**Land use conversion increases network complexity and stability of soil microbial communities in a temperate grassland**

Carolyn R. Cornell1,2,7,8,, Ya Zhang1,2,8, Daliang Ning1,2, Naijia Xiao1,2, Pradeep Wagle3, Xiangming Xiao1, Jizhong Zhou1,2,4,5,6\*

1Department of Microbiology and Plant Biology, University of Oklahoma, Norman OK, USA.

2Institute for Environmental Genomics, University of Oklahoma, Norman, OK, USA.

3USDA, Agricultural Research Service, Oklahoma and Central Plains Agricultural Research Center, El Reno, OK, USA

4School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, Oklahoma, USA.

5School of Computer Science, University of Oklahoma, Norman, Oklahoma, USA

6Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

7Department of Civil and Environmental Engineering, Rice University, Houston, TX, USA

8C.R.C. and Y.Z. contributed equally to this work.

\*To whom correspondence may be addressed: Jizhong Zhou, [jzhou@ou.edu](mailto:jzhou@ou.edu)

**Contents**

1. **Supplementary Figures 1-6**
2. **Supplementary Tables 1-8**
3. **Supplementary Text A-E**

**Supplementary Figures**

**A picture containing text, circle, diagram, colorfulness

Description automatically generated**

**Figure S1. Visualization of soil microbial networks over time.** MENs were constructed for nineteen sampling months from June 2017 – December 2018. As networks on the native TGP system remained relatively similar over, a single visual representation of the control site is outlined in a black square. The other networks depict the temporal differences of the nineteen MENs in the CT wheat land use. Large modules with ≥10 node are shown in different colors, and smaller modules are shown in gray. The average K and modularity are shown above each network. Detailed network topological attributes are listed in Supplementary Table 1.

A picture containing text, diagram, screenshot, plot

Description automatically generated

**Figure S2. Community structure and alpha-diversity comparisons between networked and whole microbial community under CT wheat and native TGP land use. a**) Principal coordinate analysis (PCoA) of the structure of the networked communities compared to whole microbial community based on Bray-Curtis distance. Networked communities were significantly different (*p* ≤ 0.05) based on land use and sampling month according to three non-parametric permutations tests found in Table 1. Red circles represent networks under CT wheat land use and blue squares represent native TGP (control) land use. Yellow triangles represent the whole microbial community under CT wheat land use and green diamonds represent the whole microbial community under native TGP land use. **b**)Comparisons of microbial community richness. phylogenetic diversity (Faith’s PD), and Shannon diversity. Richness for networked microbial communities represented by number of ASVs present in 6 of 8 replicates used for network construction. CT wheat show in red and native tallgrass prairie shown in blue. For each land use, all indices were significantly different (*p* < 0.001) between the networked and whole microbial community based on Mann-Whitney U test.

A picture containing screenshot, text, colorfulness, design

Description automatically generated

**Figure S3. Taxonomic composition of networked microbial communities under CT wheat cropland and native prairie.** Microbial community relative abundance at the **a**) phylum level and **b**) class level. Detailed significant (*p* ≤ 0.05) changes between taxa based on Mann-Whitney U test due to land use conversion for long-term CT wheat land use in Supplementary Table 2.

A picture containing text, diagram, screenshot, line

Description automatically generated

**Figure S4. Relationship between network topological properties and relative modularity.** Spearman correlations were used to compare network topologies including nodes, links, average K, average CC, GD, and Con to the relative modularity. Red circles represent networks under CT wheat land use and blue squares represent native TGP (control) land use. Correlation coefficients (rho) are shown in corresponding colors followed by \* *p* ≤ 0.05, \*\* *p* ≤ 0.01, \*\*\* *p* ≤ 0.001.

**A picture containing screenshot, diagram, plot, text

Description automatically generated**

**Figure S5. Additional temporal dynamics of network stability. a**) Robustness measured by removing five module hubs from each of the empirical MENs. For April 2018, only four module hubs removed as that was the total module hubs in CT wheat for that network. CT wheat shown in red and native TGP shown in blue for all graphs. Error bars represent the standard deviation of 100 repetitions of the simulation. Robustness for each timepoint was compared between CT wheat and native TGP land use using a two-sided t-test. Significant differences are indicated by \* *p* ≤ 0.05, \*\* *p* ≤ 0.01, \*\*\* *p* ≤ 0.001. **b**) Compositional stability of the networked community over time shown as consecutive monthly comparisons. Overall compositional stability between CT wheat and native TGP land use compared using Mann-Whitney U Test. **c**) The number of overlapping ASVs in the whole community under CT wheat and native TGP land use among different numbers of sampling times (that is, orders). For example, for order=2, the overlapping ASVs were between any two sampling times. **d**) Spearman correlations between network stability and network complexity indices under TGP land use. Significant correlations (*p* ≤ 0.05) are shown in blue for positive correlations and green for negative correlations. Inside the cells are the corresponding correlation coefficients. Non-significant correlations are shown in gray. **e**) Network node constancy. Each box shows the constancy distribution of all nodes, averaged between land uses. **f**) Unweighted network link constancy. Each box shows the constancy distribution of the links in the networks under each land use. For e and f, Mann-Whitney U Test was used to compare differences in constancy between the CT wheat and native TGP land use.

