >10 nodes. A red link indicates positive covariation between two individual nodes, whereas a blue link indicates negative covariation. Nodes in the middle of modules are the module hubs, and nodes in the black boxes are connectors. Module hub enclosed by a white box was *c*. 99% similar to an isolate with demonstrated quorum sensing potential (Genbank EU723095.1).

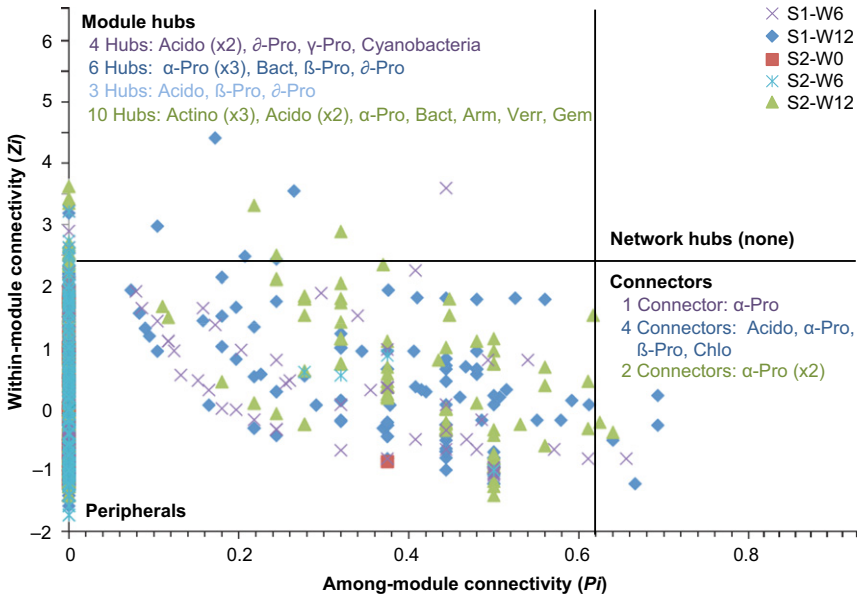


**Figure 5** Classification of nodes to identify putative keystone species within the rhizosphere networks. Each symbol represents an OTU from the five networks selected for detailed module analysis (Fig. 4). Modules hubs have $Zi > 2.5$, whereas connectors have $Pi > 0.62$. The phylogenetic affiliation of the hubs and connectors is listed on the plot using the following abbreviations: Acid, Acidobacteria; Actino, Actinobacteria; Arm, Armatimonadetes; Bact, Bacteriodetes; Chlo, Chloroflexi; Gem, Gemmatimonadetes; Pro, Proteobacteria; Verr, Verrucomicrobia. Detailed taxonomic information for module hubs and connectors is listed in Table S6.

(primarily Alphaproteobacteria), and the other two were from Acidobacteria and Chloroflexi (Figs 3 and 4; Table S6). No network hubs were detected in any of the networks, as no single node had $Pi > 0.62$ and $Zi > 2.5$ (Olesen *et al.* 2006).

Module hubs and connectors have been proposed to be keystone taxa due to their important roles in network topology (Deng *et al.* 2012). Based on this criterion, members of Proteobacteria phyla would be the most prominent keystone taxa in the rhizosphere networks, as they accounted for approximately half (48%) of all network hubs and connectors. Putative keystone taxa include taxa from the orders *Rhizobiales* (Alphaproteobacteria), *Burkholderiales* (Betaproteobacteria) and *Pseudomonadales* (Gammaproteobacteria) (Table S6). Interestingly, no single taxon acted as a keystone for multiple networks (Table S6). However, taxa from the same genera did reoccur between different networks. Two taxa from the genus *Rhizobacter* were classified as module hubs in week 6 rhizosphere networks for both seasons, and three taxa from the genus *Mesorhizobium* were identified as module hubs and connectors in week 12 rhizosphere networks for both seasons. Module hubs spanned a range of relative abundances (RA, from 0.01 to 2.53%), with many of the hub-taxa present in low relative abundance (Fig. 6, Table S6). Most of the connectors had low relative abundance (0.007 to 0.29%).

