

Metaphenomic Responses of a Native Prairie Soil Microbiome to Moisture Perturbations

 Taniya Roy Chowdhury,^{a*} Joon-Yong Lee,^a Eric M. Bottos,^a Colin J. Brislawn,^a Richard Allen White III,^a Lisa M. Bramer,^b Joseph Brown,^a Jeremy D. Zucker,^a Young-Mo Kim,^a Ari Jumpponen,^c Charles W. Rice,^c Sarah J. Fansler,^a Thomas O. Metz,^a Lee Ann McCue,^a Stephen J. Callister,^a Hyun-Seob Song,^a Janet K. Jansson^a

^aEarth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, Washington, USA

^bNational Security Directorate, Pacific Northwest National Laboratory, Richland, Washington, USA

^cDivision of Biology, Kansas State University, Manhattan, Kansas, USA

ABSTRACT Climate change is causing shifts in precipitation patterns in the central grasslands of the United States, with largely unknown consequences on the collective physiological responses of the soil microbial community, i.e., the metaphenome. Here, we used an untargeted omics approach to determine the soil microbial community's metaphenomic response to soil moisture and to define specific metabolic signatures of the response. Specifically, we aimed to develop the technical approaches and metabolic mapping framework necessary for future systematic ecological studies. We collected soil from three locations at the Konza Long-Term Ecological Research (LTER) field station in Kansas, and the soils were incubated for 15 days under dry or wet conditions and compared to field-moist controls. The microbiome response to wetting or drying was determined by 16S rRNA amplicon sequencing, metatranscriptomics, and metabolomics, and the resulting shifts in taxa, gene expression, and metabolites were assessed. Soil drying resulted in significant shifts in both the composition and function of the soil microbiome. In contrast, there were few changes following wetting. The combined metabolic and metatranscriptomic data were used to generate reaction networks to determine the metaphenomic response to soil moisture transitions. Site location was a strong determinant of the response of the soil microbiome to moisture perturbations. However, some specific metabolic pathways changed consistently across sites, including an increase in pathways and metabolites for production of sugars and other osmolytes as a response to drying. Using this approach, we demonstrate that despite the high complexity of the soil habitat, it is possible to generate insight into the effect of environmental change on the soil microbiome and its physiology and functions, thus laying the groundwork for future, targeted studies.

IMPORTANCE Climate change is predicted to result in increased drought extent and intensity in the highly productive, former tallgrass prairie region of the continental United States. These soils store large reserves of carbon. The decrease in soil moisture due to drought has largely unknown consequences on soil carbon cycling and other key biogeochemical cycles carried out by soil microbiomes. In this study, we found that soil drying had a significant impact on the structure and function of soil microbial communities, including shifts in expression of specific metabolic pathways, such as those leading toward production of osmoprotectant compounds. This study demonstrates the application of an untargeted multi-omics approach to decipher details of the soil microbial community's metaphenotypic response to environmental perturbations and should be applicable to studies of other complex microbial systems as well.

KEYWORDS metaphenome, metatranscriptome, multi-omics, soil microbiome

Citation Roy Chowdhury T, Lee J-Y, Bottos EM, Brislawn CJ, White RA, III, Bramer LM, Brown J, Zucker JD, Kim Y-M, Jumpponen A, Rice CW, Fansler SJ, Metz TO, McCue LA, Callister SJ, Song H-S, Jansson JK. 2019. Metaphenomic responses of a native prairie soil microbiome to moisture perturbations. *mSystems* 4:e00061-19. <https://doi.org/10.1128/mSystems.00061-19>.

Editor Steven J. Hallam, University of British Columbia

Copyright © 2019 Roy Chowdhury et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Janet K. Jansson, janet.jansson@pnnl.gov.

* Present address: Taniya Roy Chowdhury, Environmental Science and Technology Department, University of Maryland, College Park, Maryland, USA.

 The phenotypic response of the soil microbiome to changes in soil moisture was determined using an untargeted multi-omics approach

Received 30 January 2019

Accepted 12 May 2019

Published 11 June 2019

Climate change is predicted to result in increasing drought frequency and more infrequent, but higher-intensity, rainfall events in the highly fertile and agriculturally productive grasslands of the central United States (1). The southwestern and south-central grassland regions of North America are predicted to experience reduced precipitation in contrast to predicted higher precipitation in the central and northern grasslands (2, 3). As precipitation regimes with fewer, but larger, events amplify fluctuations in soil water content, the consequences in mesic grassland ecosystems will be prolonged dry periods between rainfall events and an increase in the length and occurrence of drought stress (4). These grassland soils store large amounts of carbon, up to 39% of the organic carbon stocks in the conterminous United States (5), and the fate of this stored C is uncertain as climate changes. Reportedly, grasslands in the U.S. Great Plains function as net carbon sinks (6, 7) but can become net carbon sources under drought (7, 8). Results from a series of modeling experiments simulating various drought magnitudes showed that short-term droughts caused greater carbon loss from grassland ecosystems (9, 10). However, contrasting results were obtained from a long-term rainfall intensity manipulation experiment at the Konza Long-Term Ecological Research (LTER) field station in Kansas (11), where soil microbial biomass and carbon use efficiency were higher in the treatments that had an extended duration between rainfall events, compared to ambient rainfall frequency. This study suggested that shifts in the soil microbiome following drying could result in a reduction in soil carbon loss and hypothesized that this could be due to an ecological selection for microbial species that are adapted to low-soil-water conditions (11).

It is well known that bacteria have evolved various life strategies to cope with soil desiccation, such as osmoregulation, dormancy/reactivation, and extracellular enzyme synthesis to accumulate substrates (12). As soil matric potentials become more negative, osmotic adjustments allow microbial cells to maintain turgor and function (13). Therefore, we predicted that soil drying would result in shifts in metabolic pathways and metabolites for osmoregulation (13). Understanding the physiological responses of the soil microbiome to changes in soil moisture is ultimately important for predicting and managing impacts of climate change on soil biogeochemistry (14) and grassland ecosystem productivity (15–17).

To date, most studies of soil microbial responses to wetting and drying have largely relied on bulk measurements of soil respiration, without knowledge of the underlying details of metabolic pathways that result in the measured release of CO₂. More recently, 16S rRNA gene sequencing (16S) and metagenomic (MG) sequencing have revealed the identities and potential functions of members of soil microbial communities. However, bulk DNA extraction methods can include DNA from dead cells and dormant cells that are not actively growing at the time of extraction. Importantly, dormant cells represent a seed bank that can be resuscitated to an active state if the environmental conditions are suitable. Thus, a critical knowledge gap exists about how indigenous, active, and naturally interacting members of soil microbial communities respond to changes in soil moisture by changing levels of gene expression, as measured by metatranscriptomics (MT).

Here, we aimed to determine the metabolic pathways underpinning the soil microbial community's phenotypic response to wetting and drying, i.e., to determine the metaphenome (18), using native prairie soils collected from three locations at the Konza LTER field station in Kansas (see Fig. S1 in the supplemental material). The moisture perturbations used in this study are particularly relevant as this central mesic grassland region is at the crossroads of predicted precipitation regime changes (2, 19). Our central hypothesis was that wetting and drying would result in consistent shifts in specific metabolic pathways utilized by the soil microbiome across different prairie soil field locations. Our simple experimental design, with three field locations and three treatments (control, wet, and dry), enabled us to optimize the omics and modeling approaches necessary to determine key metabolic pathways expressed by active members of the prairie soil microbiome and to determine how those pathways were impacted by changes in soil moisture. For example, we predicted that there would be

similar shifts across field locations in expression of metabolic pathways and production of metabolites for adaptation to soil drying, such as an increase in pathways for osmolyte production (13), as also shown in studies using biocrust isolates (20). By further integration across omics data types using reaction network approaches, we aimed to use this information to help define the metaphenome.