**A picture containing text, diagram, screenshot, line

Description automatically generated**

**Figure S6. Associations between networked community structure and environmental properties**. Correlations of the networked community structures (Bray-Curtis distance) and soil and climate properties for the TGP land use. Edge width corresponds to Mantel’s r value and the edge color represents the statistical significance. Pairwise spearman correlations of the variables are shown with a color gradient representing the correlation coefficients

**Supplementary Tables**

**Table S1.** Topological properties of the MENs of converted cropland (CT wheat) and native tallgrass prairie (TGP, control)

**Table S2.** Mann-Whitney U tests on relative abundances of highly abundant microbial phyla and classes between CT wheat and TGP.

**Table S3.** List of preserved module pairs, the number of overlapping and non-overlapping nodes in paired modules.

**Table S4.** Number of keystone nodes in CT wheat and TGP MENs.

**Table S5.** List of network hubs, their taxonomic information (% identity), and relative abundances.

**Table S6.** List of module hubs, their taxonomic information (% identity), and relative abundances.

**Table S7.** List of connectors, their taxonomic information (% identity), and relative abundances.

**Supplementary Text**

1. **Detailed site description and sampling strategy**

The tallgrass prairie (TGP) plant community is predominantly warm-season mixed grasses native to Oklahoma including big bluestem (*Andropogon gerardii* Vitman.), little bluestem (*Schizachyrium scoparium* (Michx.) Nash), Indiangrass (*Sorghastrum nutans* (L) Nash), and switchgrass (*Panicum vergatum* L.). The soil is classified as Norge loamy prairie (Fine, mixed, thermic Udertic Paleustalf) with a high water holding capacity. Grazing in the tallgrass prairie is under a year-round rotation by 50 cow-calf pairs being grazed for 30-day periods followed by 90-day rest periods. Prescribed spring burns are implemented on 4-year rotations as part of routine management. Burns generally occur in early spring before the initiation of growth of warm-season grasses. Recent prescribed burns occurred in February 2013 and February 2018.

Native tallgrass prairies are often converted into cultivated ecosystems dominated by the cool-season crop winter wheat in the U.S. Southern Plains. Winter wheat has been planted in the study site under conventional tillage (CT) management since the late 1990s. Concerning the studied period (June 2017-December 2018), the CT wheat field was grain-only during the 2015-2016 growing season and graze-out wheat (no grain production; cattle grazing from November through April/May) during the 2016-2017, 2017-2018, and 2018-2019 growing seasons. The soil in the wheat field was disced with a tandem disc harrow (10-12 cm depth) in May 2017 and June 2018 after the wheat growing season. Before planting in late September, the field was fertilized, tilled, and seedbed prepared. Since wheat was never harvested for grain, a burndown herbicide was applied in May 2018 between the 2017-2018 and 2018-2019 growing season with the field being resowed in November 2018.

Although soils are highly heterogeneous at the microscale (1), diversity and community structure are also manifested at larger scales (2). It has been demonstrated that at the ecosystem scale that changes in bacterial community composition and structure are more subtle (3). Therefore, if sampling resolution is carefully chosen, one can then give an interpretation of the estimated co-occurrence patterns (4). Additionally, a relatively larger number of samples is essential for co-occurrence patterns and should ideally include spatial or temporal gradients in environmental conditions to achieve sufficient variability in taxon abundances to resolve co-occurrence patterns (5). For these reasons, we chose to collect multiple soil samples located farther apart (~20 m) with each replicate made of several pooled soil cores to try to cover more of the distribution range of species in the site (4) and minimize the impact of the difference that may be caused by uneven management, plant diversity, and variation in soil properties.