Spurred by our previous work (DeAngelis *et al.* 2008), in which quorum sensing (QS) was identified as a potential communication strategy in the *Avena fatua* rhizosphere, we compared the hubs identified in this study to isolates with demonstrated QS capabilities (DeAngelis *et al.* 2008). Notably, the hub of the largest module in this study (OTU_175932, Fig. 4, Season 1, Week 12, enclosed by white square) was a *Rhizobium* that is 98.7% similar to an isolate in which QS activity was detected by a whole-cell biosensor (Genbank EU723095.1) (DeAngelis *et al.* 2008). This organism was highly abundant (2.3% relative abundance). Approximately 50% of the other taxa in this module were other Alphaproteobacteria, of which 24 of 49 were > 97% similar to QS isolates from the DeAngelis *et al.* (2008) study (Fig. S3).

## DISCUSSION

We found that rhizosphere assemblages formed significantly larger and more complex networks than surrounding bulk soil communities, and that rhizosphere networks developed over time as the plant grew. These patterns reoccurred over two successive seasons of plant growth; soil in second season included root detritus from the previous season. We identified modules within the networks that likely result from microbe–microbe interactions or covariation in response to shared niches in the rhizosphere**.** Increased rhizosphere network connectivity and complexity are previously undescribed properties of rhizosphere bacterial assemblages, and represent fundamental differences between the rhizosphere microhabitat and the surrounding soil.

Multiple mechanisms may be responsible for increasing network size and complexity in the rhizosphere. Networks represent coordinated variability, where the members' abundances covary in response to interactions among the members or in response to environmental factors. Changes in environmental properties, such as pH and hydrological characteristics, have been shown to alter ecological networks (Tylianakis *et al.* 2007; Barberan *et al.* 2012). In particular, both macrobiological and microbiological studies have shown that resource and food availability are important drivers of social network structures (Henzi *et al.* 2009; Foster *et al.* 2012). For example

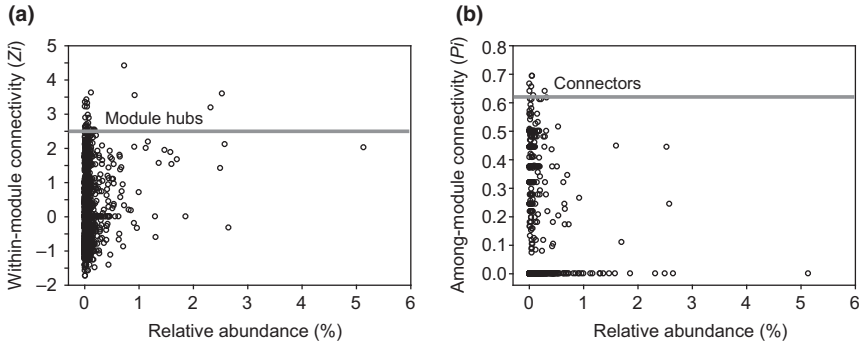


**Figure 6** The relative abundance of (a) module hubs and (b) module connectors in the rhizosphere networks. Grey lines indicate the threshold above which nodes are classified either as (a) module hubs based on within-module connectivity ($Zi > 2.5$), or as (b) module connectors based on among-module connectivity ($Pi > 0.62$).

elevated $CO_2$ has been shown to increase the phylogenetic and functional complexity of microbial networks in soil, which was likely due to the increased amount of C input into soil under elevated $CO_2$ (Zhou *et al.* 2010, 2011). Plant roots alter the immediate soil environment by changing pH, moisture and oxygen over time, as well as by inputting a significant amount of C into soil (Hinsinger *et al.* 2009; Chaparro *et al.* 2014). While the rhizosphere environment likely fosters greater direct and indirect interactions than bulk soil, it is also likely that some of the covariations detected represent niche-sharing within the gradients surrounding roots (Berry & Widder 2014). Therefore, we interpret increased network size and complexity in the rhizosphere as increased community organisation, which is combination of both increased bacterial interactions and the development of shared guilds or niches, and represents a fundamental difference between rhizosphere and bulk soil.

In contrast to the rhizosphere, networks in the surrounding soil remained relatively static and simple over time. Bulk soil organisms are thought to occupy heterogeneous, disconnected habitats that exist at the microscale (Torsvik *et al.* 2002; Fierer & Lennon 2011). The lack of networks in the bulk soil could indicate that interactions or niche sharing is minimal in the bulk soil relative to the rhizosphere, or it is possible that our sampling size was too coarse to detect microscale covariations. Alternatively, the lack of networks in bulk soil could also reflect the inactive or dormant state of many soil bacteria (Fierer & Lennon 2011).