RESULTS

Native prairie soils were collected in triplicate from three field locations (soils A, B, and C; see Fig. S1 in the supplemental material) across a landscape gradient at the Konza Long-Term Ecological Research (LTER) site in Kansas (11). The soil samples were subjected to contrasting moisture perturbations, i.e., wetting to saturation or desiccation by air drying, over a 15-day period. Soils from the three locations differed in their initial gravimetric water contents and soil texture (Table S1). This allowed us to test the influence of changes in soil moisture content on the soil metaphenome using soils with different initial water contents, i.e., at time zero (Fig. S2a). Soil from site B had the lowest water content (18%) and the highest clay content (74%). Soils A and C had similar water contents (37% and 31%, respectively), although soil C had higher clay content than soil A. Concentrations of nitrate (<0.06 ppm) and sulfate (2.77 to 5.45 ppm), two of the major anions measured, were low in the three soils. Total carbon-to-nitrogen ratios were comparable between the three soils and ranged from 12 to 18.

We measured soil respiration during the 15-day incubation (Fig. S2b). Carbon dioxide release ($\mu\text{mol CO}_2 \text{ g}^{-1}$ [dry weight] soil) was significantly ($P < 0.05$) higher in response to drying than that in the wet and control treatments in soils B and C. The amounts of CO_2 released at each time point were similar in the wet and control treatments. Although we did not measure soil oxygen levels, the increase in respiration in the dry soils could be indicative of higher rates of oxidation of organic substrates due to higher oxygen availability. Conversely, the lack of difference suggests that the moisture levels in the control soils were sufficiently high that they could not be distinguished from the wet treatments when additional water was added.

Soil microbial community shifts with change in soil moisture. Although we detected both *Bacteria* and *Archaea* in our samples, based on 16S sequencing, the majority of the sequences corresponded to *Bacteria*. We recognize that fungi are also important in C cycling, but we did not target fungi in these experiments. Therefore, the results presented below are focused on *Bacteria*.

By multivariate statistical analysis of the 16S data, we found that soil drying had a significant impact on bacterial β -diversity, as shown on the second axis of the non-metric multidimensional scaling (NMDS) plot, which clearly separates all of the dried samples from the rest (Fig. 1a). In contrast, the wet samples clustered together with the control soils in ordination space, indicating that excess soil moisture had negligible impact on microbial community composition. Samples from the same soil location also grouped together, indicating that soil-specific factors were likely driving community composition.

Across all soil locations, there were significant ($P < 0.01$) shifts in specific taxa in response to drying, with significant decreases in operational taxonomic units (OTUs) corresponding to *Acidobacteria*, *Bacteroidetes*, *Gemmatimonadetes*, *Nitrospirae*, and *Proteobacteria* (Fig. 1b). Taxa that significantly increased after drying included members of the *Actinobacteria*, *Chloroflexi*, *Proteobacteria*, and WS3 candidate phylum. There were fewer significant shifts in taxa following wetting. Taxa that decreased after wetting included members of the *Verrucomicrobia*, *Bacteroidetes*, and *Acidobacteria*, while members of the *Proteobacteria* significantly increased (Fig. 1b). Drying also resulted in significant shifts in α -diversity (Kruskal-Wallis, $P > 0.05$) (Fig. 1c).

At a finer level of taxonomic resolution, the different field locations exhibited site-specific shifts in several OTUs in response to wetting or drying. Figure S3 shows specific OTUs that significantly ($P < 0.01$) increased or decreased in response to wetting or drying for each of the three soil locations. For example, members of the *Xanthomonadaceae*, *Sphingomonadaceae*, and *Chitinophagaceae* families consistently decreased in

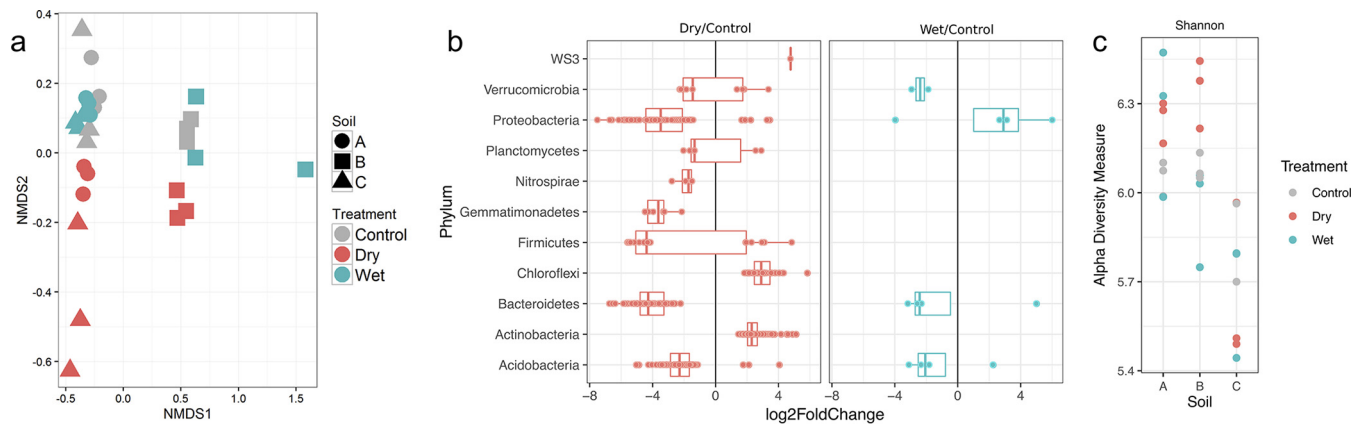


FIG 1 Soil microbiome response to wetting and drying. (a) Nonmetric multidimensional scaling (NMDS) plot of Bray-Curtis dissimilarities showing the microbial community structure estimated by 16S rRNA gene sequencing of soils collected from 3 Kansas prairie field locations. Site (A, circles; B, squares; C, triangles) and treatment (blue, wet; red, dry; gray, ambient field-moisture control). The stress value for the NMDS is 0.076. All sequence count data were normalized to the upper 75th quantile. (b) Differential abundances of OTUs within phyla that were observed to significantly shift in response to drying and wetting relative to continuous moisture control conditions (log₂ fold change, adjusted *P* value < 0.01). (c) Alpha diversity (Shannon's index) of the soil microbiome in soils A, B, and C, for control (gray), dry (red), and wet (blue) treatments.

abundance following drying across all 3 soil locations. These families are comprised of common decomposers. Soil C had few OTUs that increased after drying, whereas soils A and B had increases in members of commonly stress-tolerant *Actinobacteria*, such as *Micromonosporaceae*, *Gaiellaceae*, *Frankiaceae*, *Dolo_23*, and an unidentified family (Fig. S3a). Note that there were no significant shifts in taxa following wetting in soil C (Fig. S3b, data not shown), whereas soils A and B exhibited few and inconsistent shifts (Fig. S3b).

Transcriptional response to changes in soil moisture. We obtained an average of 1,857 transcripts, annotated as Enzyme Commission (EC) numbers, from soils A and C. Unfortunately, we did not obtain sufficient RNA from site B to include that location in the analysis. Ordination of the data revealed that gene expression was significantly influenced by soil moisture changes in both soils A and C (Fig. 2a). The majority, >97%, of the transcripts corresponded to rRNA, but there were sufficient transcripts to also assess shifts in expression of the functional genes. A heat map of the top 20 most