1. **Network properties**

As an important part of many ecosystems, microbial communities are inherently complex (6). Despite the many types of complex systems, the architecture of their networks shares many common properties such as scale-free, small-world, modularity, and hierarchy (7-9). Overall, the networks in this study were scale-free, small-world, and modular under CT wheat and native land use (Table S1). Scale-free networks are highly non-uniform with most nodes or species having few links or interactions while few nodes have many links (10, 11). Additionally, in complex networks, most nodes can be reached through a path of few links (i.e. “small-world effect”) which has been observed in many network types including microbial networks (12), neural networks (13), and the World Wide Web (7). Lastly, a module in a network is a group of nodes that are highly connected within the group but have very few connections outside the group. Modules are often thought of as communities or functional units (14) which have been suggested to arise from habitat heterogeneity, specificity of interactions, ecological niche overlap, resource partitioning, phylogenetic relatedness, and/or natural selection (15). In this study, while the native TGP MENs were more modular (0.925 – 0.959 and 0.348 – 0.940, respectively), MENs of both land uses were modular based on modularity (> 0.35) and relative modularity (> 0, modularity of empirical MENs greater than modularity of random MENs).

**C. “Biotic interactions” in MENs community assembly**

Microorganisms are thought to follow similar patterns of established ecological theory for large organisms including patterns of community assembly (6). Microbial community assembly processes are important for determining community composition and linking microbes to ecosystem function (16, 17) with environmental filtering, dispersal limitation, and biotic interactions highly contributing to observed community patterns (18). While there are many methods for predicting biotic interactions (19-21), when using co-occurrence or correlations methods, it is thought that biotic interactions provide enough of a regular signal in the data to be detected over ecological and environmental processes for larger organisms (4, 22). Although, caution is still greatly needed when interpreting “biotic interactions” since the associations could depend on many other factors such as community complexity, population abundance, environmental conditions, and spatial and temporal scales (12, 23).

In order to detect the potential contributions of environmental filtering, dispersal limitations, and biotic interactions in the constructed molecular ecological networks (MENs), we used a combination of statistical approaches that have previously been applied for the same purpose (12). First, canonical correspondence analysis (CCA) was used to determine if soil properties, climate variables, and distance between samples played a significant role in controlling the responses of the networked communities to land conversion. Only non-redundant variables that generated a significant or moderately significantly CCA model (*p* ≤ 0.05 or *p* ≤ 0.1) by term were included. Overall, soil properties including soil pH, topsoil nitrate (NO3-), soil organic matter (OM), and available phosphorus (P) played a highly significant (*p* = 0.001) role in influencing the MENs. Variation partitioning analysis (VPA) based on CCA (Figure 2a) also indicated soil properties explained (17.4%) the largest amount of variation of the soil MEN compared to climate variables (3.4%) and distance between samples (3.7%). Nevertheless, the majority of the variation could not be explained by the measured environmental variables.

Lastly, we tested whether each link in a network was related to environmental filtering or dispersal limitation using the Link Test for Environmental filtering or Dispersal limitation (LTED) (12, 24). For example, if a link between two nodes is due to environmental filtering, then a strong correlation between each node and the environmental variable of interest should be detected. All links in the seven MENs where environmental variables were measured for both land uses were tested using the significant soil and climate variables used in CCA and VPA at the network correlation cutoff (St = 0.96) and a lower correlation threshold cutoff of |*r*| ≥ 0. For dispersal limitation, the abundance distribution of two species or linked nodes would both be expected to covary with spatial distance if impacted by dispersal limitations. This idea was applied to all links in the networks used for environmental filtering using either *r* ≥ 0 or *r* ≥ 0.5 (Pearson correlation) with *p* < 0.05. Using the LTED test, only very minor contributions from environmental filtering (St = 0.96 and St = 0.8) were observed (Figure 2b), and a minimal portion of links (Figure 2c) in the networks on average in both fields (*p* ≤ 0.05, *r* > 0) were impacted by dispersal limitation. However, while it is probable that biotic interactions could be a key part of CT wheat and native TGP networks based on our results, we cannot rule out potential contributions from unmeasured environmental variables and differences in micro-scale environmental conditions in shaping MENs.

**D. Changes in community composition and keystone nodes**

The presence of microbes in the soil environment is immensely affected by many factors such as abiotic stressors, competition with other microbes, disturbance, and uneven resource distribution (25). Conversion of natural ecosystems to agricultural lands can exacerbate many of these factors resulting in a rapidly fluctuating environment which has been shown to have a significant impact on microbial community composition and diversity (26). In this study, converting native land to long-term CT wheat use resulted in significant (*p* ≤ 0.05) changes in the community composition including increases in the relative abundance of Actinobacteria, Armatimonadetes, Bacteroidetes, Chloroflexi, Cyanobacteria, Gemmatimonadetes, Alphaproteobacteria, and Betaproteobacteria accompanied by significant decreases in the relative abundance of Acidobacteria, Thaumarcheota, Verrucomicrobia, and Deltaproteobacteria (Figure S3, Table S2).