We previously noted that bacterial richness and diversity decrease in rhizosphere communities as *A. fatua* grows (Shi *et al.* 2015); similar decreases in rhizosphere diversity have been observed in other studies (Kowalchuk *et al.* 2002; Chaparro *et al.* 2014). Here we report that diversity decreases as network size and connectivity increases in rhizosphere soil. These findings are not contradictory, as both likely result from the root acting as a strong environmental filter during rhizosphere microbial community assembly (Nuccio *et al.* 2016). In Shi *et al.* (2015) we show that even though rhizosphere bacteria were numerically more abundant than bulk soil bacteria, rhizosphere communities were less rich than bulk soil communities because particular taxa became dominant in rhizosphere soil over time. We postulate that roots promote the development of niches populated by dominant taxa, which would concurrently yield decreased diversity, greater interactions, greater co-variations due to shared niches, and overall result in more complex co-occurrence patterns over time. The inverse relationship between diversity and network connectivity highlights the importance of studying the relationships among organisms, as they are a crucial dimension of community organisation not captured by univariate diversity metrics (Zhou *et al.* 2010).

Because of the importance of modules in ecology and evolutionary biology, many studies with macroorganisms as well as microorganisms have focused on identifying modules in networks (Olesen *et al.* 2007; Dupont & Olesen 2009; Zhou *et al.* 2010, 2011; Deng *et al.* 2012). Within rhizosphere modules, we identified a small number of module hubs (i.e. nodes highly connected within a module) and connectors (i.e. nodes linking different modules together). Previous work has indicated that these organisms may function as keystone taxa, as they have disproportionately important roles in maintaining network structure relative to the other taxa in the network (Olesen *et al.* 2007; Faust & Raes 2012). The disappearance of these putative keystone taxa may cause modules and networks to disassemble (Paine 1995; Power *et al.* 1996), and thus keystone taxa may play a role in maintaining ecosystem stability (Olesen *et al.* 2007; Lu *et al.* 2013). In this study, no taxa acted as hubs or connectors in two different networks, which suggests that the conditions present were not identical over time, and supports the context dependency theory that keystone species play critical roles only under certain conditions (Power *et al.* 1996). Previous studies have found that the putative keystone species changed as conditions changed (Lu *et al.* 2013; Lupatini *et al.* 2014). Alternatively, functional redundancy may explain the unique keystone taxa detected in the rhizosphere networks; that is, different organisms may play the same functional role over time in different modules.

Interestingly, we found that the covariations occurring within modules were predominantly positive (> 80%), which represent both positive interactions as well as organisms occupying similar guilds or niches. Interactions encompass a spectrum ranging from antagonistic to cooperative, as exemplified

by resource competition verses quorum controlled activities (Berry & Widder 2014). For example in a groundwater study that used the same analysis algorithm employed in this study (Deng *et al*. 2015), negative co-occurrence patterns (co-exclusion) predominated and suggested that substrate injections triggered bacterial competition. There has been substantial recent discussion of positive interactions occurring among natural populations of bacteria (Morris *et al*. 2012; Hallam & McCutcheon 2015), including soil bacteria (Ren *et al*. 2015). The positive characteristic of the bacterial co-occurrence patterns in rhizosphere microbiomes in this study is consistent with cooperative or syntrophic interactions, and suggests the potential for extensive mutualistic interactions among bacteria in rhizosphere assemblages.

The modules detected in this study may contain organisms that interact, either directly or indirectly (Duran-Pinedo *et al*. 2011; Zhou *et al*. 2011). Microorganisms can communicate with each other and their eukaryotic partners through various signal molecules (Keller & Surette 2006). Density-dependent behaviour, such as quorum sensing (QS), is one example of microbial communication shown to control competitive as well as cooperative behaviours in microbial communities (Bassler & Losick 2006; Keller & Surette 2006). Rhizosphere microorganisms have been shown to use QS for a variety of ecological roles (e.g. virulence trait expression, biofilm formation, extracellular enzyme production, exopolysaccharide production) (Loh *et al*. 2002; DeAngelis *et al*. 2008), and rhizosphere microorganisms are more competent at producing signal molecules than bulk soil microorganisms [e.g. N-acyl-homoserine lactones (AHLs)] (Elasri *et al*. 2001). While bacterial traits and functions cannot be definitively predicted by phylogeny, many of the putative keystone taxa identified in this study are affiliated with groups previously shown to include taxa that use quorum sensing as a communication strategy (*Rhizobium*, *Burkholderiales, Pseudomonadales*) (Elasri *et al*. 2001; DeAngelis *et al*. 2008). Directly relevant to this study, DeAngelis *et al*. (2008) isolated AHL-producing microorganisms from the *A. fatua* rhizosphere growing in the same soil and greenhouse conditions as this study. One of these QS-capable isolates is highly similar to the hub within the largest module detected in this study (*c*. 99% similar to *Rhizobium* spp. by 16S rRNA gene) and this hub is one of the most abundant hubs identified. In addition, unlike the other modules found at this time point, approximately half of the alpha-Proteobacterial peripherals within this module were > 97% similar to the QS isolates. Our results are consistent with previous work that suggests QS may be a relevant interaction mechanism in rhizosphere communities.