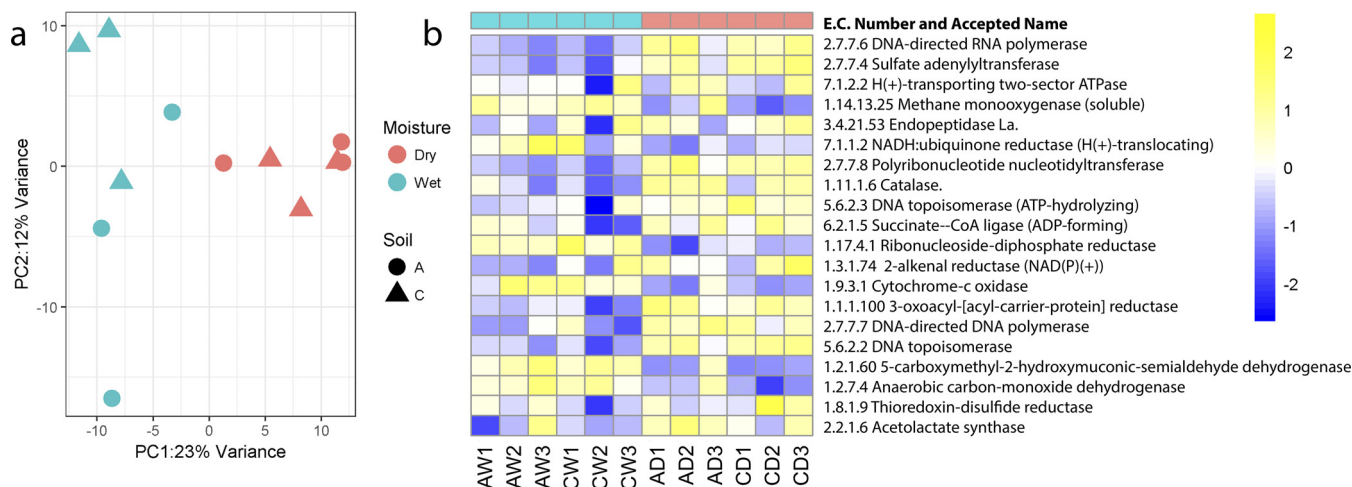


FIG 2 Response of soil metatranscriptome to moisture perturbations. (a) Metatranscriptome data shown as a PCoA ordination of Bray-Curtis dissimilarities of sequence data categorized by function, i.e., Enzyme Commission (EC) number. Site (A, circles; C, triangles) and treatment (blue, wet; red, dry). All sequence count data were normalized to the upper 75th quantile. (b) Heat map showing the top 20 most abundant transcripts (ECs) observed under dry relative to wet conditions. The x axis indicates soil sample (A or C), treatment (W, wet; D, dry), and replicate (1, 2, or 3). Moisture conditions are indicated by the header row in blue (wet) or red (dry). The color gradient for each cell is scaled to a log₂ fold change of -2 to 2.

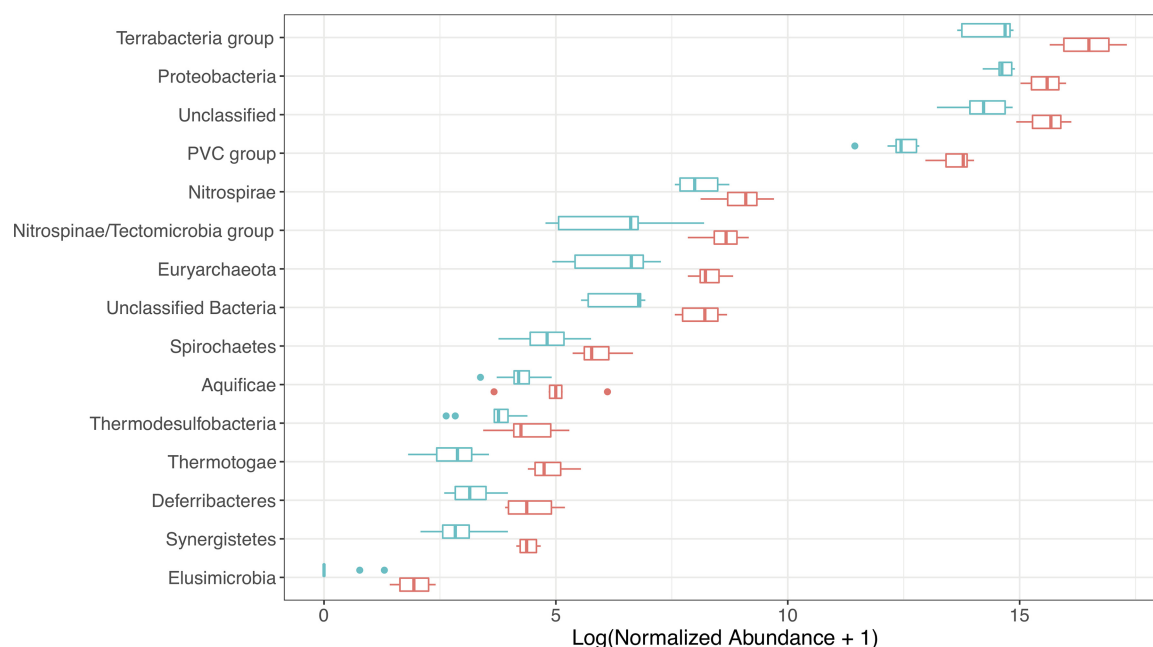


FIG 3 Transcriptional response of different phyla to soil wetting and drying. Log-normalized abundance of active members of the soil microbial community that significantly shifted in abundance in response to drying relative to the wet conditions, as revealed by mapping transcripts to contig-level taxonomies derived from the soil metagenome. All sequence count data were normalized to the upper 75th quantile. Log₁₀ values of the normalized read counts are presented on the x axis, and taxa that shifted significantly in abundance at the phylum level are shown on the y axis. Changes in read count were considered significant for log₂ fold change of >2.0 and P_{adj} of <0.05.

abundant transcripts (ECs) that were grouped according to treatment (Fig. 2b) showed that most of the abundant transcripts were higher under dry than wet conditions. These included several transcripts involved in nucleotide metabolism, including DNA-directed RNA polymerase (EC 2.7.7.6), sulfate adenylyltransferase (EC 2.7.7.4), polyribonucleotide nucleotransferase (EC 2.7.7.8), DNA-directed DNA polymerase (EC 2.7.7.7), and DNA topoisomerase (EC 5.99.1.3). Other abundant transcripts that were higher or lower under wet or dry conditions included several for central and peripheral carbon metabolism pathways. In addition to the most dominant transcripts, we screened for transcripts that significantly differed under dry compared to wet conditions (Table S2). These included several ECs corresponding to enzymes involved in carbon metabolism, and for biosynthesis of secondary metabolites and sugars, indicating that the community was responding by increasing pathways for production of osmolytes. For example, transcript levels for genes involved in trehalose production were significantly higher after drying (Table S2). This included transcripts for genes encoding trehalose phosphatase (EC 3.1.3.12) (adjusted P value [P_{adj}] = 0.033 and 0.046 in soils A and C, respectively) and trehalose synthase (EC 5.4.99.16) (P_{adj} = 0.014 and 0.000 in soils A and C, respectively). Additional transcripts that significantly (P_{adj} < 0.05; see Table S2 for specific P values) increased after drying were the following: for soil A, pyruvate carboxylase (EC 6.4.1.1) and a hydrolase (EC 3.4.24.25), and for soil C, 5-carboxymethyl-2-hydroxymuconic-semialdehyde dehydrogenase (EC 1.2.1.60) and malate dehydrogenase (EC 1.1.1.40). However, other transcripts involved in carbon metabolism were significantly lower (P_{adj} < 0.05) in both soils after drying, including pyruvate dehydrogenase (EC 1.2.2.2), a hydrolase (EC 3.1.4.4), and kojibiose phosphorylase (EC 2.4.1.230).

To determine which members of the community were active under the different soil moisture conditions, all of the metatranscriptomic reads from both soils A and C were mapped to contigs having phylogenetic assignments, and the read abundances were determined for each condition (Fig. 3). Most taxa had higher levels of expression under dry than wet conditions. Representatives of the *Terrabacteria* group, in particular, showed the highest levels of expression under both wet and dry conditions, although

significantly higher under dry conditions ($P_{\text{adj}} < 0.001$). This large group encompasses nearly two-thirds of bacterial species, including *Actinobacteria*, *Firmicutes*, *Cyanobacteria*, *Chloroflexi*, and *Deinococcus-Thermus*.

Shifts in metabolite profiles following changes in soil moisture. Subsequently, we applied an untargeted metabolomics approach to determine the impact of soil wetting and drying on the soil metabolome. A total of 165 metabolites were detected, although only 70 could be identified in reference databases (see Table S3 for the full list of metabolites and P values). Global metabolome data shown as a projection pursuit principal-component analysis (PPCA) of all detected metabolites revealed a separation of the dry soils from the wet and control soils (Fig. 4a). This distinction was more evident by using PPCA of metabolites with significant treatment effects for at least one site (Fig. 4b), where the dry samples clearly clustered separately from the other samples. Fifteen of the 70 detected metabolites changed significantly ($P < 0.05$) under the wet or dry treatments compared to control (Fig. 4c). Soil desiccation resulted in consistently significant increases in several sugars and sugar alcohols across all 3 sampling locations (Fig. S4). In general, the relative abundances of simple sugars and carboxy acids (e.g., glucose, sucrose, fructose, mannose, xylose, trehalose, hydroxybenzoic acid, threonic acid, and toluic acid) were significantly higher after drying compared to the control soil (Fig. S4a). Other metabolites that increased in relative abundance after drying in individual locations included acids, amines, and alcohols (Fig. S4b to d).