Similarly, the identity of the keystone nodes also differed between land uses. The phyla Proteobacteria, Actinobacteria, and Acidobacteria represented the majority of keystone nodes in both sites and frequently are recognized as keystone taxa (27, 28). The native TGP keystone nodes overlapped with many of those previously detected in grasslands (Sphingobacteriales, Actinomycetales, Acidobacteria GP4) while the keystone nodes of the CT wheat field corresponded to those for other agricultural lands (Rhizobiales, Solirubrobacterales) and plant-associated microbes (Acidobacteria GP1, Acidobacteria GP3) (Table S5-S7) (29). Therefore, land use conversion for long-term cropland not only shifted the microbial community structure but also changed the keystone nodes which drove community composition regardless of their abundance.

**E. Network complexity and stability**

Compared with the native TGP land use, CT wheat underwent many major management events over the 19-month sample period (Figure 1c). In summary, after the wheat spring growing season ended in the CT land use, the field was left fallow. It was tilled several times between June and September to mix the soil and prepare the seedbed. Seedbed preparation was closely accompanied by fertilizer application and seed planting. For the 2018-2019 season, the field was again left fallow after the spring and prepared using herbicide and soil tillage. Finally, the field was fertilized during the fall and bare spots resowed in late November. While it has been reported that land use and management intensity alters microbial community structure and functional potential in these fields (26), it also appeared to greatly impact the complexity of the microbial community associations. For example, the average K, number of links, and modularity highly fluctuated throughout the sampling period for CT wheat (Figure 1c) with increases in associations (i.e. links and avgK) and decreases in modularity during times of intensive disturbance (e.g. tillage). Although these changes in complexity could also be due to environmental conditions, it is likely that management input played a larger role given that the native TGP experienced the same climatic conditions with no noticeable change in complexity.

As connectivity and network size were considerably different among the CT wheat MENs, we calculated the relative modularity which takes into account the mean expected modularity resulting in a more reliable measure of modularity to compare across MENs. The RM of the CT wheat cropland had a greater slope and was significantly correlated (*p* ≤ 0.05) with many of the network topological properties including nodes, links, average connectivity, average clustering coefficient, path distance, and connectance (Figure S4). In comparison, the RM of the native TGP networks only significantly correlated (*p* ≤ 0.05) with links and average connectivity. For these two significant correlations, there was very minimal change in the RM and associated complexity indices. Therefore, CT wheat land use significantly enhanced the relationship between relative modularity and network complexity.

Determining the stability of a network is challenging and often depends on many varying network properties including the networks features (i.e. small world, scale free, and modular) and the complexity of the network (12). Through simulations, CT wheat MENs appeared to be more stable with greater robustness and less vulnerability to node loss (Figure 3, Figure S5). Nevertheless, the CT wheat networks were highly dependent on several keystone nodes connecting the networks as opposed to the native TGP networks which were very modular (30). It remained unknown how the CT wheat MENs would respond to loss of one or more network hubs. This is important as scale-free networks are generally unstable when a keystone node is lost compared to random node loss (7). On the one hand, greater network complexity as observed for CT wheat has also been linked to greater resistance (31, 32) and decreased resilience (33). On the other hand, greater network size is part of greater complexity with TGP networks being significantly larger. Therefore, network stability is impacted by different dimensions of microbial community structures and interactions, with management further influencing stability by making network stability highly dependent on network complexity under CT wheat land use.

**References in Supplementary Section**

1. Zhou J, Xia B, Treves DS, Wu L-Y, Marsh TL, O’Neill RV, et al. Spatial and resource factors influencing high microbial diversity in soil. Appl Environ Microbiol. 2002;68(1):326-34.

2. Kaiser K, Wemheuer B, Korolkow V, Wemheuer F, Nacke H, Schöning I, et al. Driving forces of soil bacterial community structure, diversity, and function in temperate grasslands and forests. Sci Rep. 2016;6(1):33696.

3. O'Brien SL, Gibbons SM, Owens SM, Hampton-Marcell J, Johnston ER, Jastrow JD, et al. Spatial scale drives patterns in soil bacterial diversity. Environ Microbiol. 2016;18(6):2039-51.

4. Blanchet FG, Cazelles K, Gravel D. Co-occurrence is not evidence of ecological interactions. Ecol Lett. 2020;23(7):1050-63.

5. Barberán A, Bates ST, Casamayor EO, Fierer N. Using network analysis to explore co-occurrence patterns in soil microbial communities. ISME J. 2012;6(2):343-51.