Over last two decades, soil microbial ecology studies have commonly focused on relatively abundant taxa, although it is questionable whether abundant taxa are the most functionally important members of microbial communities. Using next-generation sequencing, we were able to investigate the importance of both high and relatively low abundance taxa in microbial communities (Caporaso *et al*. 2012). Interestingly, the majority of putative keystone taxa had relatively low abundances, which suggests that low abundance taxa may play important roles in maintaining network structures in rhizosphere microbial communities. Similarly, many macroecological network studies have reported that less abundant or even rare species were likely the keystone species in various ecosystems (see review by Power *et al*. (1996)). Lupatini *et al*. (2014) recently reported that rare microorganisms might act as important keystone taxa in the soil networks. Thus, less abundant taxa can be as important or more important than the abundant ones in maintaining microbial networks (Lyons & Schwartz 2001; Pester *et al*. 2010).

The analyses of networks based on 16S rRNA gene sequences limited this study to bacteria and archaea present in the soil. Clearly other members of the soil food web such as mycorrhizae (e.g. arbuscular mycorrhizae in this system), fauna and viruses are major biotic forces not explicitly included in our analyses. However, network analysis has much potential for exploring these multi-domain interactions.

In summary, this study reports previously undocumented network complexity in rhizosphere soils; this complexity developed over time, and reoccurred over two growth cycles of an annual plant. In contrast, the microbial networks in the surrounding bulk soil were relatively simple and static, and this difference in network development likely reflects fundamental properties of each habitat. Increases in network complexity were concurrent to decreases in bacterial diversity, which emphasises the need to characterize community organisation in addition to quantifying diversity. The co-occurrence patterns identified were predominantly positive, and quorum sensing was identified as one possible interaction strategy. We propose that network complexity represents an important, previously unrecognised dimension of rhizosphere microbial communities.

## ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research Genomic Science Program under Award Numbers DE-SC0004730 and DE-SC0010570. Work at LLNL was performed under the auspices of the U.S. DOE at Lawrence Livermore National Laboratory under Contract DE-AC52- 07NA27344, with funding provided by the U.S. DOE Genomic Science Program under contract SCW1421. Work at the University of Oklahoma was funded by the U.S. DOE under UC-subcontract number 00008322. We thank Jiabao Li (OU) for the MiSeq sequencing, Ye Deng (OU) for the assistance in using the network pipeline, and Nhu Nguyen (UCB) for a thoughtful review of this manuscript. The authors declare no conflict of interest.

## STATEMENT OF AUTHORSHIP

EN and SS both 1st authors. SS and MKF designed study. SS carried out experiment. JZ, ZH, ZS provided guidance on network analyses. SS and EN analysed and presented data. SS prepared the first draft, EN prepared the final manuscript, and all authors contributed substantially to revisions.

## REFERENCES


Barberan, A., Bates, S.T., Casamayor, E.O. & Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.*, 6, 343–351.

Bassler, B.L. & Losick, R. (2006). Bacterially Speaking. *Cell*, 125, 237–246.

Bastian, M., Heymann, S. & Jacomy, M. (2009). *Gephi: An Open Source Software for Exploring and Manipulating Networks*. International AAAI Conference on Weblogs and Social Media.

Berry, D. & Widder, S. (2014). Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front. Microbiol.*, 5, 1–14.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N. *et al.* (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.*, 6, 1621–1624.

Chaparro, J.M., Badri, D.V. & Vivanco, J.M. (2014). Rhizosphere microbiome assemblage is affected by plant development. *ISME J.*, 8, 790–803.