Identification of reaction networks responding to soil wetting and drying. To characterize the soil metaphenome, we incorporated metatranscriptomes and metabolites into a metabolic reaction network model (Fig. S5a). Using this method, we identified condition-specific reaction pathways that were activated under the dry and wet conditions, respectively, relative to the control conditions, in soils A and C (Fig. S5b). Based on a hierarchy of reference pathways described in the KEGG database, we classified the identified reactions into metabolite-reaction networks. Finally, we converted the identified metabolite-reaction networks into undirected bipartite graphs (Fig. 5). This graphical representation identified four clusters associated with dry conditions in soil A: (i) “starch and sucrose metabolism”; (ii) “cyanoamino acid metabolism”; (iii) “carbon metabolism, pentose phosphate pathway, and carbon fixation”; and (iv) “glycerophospholipid metabolism” and “ether lipid metabolism.” Metabolite-reaction networks in soil C under dry conditions also included “starch and sucrose metabolism” and “carbon metabolism, pentose phosphate pathway, and carbon fixation.” In addition, “fatty acid metabolism” and “chloroalkane and chloroalkene degradation” were uniquely found in soil C under dry conditions and wet conditions, respectively. In contrast, no such distinct clusters were observed in the wet samples from soil A. In both soils, we predicted reactions that were uniquely associated with wet or dry soils in both soils A and C (Fig. 5c). Interestingly, these reactions were for synthesis of trehalose in dry soils and trehalose degradation in wet soils.

Thus, our approach to integrate the empirical observations from the soil metatranscriptome and metabolome into metabolic networks was in alignment with findings from the individual data sets and also highlighted specific reaction modules of potential significance under soil drought for further targeted investigations, e.g., cyanoamino acid metabolism observed only in soil A under dry conditions and chloroalkane and chloroalkene degradation observed only in soil C under wet conditions and trehalose metabolism in both soils.

The activated and enriched reaction modules for each soil-moisture treatment combination are presented individually, i.e., soil A-dry, soil A-wet, soil C-dry, and soil C-wet (Fig. S5b). Reaction module distributions between the two wet conditions in soils A and C showed a high correlation (i.e., correlation coefficient of 0.71), which implies that wet biochemistry was not very different across the two soil sites. However, there was a low correlation between reaction module distributions from the two soils after drying (i.e., correlation coefficient of 0.56), which implies that the effect of drying may be different depending on soil type or location. Conversely, the within-soil correlation

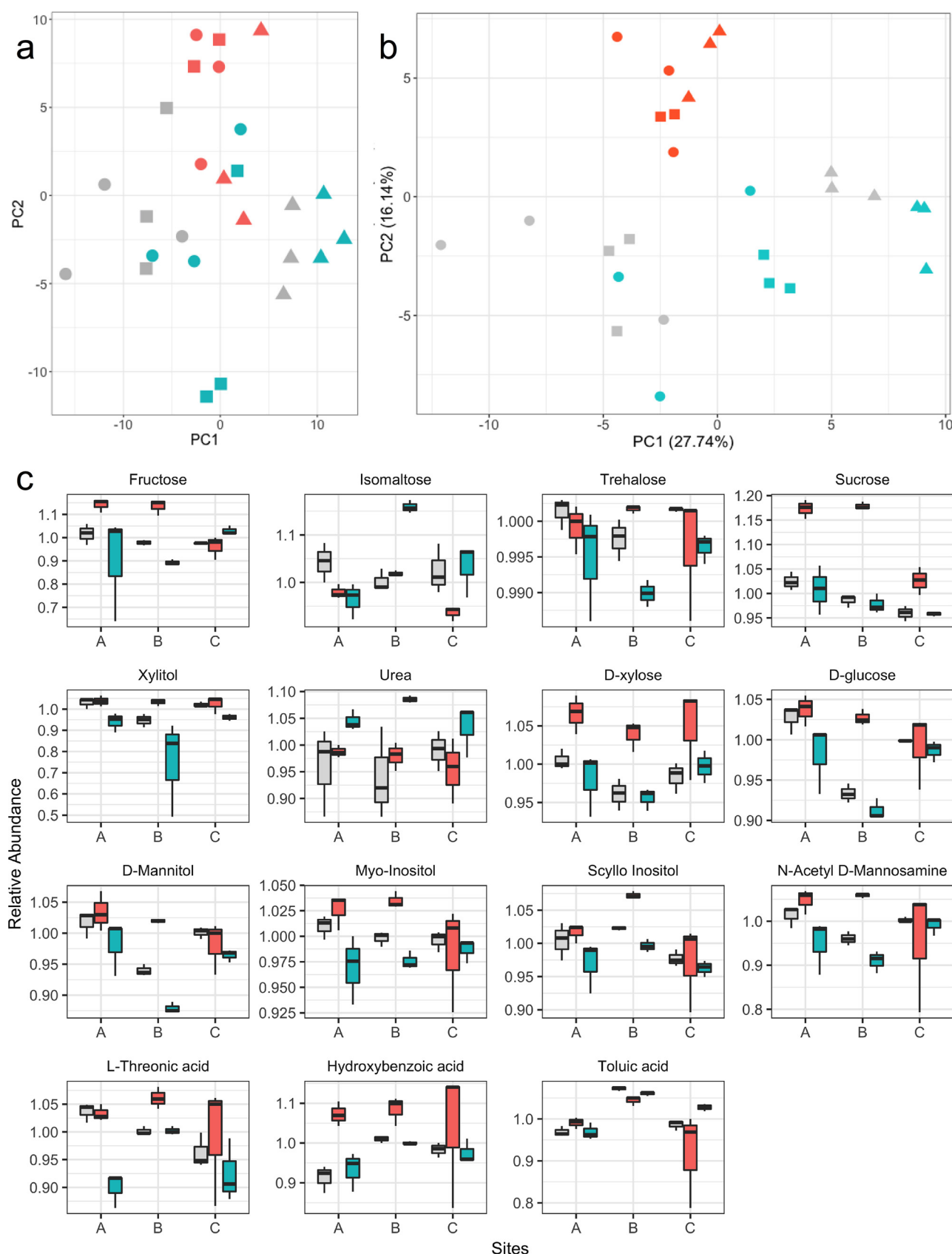


FIG 4 Impact of soil moisture treatments on the soil metabolome. (a) Global metabolome data shown as a projection pursuit principal-component analysis (PPCA) of all detected metabolites. Relative abundance data for metabolites were $\log_2$ transformed and median centered. The % on the axes (Continued on next page)