6. Fuhrman JA. Microbial community structure and its functional implications. Nature. 2009;459(7244):193-9.

7. Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet. 2004;5(2):101-13.

8. Alon U. Biological networks: the tinkerer as an engineer. Science. 2003;301(5641):1866-7.

9. Clauset A, Moore C, Newman MEJ. Hierarchical structure and the prediction of missing links in networks. Nature. 2008;453(7191):98-101.

10. Barabási A-L, Albert R. Emergence of scaling in random networks. Science. 1999;286(5439):509-12.

11. Deng Y, Jiang YH, Yang Y, He Z, Luo F, Zhou J. Molecular ecological network analyses. BMC Bioinform. 2012;13:113.

12. Yuan MM, Guo X, Wu L, Zhang Y, Xiao N, Ning D, et al. Climate warming enhances microbial network complexity and stability. Nat Clim Change. 2021;11(4):343-8.

13. Watts DJ, Strogatz SH. Collective dynamics of ‘small-world’ networks. Nature. 1998;393(6684):440-2.

14. Luo F, Zhong J, Yang Y, Zhou J. Application of random matrix theory to microarray data for discovering functional gene modules. Phys Rev E. 2006;73(3):031924.

15. Olesen JM, Bascompte J, Dupont YL, Jordano P. The modularity of pollination networks. Proc Natl Acad Sci USA. 2007;104(50):19891.

16. Graham EB, Knelman JE, Schindlbacher A, Siciliano S, Breulmann M, Yannarell A, et al. Microbes as engines of ecosystem function: When does community structure enhance predictions of ecosystem processes? Front Microbiol. 2016;7.

17. Ferrenberg S, O'Neill SP, Knelman JE, Todd B, Duggan S, Bradley D, et al. Changes in assembly processes in soil bacterial communities following a wildfire disturbance. ISME J. 2013;7(6):1102-11.

18. D'Amen M, Mod HK, Gotelli NJ, Guisan A. Disentangling biotic interactions, environmental filters, and dispersal limitation as drivers of species co-occurrence. Ecography. 2018;41(8):1233-44.

19. Ulrich W, Gotelli NJ. Pattern detection in null model analysis. Oikos. 2013;122(1):2-18.

20. Zhou J, Deng Y, Luo F, He Z, Tu Q, Zhi X. Functional molecular ecological networks. mBio. 2010;1(4).

21. Montesinos-Navarro A, Hiraldo F, Tella JL, Blanco G. Network structure embracing mutualism–antagonism continuums increases community robustness. Nat Ecol Evol. 2017;1(11):1661-9.

22. Gotelli NJ, McCabe DJ. Species co-occurence: a meta-analysis of J. M. Diamond’s assembly rules model. Ecology. 2002;83(8):2091-6.

23. Freilich MA, Wieters E, Broitman BR, Marquet PA, Navarrete SA. Species co-occurrence networks: Can they reveal trophic and non-trophic interactions in ecological communities? Ecology. 2018;99(3):690-9.

24. Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, et al. Ocean plankton. Determinants of community structure in the global plankton interactome. Science. 2015;348(6237):1262073.

25. Fierer N. Embracing the unknown: disentangling the complexities of the soil microbiome. Nat Rev Microbiol. 2017;15(10):579-90.

26. Cornell CR, Zhang Y, Ning D, Wu L, Wagle P, Steiner JL, et al. Temporal dynamics of bacterial communities along a gradient of disturbance in a U.S. southern plains agroecosystem. mBio. 2022;13(3):e03829-21.

27. Lupatini M, Suleiman AKA, Jacques RJS, Antoniolli ZI, de Siqueira Ferreira A, Kuramae EE, et al. Network topology reveals high connectance levels and few key microbial genera within soils. Front Environ Sci. 2014;2.

28. Ma B, Wang H, Dsouza M, Lou J, He Y, Dai Z, et al. Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. ISME J. 2016;10(8):1891-901.

29. Banerjee S, Schlaeppi K, van der Heijden MGA. Keystone taxa as drivers of microbiome structure and functioning. Nat Rev Microbiol. 2018;16(9):567-76.

30. Kitano H. Biological robustness. Nat Rev Genet. 2004;5(11):826-37.

31. Landi P, Minoarivelo HO, Brännström Å, Hui C, Dieckmann U. Complexity and stability of ecological networks: a review of the theory. Popul Ecol. 2018;60(4):319-45.

32. Okuyama T, Holland JN. Network structural properties mediate the stability of mutualistic communities. Ecol Lett. 2008;11(3):208-16.

33. Pimm SL. The complexity and stability of ecosystems. Nature. 1984;307(5949):321-6.