DeAngelis, K.M., Lindow, S.E. & Firestone, M.K. (2008). Bacterial quorum sensing and nitrogen cycling in rhizosphere soil. *FEMS Microbiol. Ecol.*, 66, 197–207.

Deng, Y., Jiang, Y.-H., Yang, Y., He, Z., Luo, F. & Zhou, J. (2012). Molecular ecological network analyses. *BMC Bioinformatics*, 13, 113.

Deng, Y., Zhang, P., Qin, Y., Tu, Q., Yang, Y., He, Z. *et al.* (2016). Network succession reveals the importance of competition in response to emulsified vegetable oil amendment for uranium bioremediation. *Environ. Microbiol.*, 18, 205–218.

Dupont, Y.L. & Olesen, J.M. (2009). Ecological modules and roles of species in heathland plant–insect flower visitor networks. *J. Anim. Ecol.*, 78, 346–353.

Duran-Pinedo, A.E., Paster, B., Teles, R. & Frias-Lopez, J. (2011). Correlation Network Analysis Applied to Complex Biofilm Communities. *PLoS ONE*, 6, e28438.

Elasri, M., Delorme, S., Lemanceau, P., Stewart, G., Laue, B., Glickmann, E. *et al.* (2001). Acyl-Homoserine Lactone Production Is More Common among Plant-Associated Pseudomonas spp. than among Soilborne Pseudomonas spp. *Appl. Environ. Microbiol.*, 67, 1198–1209.

Faust, K. & Raes, J. (2012). Microbial interactions: from networks to models. *Nat. Rev. Microbiol.*, 10, 538–550.

Faust, K., Sathirapongsasuti, J.F., Izard, J., Segata, N., Gevers, D., Raes, J. *et al.* (2012). Microbial Co-occurrence Relationships in the Human Microbiome. *PLoS Comput. Biol.*, 8, e1002606.

Fierer, N. & Lennon, J.T. (2011). The generation and maintenance of diversity in microbial communities. *Am. J. Bot.*, 98, 439–448.

Foster, E.A., Franks, D.W., Morrell, L.J., Balcomb, K.C., Parsons, K.M., van Ginneken, A. *et al.* (2012). Social network correlates of food availability in an endangered population of killer whales, *Orcinus orca*. *Anim. Behav.*, 83, 731–736.

Fuhrman, J. (2009). Microbial community structure and its functional implications. *Nature*, 459, 193–199.

Guimera, R. & Amaral, L. (2005). Functional cartography of complex metabolic networks. *Nature*, 433, 895–900.

Hallam, S.J. & McCutcheon, J.P. (2015). Microbes don't play solitaire: how cooperation trumps isolation in the microbial world. *Environ. Microbiol. Reports*, 7, 26–28.

Henzi, S.P., Lusseau, D., Weingrill, T., Schaik, C.P. & Barrett, L. (2009). Cyclicity in the structure of female baboon social networks. *Behav. Ecol. Sociobiol.*, 63, 1015–1021.

Hinsinger, P., Bengough, A.G., Vetterlein, D. & Young, I. (2009). Rhizosphere: biophysics, biogeochemistry and ecological relevance. *Plant Soil*, 321, 117–152.

Keller, L. & Surette, M.G. (2006). Communication in bacteria: an ecological and evolutionary perspective. *Nat. Rev. Microbiol.*, 4, 249–258.

Konopka, A. (2009). What is microbial community ecology? *ISME J.*, 3, 1223–1230.

Kowalchuk, G.A., Buma, D., de Boer, W., Klinkhamer, P.L. & van Veen, J. (2002). Effects of above-ground plant species composition and diversity on the diversity of soil-borne microorganisms. *Antonie Van Leeuwenhoek*, 81, 509–520.

Loh, J., Pierson, E.A., Pierson Iii, L.S., Stacey, G. & Chatterjee, A. (2002). Quorum sensing in plant-associated bacteria. *Curr. Opin. Plant Biol.*, 5, 285–290.

Lu, L., Yin, S., Liu, X., Zhang, W., Gu, T., Shen, Q. *et al.* (2013). Fungal networks in yield-invigorating and -debilitating soils induced by prolonged potato monoculture. *Soil. Biol. Biogeochem.*, 65, 186–194.

Lupatini, M., Suleiman, A., Jacques, R., Antoniolli, Z., Ferreira, A., Kuramae, E.E. *et al.* (2014). Network topology reveal high connectance levels and few key microbial genera within soils. *Front Environ. Sci.*, 2, 1–11.