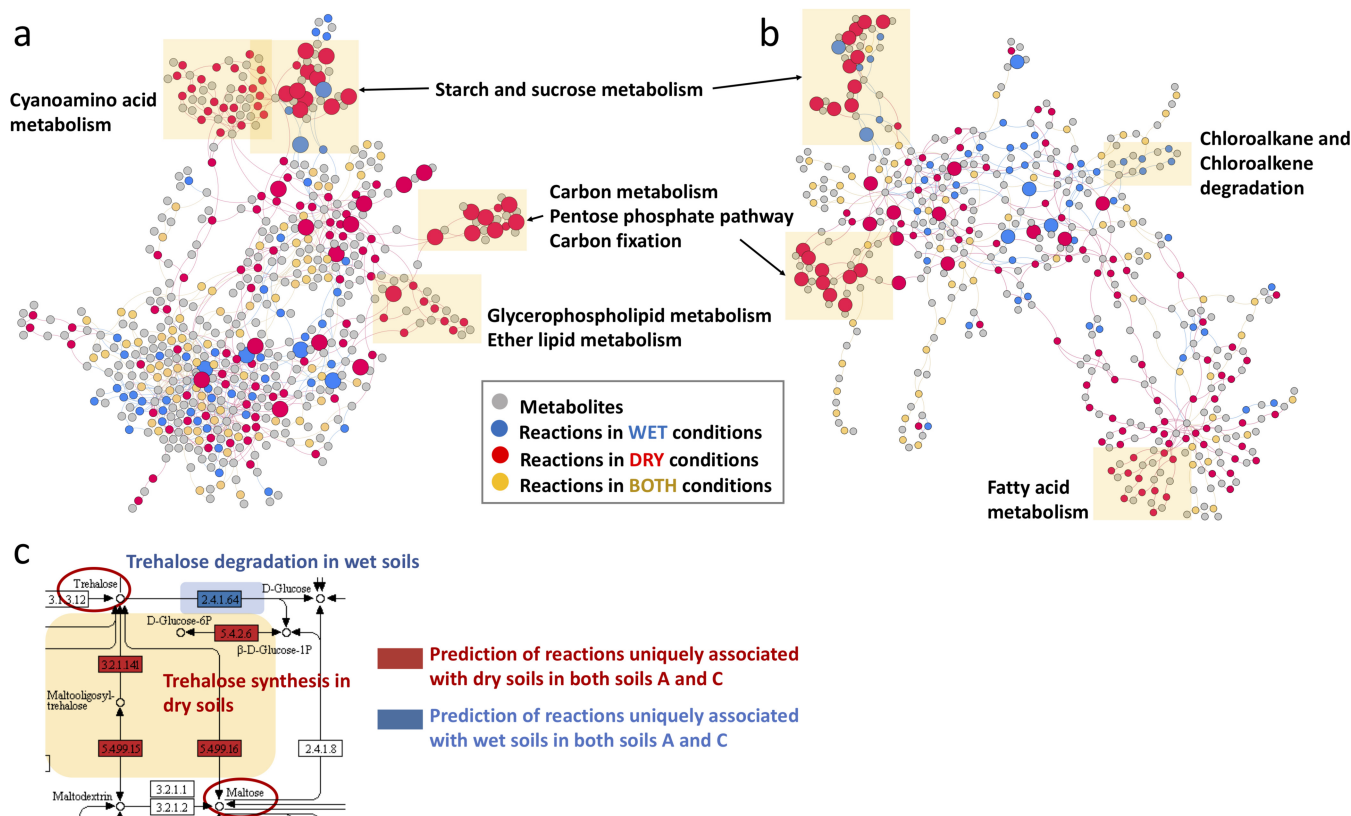


FIG 5 Prediction by MEMPIS of the moisture impact on biochemistry in native prairie soil. (a and b) Reaction-metabolite integrative bipartite networks for soils A (a) and C (b) are shown. Gray symbols indicate metabolites, and lines indicate relationships between reactions or metabolites based on KEGG annotation. Colored symbols indicate reactions that are more prevalent under specific incubation conditions: blue, wet; red, dry; yellow, both. Yellow shading highlights specific pathways that are more prevalent under dry conditions. Larger nodes represent the commonly predicted reactions in both soils A and C, while small nodes represent the uniquely predicted reactions in the individual soil. (c) Predicted reactions for trehalose synthesis and degradation in the starch and sucrose metabolism pathways are shown. Colored boxes indicate the predicted reactions uniquely associated with dry and wet conditions in both soils A and C.

between wet and dry treatments in soil A was 0.5 and in soil C was 0.62, also indicating a soil-specific response to such moisture shifts, such as cyanoanimo acid metabolism that was found only in soil A under dry conditions as mentioned above. These findings are counter to our overarching hypothesis that there would be similar metaphenomic responses across field locations to changes in soil moisture, other than those found for trehalose metabolism (Fig. 5c).

DISCUSSION

We developed an untargeted omics approach to determine the soil microbial community’s metaphenomic response to soil moisture and to define specific metabolic signatures of the response. By incorporating both metatranscriptomics and metabolomics into the same metabolic reaction networks, we aimed to identify reaction pathways distinctively activated in response to the moisture perturbations and in an ecosystem-specific manner. We refer to these reaction pathways as being “context-specific” reaction modules, i.e., soil- and treatment-specific groups of reactions that are functionally related. Using this approach, we demonstrate that despite the high com-

FIG 4 Legend (Continued)
represents the variance explained by each of the coordinates. Site (A, circles; B, triangles; C, squares) and treatment (gray, control; blue, wet; red, dry). (b) Metabolite data shown as a projection pursuit principal-component analysis (PPCA) of metabolites with significant treatment effects for at least one site. (c) Fifteen of 70 detected metabolites that changed significantly under the wet or dry treatments compared to control ($P < 0.05$, one-way ANOVA). Whiskers indicate the most extreme values within 1.5 multiplied by the interquartile region. Box, 25% quartile; median, 75% quartile. Pairwise comparisons of means to test treatment effects were performed after outlier removal.

plexity of the soil habitat, it is possible to generate an integrated understanding of the effect of environmental change on the soil microbiome and the soil metaphenome.

Our field sampling across the Konza native prairie site took advantage of natural variability across the site, including locations that were naturally wetter or drier (see Fig. S1 and Table S1 in the supplemental material). Although the soil microbial community composition profiles, gene expression patterns, and metabolites shifted upon drying across all soil locations, we observed minor differences between wet and control soils. These results suggest that the soil microorganisms at our sampling sites were more adapted to wet conditions at the times of sampling. Our data also suggested that soil location was the primary driver and that change in soil moisture was a secondary driver of soil microbial composition and function. In contrast, in a recent study, we found that samples collected from the Konza prairie site at a maximum distance of 10 m from each other clustered together and separately from adjacent cultivated soils (21). In the same study, we also found that prairie soils from different states (Iowa and Wisconsin) exhibited similar soil microbiome structures compared to adjacent cultivated soils. Therefore, we proposed the concept of a prairie microbiome that was relatively consistent across locations (21). Previous studies that experimentally manipulated durations of drought and rainfall events in the Konza prairie offered a limited explanation of the variation in soil respiration due to moisture and implicated the whole-community microbial response as a major driver (22). Here, we find, however, the need to account for field-scale heterogeneity, in particular that of hydrology, across a site in order to better understand how the microbiome is impacted by changes in soil moisture.

Our results showed that soil drying, in particular, had significant impacts on all omics levels that we investigated, from community composition to gene expression and metabolite profiles. Although the Konza LTER site had not experienced a recent drought at the time of our sampling, legacy effects and microbial dormancy might explain some of our results. The central grassland regions experience drought every 20 to 30 years, and the last drought was recorded in the 2000s based on published data (23). Our short-term (15-day) drying experiment was designed to reveal the initial response of the soil microbiome to desiccation. However, longer-term experiments are needed to predict the metaphenomic response to periods of drought that can last months to years.

Increasing the soil moisture contents to saturation had little impact on the soil microbiome. Therefore, drying may have a more pronounced impact over a shorter time frame than wetting in the soils we studied. This could be due to negligible changes in abiotic conditions between the wet control and water-logged treatments. Alternatively, the lack of pronounced differences upon soil saturation relative to the control could be attributed to a soil microbiome acclimatized to higher antecedent soil moisture status due to frequent seasonal precipitation events. Indeed, soils A and C were located in low-lying regions of the landscape (Fig. S1) with relatively wet (~30%) *in situ* soil moisture observed at the time of collection. In contrast, soil B had a lower gravimetric water content and was harder to wet up, presumably due to the higher clay content that would slow water permeation. Previous studies from sites adjacent to our field location found that when the soil microbiome was preadapted to heavy rainfall, there was a negligible response to laboratory saturation (14). This suggests that the historical soil moisture regime may impact current rates of microbial processes (e.g., respiration activity) by aggregating community-level traits that control carbon use efficiency, soil moisture sensitivity, and stress tolerance (24).

Our results corroborate other research that has reported relative decreases in the Gram-negative phyla *Proteobacteria*, *Verrucomicrobia*, and *Bacteroidetes* (12, 25–27) and increases in the Gram-positive phyla *Firmicutes* and *Actinobacteria* (25, 28–30) after soil drying. These shifts in relative abundance at the phylum level are driven by increases or decreases in abundances of specific groups within each respective phylum (12, 30), depending on their respective metabolic capacity to respond to the change in environmental conditions (27, 31). However, most of these phylum-level distinctions in

relative abundance were not mirrored at the transcriptional level. Instead, drying increased the transcriptional activity across different phyla. We found that members of the *Terrabacteria* group and *Proteobacteria* were the most transcriptionally active in general and in response to soil drying. The *Terrabacteria* group consists of the Gram-positive *Actinobacteria* and *Firmicutes*, in addition to *Chloroflexi* and *Cyanobacteria* (32), several species of which are known to be desiccation and heat tolerant and have been reported under warmer (33) and drier (34) soil conditions. Additionally, several genera within the *Actinobacteria* are known to be capable of survival in arid soils (27, 30, 35), which supports our findings.

We performed an integrative analysis of the multi-omics data to generate metabolic networks activated in response to the soil moisture perturbations. The metabolite data were the most discriminating data set with the highest number of significant features that were consistently significantly different across the three sampling locations. In contrast, the metatranscriptome data were highly variable between sampling sites. One of the limitations of this study was that we obtained high-quality RNA from only 2 of the 3 field locations. Therefore, the significance of our metatranscriptome data was often difficult to assess, particularly since there were many opposite trends in the community expression data when comparing the two sites. This could be due to the transient nature of the transcripts that are highly dependent on *in situ* environmental variables that govern gene expression, as well as the relatively low depth of sequence coverage of the transcripts. The metabolite data were also limited because of the high number of unknown metabolites in the samples.

When we used our data-integrated metabolic network modeling approach, it was, however, possible to reveal specific pathways that were expressed in the soil microbial community as a response to moisture perturbation. In contrast to discrete data-reaction mapping, this framework extended the series of localized snapshots of information to a global metabolic reaction network scale, enabling the identification of context-specific reaction modules. Indeed, distributions of reaction modules among dry/wet soil A and dry/wet soil C were shown to be fundamentally different (Fig. S5). Importantly, beyond simple quantification of the cross-condition differences in biochemistry, our metabolic network analysis pointed out specific reactions that shifted in response to change in soil moisture and provides the basis for formulating new hypotheses to test for validation of model predictions. For example, we found that transcripts for cyanoamino acid metabolism were more prevalent after drying in soil A. Although the significance of cyanoamino acid accumulation is not known for soil microbes, transcription of cyanoamino acids has been shown to increase during salt stress in desert poplar trees (64), suggesting that these atypical amino acids are protective under desiccation conditions. This is a hypothesis to test in future experiments.

In summary, we were able to extract information about the specific physiological responses of the soil microbiome and metapenome to moisture perturbations. We were able to capture a contingent of metabolic mechanisms that the soil bacterial community uses in response to water deficit such as alteration of metabolism, osmolyte biosynthesis, and transcription control. The integration of multiple omics technologies with metabolic pathway modeling now provides a richer knowledge base for better understanding of the soil metapenome.

MATERIALS AND METHODS

Site description and sample collection. Soil samples were collected at the Konza Prairie Biological Station (KPBS), a Long-Term Ecological Research (LTER) site representative of native tallgrass prairie in the Flint Hills of eastern Kansas, USA, from three field locations: site A, 39°06'11"N, 96°36'48"W, 339 meters above sea level (MAMS); site B, 39°04'39"N, 96°36'29"W, 413 MAMS; and site C, 39°04'20"N, 96°34'33"W, 415 MAMS (see Fig. S1 in the supplemental material). These sampling locations or sites remain undisturbed by agriculture and are dominated by perennial C4 grasses characteristic of native lowland tallgrass prairies in this region. Our selected sampling sites were burned annually (sites A and C) or at 4-year intervals (site B). The Flint Hills are generally characterized by shallow soils overlaying chert-bearing limestones and shales (36), and the site is classified as typical chernozem according to the Food and Agriculture Organizations (FAO) soils classification used by the United Nations and belongs to the Mollisol order of the U.S. Soil Taxonomy. The study soils have been previously reported as low-lying Irwin

silty clay loams (fine, mixed mesic Pachic Argiustolls, USDA Natural Resources Conservation Service (NRCS) (11). The three soils used in this study represented locations across a natural hydrology gradient such that soils A and C were from the wetter areas and soil B was from a relatively drier area of the landscape (Fig. S1) and located adjacent to the Rainfall Manipulation Plots (RaMPs) at this LTER site (37, 38).

At each of the three sites, soil samples were obtained from three locations at least 10 m apart at 0- to 15-cm depth with a shovel. Roots and large rocks were removed; the soil was manually mixed, first in a 5-gal bucket and then in 1-gal Ziploc bags; and soil was composited into one bag for each site. Soils were immediately frozen under liquid nitrogen and shipped to the Pacific Northwest National Laboratory (PNNL), Richland, WA. Upon receipt at PNNL, the soils were quickly sieved through a 2-mm sieve and 50-g portions of the soils were aliquoted to 50-ml Falcon tubes and flash-frozen at -80°C to serve as baseline samples and replicates for the experiments described below.

Experimental design and laboratory incubation. We experimentally wetted and dried the soils collected from three field locations using a simple experimental design to determine the effects of extreme moisture shifts on the soil microbiome structure and function. Previously processed (as described above) soils were thawed and preincubated at 21°C for 48 h before the onset of treatments. After temperature equilibration, triplicate microcosms (20 g soil in 50-ml glass beakers) were subjected to three moisture conditions: wetted to saturation (addition of sterile deionized water), air dried to constant weight by evaporation, and maintained at ambient field-moist conditions. The 27 microcosms (3 sites $\times$ 3 treatments $\times$ 3 replicates) were maintained gravimetrically and incubated at 21°C in a biosafety cabinet (open to the atmosphere inside) for 15 days (Fig. S2b). For the wet treatment, water was added to the 3 soils such that the endpoint moisture contents for the saturation treatments allowed for excess (not inundated) unabsorbed water in the microcosms and were in the range of that reported for *in situ* delayed rainfall treatments at Konza Biological Station (14). We used percent water-filled pore spaces (% WFPS) to estimate water content and aeration status (39) using equations 1 and 2:

$$\%WFPS = \frac{\text{gravimetric water content} \times Db}{\text{total soil porosity}} \quad (1)$$

$$\text{soil porosity} = 1 - \left(\frac{Db}{Dp} \right) \quad (2)$$

where average soil bulk density (Db) = 1.1 g cm^{-3} and assumed particle density (Dp) = 2.65 g cm^{-3} .

The final WFPS ranged from $\sim 56\%$ to 84% in the saturation treatments and from $\sim 22\%$ to 47% in the dry treatments compared to $\sim 33\%$ to 68% in the control soils (Table S1). At the end of the experiment, subsamples from each replicate microcosm were collected for nucleic acid and metabolite extractions as described below. We refer to the saturated samples as “wet,” the air-dried samples as “dry,” and the field-moist soils as “control” where the weights were kept constant by addition of sterile water throughout the incubation period (Fig. S2b).

Soil respiration from the wet, dry, and control incubations was measured at regular intervals of 2 to 3 days using a G2301 Picarro GHG analyzer (Picarro, Sunnyvale, CA, USA) (Fig. S2b). Headspace gases of each sample were measured for 2 h, and fluxes were calculated from the change in concentration using the Ideal Gas Law and equation 3:

$$F = \left(\frac{dC}{dT} \times \frac{V}{M} \times \frac{P}{RT} \right) \quad (3)$$

where F is the flux ($\mu\text{mol g soil}^{-1} \text{ s}^{-1}$), dC/dT is the rate of change in CO_2 concentration (mol fraction s^{-1}), V is the total volume (cm^3) of microcosm and chamber, M is the dry weight of soil (g), P is the atmospheric pressure (kPa), R is the universal gas constant ($8.3 \times 10^3 \text{ cm}^3 \text{ kPa mol}^{-1} \text{ K}^{-1}$), and T is the temperature of incubation (K).

Nucleic acid extraction. DNA was extracted from 0.25 g (wet weight) soil using the PowerSoil DNA isolation kit (Mo Bio Laboratories Inc., Carlsbad, CA) according to the manufacturer’s instructions, and extracts were quantified using the Qubit Fluorometer 2.0 (Invitrogen, Carlsbad, CA, USA) and checked for quality using the NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA) and Bioanalyzer HS DNA chips (Agilent Technologies, Santa Clara, CA).