Lyons, K.G. & Schwartz, M.W. (2001). Rare species loss alters ecosystem function – invasion resistance. *Ecol. Lett.*, 4, 358–365.

Morris, J.J., Lenski, R.E. & Zinser, E.R. (2012). The Black Queen hypothesis: evolution of dependencies through adaptive gene loss. *mBio*, 3, 1–7.

Newman, M. (2006). Modularity and community structure in networks. *Proc. Natl Acad. Sci. USA*, 103, 8577–8582.

Nuccio, E.E., Anderson-Furgeson, J., Estera, K., Pett-Ridge, J., de Valpine, P., Brodie, E.L. *et al.* (2016). Climate and edaphic controllers influence rhizosphere community assembly for a wild annual grass. *Ecology*, 97, 1307–1318.

Olesen, J., Bascompte, J., Dupont, Y. & Jordano, P. (2006). The smallest of all worlds: pollination networks. *J. Theor. Biol.*, 240, 270–276.

Olesen, J., Bascompte, J., Dupont, Y. & Jordano, P. (2007). The modularity of pollination networks. *Proc. Natl Acad. Sci. USA*, 104, 19891–19896.

Paine, R.T. (1995). A conversation on refining the concept of keystone species. *Conserv. Biol.*, 9, 962–964.

Pester, M., Bittner, N., Deevong, P., Wagner, M. & Loy, A. (2010). A 'rare biosphere' microorganism contributes to sulfate reduction in a peatland. *ISME J.*, 4, 1591–1602.

Philippot, L., Raaijmakers, J.M., Lemanceau, P. & van der Putten, W.H. (2013). Going back to the roots: the microbial ecology of the rhizosphere. *Nat. Rev. Microbiol.*, 11, 789–799.

Power, M.E., Tilman, D., Estes, J.A., Menge, B.A., Bond, W.J., Mill, L.S. *et al.* (1996). Challenges in the Quest for keystones. *Bioscience*, 46, 609–620.

Prosser, J.I., Bohannan, B.J.M., Curtis, T.P., Ellis, R.J., Firestone, M.K., Freckleton, R.P. *et al.* (2007). The role of ecological theory in microbial ecology. *Nat. Rev. Microbiol.*, 5, 384–392.

Proulx, S.R., Promislow, D.E.L. & Phillips, P.C. (2005). Network thinking in ecology and evolution. *Trends Ecol. Evol.*, 20, 345–353.

Ren, D., Madsen, J.S., Sorensen, S.J. & Burmolle, M. (2015). High prevalence of biofilm synergy among bacterial soil isolates in cocultures indicates bacterial interspecific cooperation. *ISME J.*, 9, 81–89.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D. *et al.* (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.*, 13, 2498–2504.

Shi, S., Nuccio, E., Herman, D.J., Rijkers, R., Estera, K., Li, J. *et al.* (2015). Successional trajectories of rhizosphere bacterial communities over consecutive seasons. *mBio*, 6, 1–8.

Steele, J.A., Countway, P.D., Xia, L., Vigil, P.D., Beman, J.M., Kim, D.Y. *et al.* (2011). Marine bacterial, archaeal and protistan association networks reveal ecological linkages. *ISME J.*, 5, 1414–1425.

Team, R.C. (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria.

Thompson, J.N. (2005). *The Geographic Mosaic of Coevolution*. University of Chicago Press Chicago, IL.

Torsvik, V., Øvreas, L. & Thingstad, T.F. (2002). Prokaryotic Diversity-Magnitude, Dynamics, and Controlling Factors. *Science*, 296, 1064–1066.

Tylianakis, J.M., Tscharntke, T. & Lewis, O.T. (2007). Habitat modification alters the structure of tropical host-parasitoid food webs. *Nature*, 445, 202–205.

Zhou, J., Deng, Y., Luo, F., He, Z., Tu, Q. & Zhi, X. (2010). Functional molecular ecological networks. *mBio*, 1, 1–10.

Zhou, J., Deng, Y., Luo, F., He, Z. & Yang, Y. (2011). Phylogenetic molecular ecological network of soil miobial communities in response to elevated $CO_2$. *mBio*, 2, e00122-11.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

Editor, Nancy Johnson
Manuscript received 26 November 2015
First decision made 8 January 2016
Second decision made 6 April 2016
Manuscript accepted 8 May 2016