RNA was extracted from the soil samples (2 g) using the PowerSoil RNA isolation kit (Mo Bio Laboratories Inc., Carlsbad, CA) according to the manufacturer’s instructions. RNA was DNase treated with Turbo DNase (Life Technologies, Grand Island, NY, USA) at 37°C for 20 min. RNA was subsequently purified by phenol-chloroform extraction and recovered by precipitation with isopropanol. RNA was quantified using the Qubit Fluorometer 2.0 (Invitrogen, Carlsbad, CA, USA) and quality checked using a NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA) and Bioanalyzer Pico RNA chips (Agilent Technologies, Santa Clara, CA). First-strand synthesis of cDNA from RNA was carried out using the SuperScript VILO cDNA synthesis kit (Life Technologies) according to the manufacturer’s protocol using a 1:1 blend of $1 \mu\text{M}$ random hexamers/decamers. Second-strand synthesis was performed using the NEBNext mRNA second-strand synthesis module and T4 gene 32 (New England Biolabs, Ipswich, MA) according to the manufacturer’s protocol at 16°C for 2 h. cDNA was purified using Agencourt AMPure XP beads (Beckman Coulter, Beverly, MA). The resulting cDNA was quantified and quality checked using the Agilent RNA 6000 Pico kit on a 2100 Bioanalyzer instrument (Agilent Technologies, Santa Clara, CA).

For Illumina library construction, DNA and cDNA were sheared to approximately 450 to 500 bp by using a Covaris M220 focused ultrasonicator instrument (Covaris M220 series; Woburn, MA). Sheared DNA fragments were end repaired, A tailed, and ligated to TruSeq adapters from Illumina according to the manufacturer’s instructions (Illumina, San Diego, CA). Adapter dimers were removed twice using Agen-

court AMPure XP magnetic beads (Beckman Coulter, Danvers, MA). Libraries were checked for size and adapter dimers using a high-sensitivity DNA chip on a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) and quantified by qPCR on a StepOne real-time PCR system (Thermo Fisher Scientific) using the KAPA library quantification kit (KAPA Biosystems, Wilmington, MA) according to the manufacturer's instructions. Fragments were sequenced using the Illumina MiSeq and HiSeq sequencing platforms (3 lanes) using the V3 chemistry (Illumina, San Diego, CA) for 300 cycles, generating 2×250 -bp paired-end reads and ~ 10 million reads per sample.

For the metatranscriptome data analysis from the moisture perturbation experiment, we focused on soils A and C, because of difficulties in obtaining sufficient high-quality RNA from site B. We were also able to obtain high-quality RNA and sequence the metatranscriptomes from only one of the 3 replicates from the soil A and C controls.

16S rRNA gene sequencing and analysis. PCR amplification of the V4 region of the 16S rRNA gene was performed using the protocol developed by the Earth Microbiome Project (40) and as previously described (41), with the exception that the 12-base barcode sequence was included in the forward primer. Amplicons were sequenced on an Illumina MiSeq using the 500-cycle MiSeq reagent kit v2 (Illumina, San Diego, CA) according to the manufacturer's instructions.

Raw sequence reads were demultiplexed with EA-Utils (42) with zero mismatches allowed in the barcode sequence. Reads were quality checked with BBDuk2 (43) to remove adapter sequences and PhiX with matching kmer length of 31 bp at a hamming distance of 1. Reads were merged using USEARCH (44) with a minimum length threshold of 175 bp and maximum error rate of 1%. Sequences were dereplicated (minimum sequence abundance of 2) and clustered using the distance-based, greedy clustering method of USEARCH (44) at 97% pairwise sequence identity among operational taxonomic unit (OTU) member sequences. *De novo* prediction of chimeric sequences was performed using USEARCH during clustering. Taxonomy was assigned to OTU sequences at a minimum identity cutoff fraction of 0.8 using the global alignment method implemented in USEARCH across RDP Trainset database version 15 trained with UTX 250-bp configuration (45). OTU seed sequences were filtered against RDP Gold reference database version 9 to identify chimeric OTUs using USEARCH.

Metagenome and metatranscriptome library preparation and sequence analysis. DNA extracted from the technical replicates for each soil location prior to perturbations (nine samples in total) was pooled to construct a single long-hybrid Molecule subassembly according to the manufacturer's protocol (Illumina, San Diego, CA) and as previously described (46, 47). The data were assembled on BaseSpace using the Illumina TruSeq Long-Read Assembly application v1.0. Molecule long hybrid read subassembly, annotation, mapping, and hybrid assembly have been previously described (47).

Metagenomic and metatranscriptomic libraries were constructed from technical replicates for treatment and control conditions, as applicable for the respective perturbation experiments. Libraries were prepared using 500 ng of DNA and 150 ng of cDNA. The DNA and cDNA were sheared to approximately 450 to 500 bp using a Covaris M220 focused-ultrasonicator instrument (Covaris M220 series; Woburn, MA). Sheared DNA fragments were end repaired, A tailed, and ligated to TruSeq adapters from Illumina according to the manufacturer's instructions (Illumina, San Diego, CA). Adapter dimers were removed twice using Agencourt AMPure XP magnetic beads (Beckman Coulter, Danvers, MA). Libraries were checked for size and adapter dimers using a high-sensitivity DNA chip on a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) and quantified by qPCR (StepOne Plus system; Applied Biosystems, Foster City, CA) using the KAPA library quantification kit (KAPA Biosystems, Wilmington, MA) according to the manufacturer's instructions. Fragments were sequenced on the HiSeq sequencing platform (Illumina, San Diego, CA) with 3 runs of 250-bp paired-end reads that generated ~ 10 million reads per sample.

Replicate metatranscriptomes were quality controlled, screened for phiX, and mapped to the Molecule subassembly using Bowtie 2 (48). Genes in the assembly were annotated with function and taxonomic rank using MetaPathways v.2.5 (49). Read counts per function for each condition were generated using Kyoto Encyclopedia of Genes and Genomes Orthology (KO) and Enzyme Commission (EC) numbers for both MTs and MGs using ATLAS (<https://github.com/metagenome-atlas/atlas>). The KO and EC count matrices were used for analyses of MG and MT functional data sets.

The aligned read abundances from the metatranscriptomes from the different treatments were used to determine which members of the community were active under the different conditions. The metatranscriptomes for each moisture condition from soils A and C were mapped to metagenomic contigs, containing 810,853 open reading frames, which were assembled from existing metagenome data from the same original soil batch (40). All metatranscriptomic reads in all samples were able to be aligned to the assembly. The average number of metatranscriptomic reads per sample that mapped to open reading frames on the assembly was 116,316.

Metabolite extraction, analysis, and data processing. We conducted an untargeted analysis of polar metabolites in soil using a direct extraction method (50). Briefly, ~ 5 g wet soil was aliquoted into 50-ml screw-cap self-standing tubes (Next Advance, Averill Park, NY) containing 0.9- to 2.0-mm stainless steel beads, 0.1-mm zirconia beads, and 0.1-mm garnet beads and 8 ml of 60% (vol/vol) methanol in Nanopure water to effectively quench metabolic reactions upon lysis (51). Samples were bead beaten in a 50-ml Bullet Blender (Next Advance, Averill Park, NY) at speed 12 for 15 min at 4°C and subsequently transferred to chemical-compatible polypropylene 50-ml tubes (Olympus Plastics, Waltham, MA). Twelve milliliters of ice-cold chloroform was added, and each sample was probe sonicated at 60% amplitude (in a fume hood) for 30 s on ice, allowed to cool on ice, and sonicated again. The samples were allowed to completely cool at -80°C for ~ 10 min and then centrifuged at $4,500 \times g$ for 10 min at 4°C to separate the aqueous phase. The upper aqueous phases containing polar metabolites were transferred into glass vials using serological pipettes, dried down completely in a vacuum concentrator (Labconco, Kansas City,

MO), and stored at -20°C until chemical derivatization. Details of chemical derivatization and gas chromatography-mass spectrometry (GC-MS) analyses have been previously published (52).

GC-MS raw data files were processed using MetaboliteDetector (53). Retention indices (RIs) of metabolites were calculated based on analysis of a mixture of fatty acid methyl ester (FAME) standards (C_8 to C_{28}), followed by their chromatographic alignment across all analyses after deconvolution. Metabolites were then identified by matching GC-MS features (characterized by measured RIs and mass spectra) to the Agilent Fiehn Metabolomics Retention Time Locked (RTL) library (54). All metabolite identifications were manually validated to reduce deconvolution errors during automated data processing and to eliminate false identifications. The relative abundance of metabolites was calculated based on the peak integration of three selected mass fragments from a single metabolite peak, and the information for three quantification ions was used consistently for all the samples in the study.

A total of 165 metabolites were measured by GC-MS from the 27 samples, including biological replicates from the three sites A, B, and C for the three treatments, wet, dry, and control. Of these, 70 could be identified using the reference databases. Evaluation for outlier behavior resulted in the removal of one replicate from soil C dry treatment. A one-factor analysis of variance (ANOVA) model with main effect of treatment was performed for each metabolite and site. *Post hoc* tests for three specific differences were run: wet versus dry, wet versus control, and dry versus control.

Construction of metabolic network model. We constructed the metabolic network model by assembling all reactions downloaded from the KEGG database (termed the master network). While the metabolic network *per se* is generic, the incorporation of context-specific (i.e., soil site- and moisture-dependent) multi-omics data led to the identification of sets of reactions that were differentially activated under the wet and dry conditions compared to the control. In comparison to stand-alone data analysis performed above, this integrative approach enabled illustration of the key biochemical changes based on an expanded set of reactions by determining the most probable paths that connect the measured metabolites and transcripts (see Fig. S5). Using this method, we identified 209 and 129 activated reactions derived from the 25 and 7 overexpressed transcripts under the dry and wet conditions, respectively, compared to the control conditions in soil A. For soil C, we identified 194 and 135 activated reactions using 13 and 5 genes overexpressed under dry and wet conditions, respectively. Soil B could not be part of this analysis due to the lack of metatranscriptomic data. Based on a hierarchy of reference pathways described in the KEGG database, we classified the identified reactions into reaction module levels (also see the supplemental material).

Statistical analyses. Estimates of α -diversity (observed, Chao 1, and inverse Simpson) based on 16S data were obtained from nonrarefied data. β -Diversity analysis was based on upper 75th-quartile normalized library counts. Nonmetric multidimensional scaling or principal-coordinate analysis (PCoA) plotting was used to visualize β -diversity based on unweighted UniFrac or Bray-Curtis distances using functions supported by the package Phyloseq (55).

Count data generated from sequence analyses for MGs and MTs were screened for significant differences between treatments using a negative binomial frequency distribution model, as implemented in the package DESeq2 (56) and based on upper 75th-quartile normalized library counts. The most abundant transcripts were visualized as a heat map using a hierarchical clustering method implemented in the R package pheatmap. Differences in the abundance of taxa and the variance in expression of transcripts were characterized using the DESeq2 parameters `fitType = "local"` and an adjusted *P* value threshold of 0.05 to calculate $\log_2$ fold changes between treatments. The significant ($P < 0.05$) KOs from the MT data were subjected to pathway enrichment analyses using the package GAGE (57), and pathways were visualized using the package Pathview (58). Analysis was performed using the R programming language (v. 3.3.2) (59).

Abundance data for metabolites were $\log_2(x)$ transformed and median centered. Data were processed on the $\log_2$ scale, and missing values were not imputed. Data were evaluated for outlier behavior using rMd-PAV (60), a robust principal-component analysis (PCA)-based approach, and any outliers were removed. A two-factor analysis of variance (ANOVA) model, with main effects of treatment and site, as well as an interaction of the two main effects, was performed for each metabolite to simultaneously evaluate the statistical hypothesis that there is equal metabolite abundance by treatment and site. The inclusion of an interaction term enabled a test of equal metabolite abundances for treatment within each site and the evaluation of different treatment effects (directions) across sites. Further, a PCA approach that allows missing values, projection pursuit PCA (61), was performed on metabolites that showed a significant treatment effect for at least one site. A total of 100 metabolites were used in computing the principal coordinates using methods described above after accounting for missing values using the abovementioned PPCA approach. Differences in metabolite relative abundances between the treatment groups were tested by Kruskal-Wallis test at the 0.05 level of significance. A *post hoc* pairwise comparison was performed using Bonferroni's test. Additional description of the metabolite data analysis can be found in the supplemental material. A complete list of compounds identified in this study is provided in Table S3.

Modeling approach for identifying active reaction modules under dry and wet conditions. We identified distinct sets of reactions, i.e., reaction modules that were activated under the soil dry and wet conditions, by incorporating metabolite and transcript measurements into a metabolic network model. First, we constructed a metabolic network by collecting all reactions and metabolites obtained from the KEGG database (release date 28 May 2018). The resulting network included a total of 10,826 reactions and 8,398 compounds. We then converted reaction stoichiometry to binary representation so that the network model provides a complete description of connectivity between metabolites and reactions.

The network constructed as described above can be considered a generic model, which is applicable across treatment conditions. In order to identify the condition-specific (soil and moisture status combinations) activated pathways, we constrained the generic metabolic network to the metabolite and metatranscriptomic data collected from the control, dry, and wet conditions, respectively, for soils A and C. For measured metabolites, we forced the model to activate at least one of the associated reactions (involved in either their production or consumption). We further constrained the model by incorporating overexpressed genes under dry and wet conditions in comparison to control. For this purpose, we first identified the differentially expressed transcripts, i.e., EC numbers (dry versus control; wet versus control), using DESeq2 (56). We selected genes based on two criteria: (i) differential gene expressions that were statistically significant (i.e., false discovery rate [FDR] < 0.1) and (ii) expression levels under dry and wet conditions that were two times higher than those in control soils (i.e., $\log_2 FC > 1$). For multiple testing adjustment, the Benjamini-Hochberg (BH) (62) adjustment method was employed. For soil A, we identified 25 and 7 overexpressed genes for dry and wet conditions versus its control, resulting in 209 and 129 activated reactions, respectively. Similarly, we identified 13 and 5 genes overexpressed under dry and wet conditions for soil C, resulting in 194 and 135 activated reactions, respectively. We forced the model to activate the reactions associated with highly expressed genes thus identified.

Under the constraints described above, we formulated a Mixed Integer Linear Programming (MILP) problem to obtain minimal pathways or subnetworks that connect all metabolites and reactions incorporated (as constraints) into the model. We termed our algorithm “Metabolite-Expression-Metabolic Network Integration for Pathway Identification and Selection” (MEMPIS). We solved the formulated minimization problem using the IBM ILOG CPLEX version 12.7.1 optimization package.

For visualizing a set of predicted reactions and metabolites, after extracting the corresponding subnetwork from the generic model, the subnetwork was converted to an indirect bipartite graph with two types of nodes: reactions and compounds. We used Python package NetworkX and Gephi for graph manipulation and network visualization. Also, we used iPath3 (63) to visualize and compare the activated reactions between dry and wet conditions on KEGG global maps summarizing the comprehensive KEGG pathway maps and KEGG modules.

Availability of data. All data are made publicly available at the Open Source Framework (OSFHOME) and can be retrieved from <https://osf.io/4uvj7/>.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mSystems.00061-19>.

FIG S1, JPG file, 0.6 MB.

FIG S2, JPG file, 0.3 MB.

FIG S3, JPG file, 0.3 MB.

FIG S4, DOCX file, 0.4 MB.

FIG S5, JPG file, 0.2 MB.

TABLE S1, DOCX file, 0.01 MB.

TABLE S2, DOCX file, 0.02 MB.

TABLE S3, DOCX file, 0.02 MB.

ACKNOWLEDGMENTS

This research was supported by the Department of Energy Office of Biological and Environmental Research (BER) and is a contribution of the Scientific Focus Area “Phenotypic response of the soil microbiome to environmental perturbations.” This research was also supported by Environmental Molecular Sciences Laboratory (EMSL) open call proposal number 48784, Dissecting and Deciphering of the Soil Microbiome. EMSL is a U.S. Department of Energy (DOE) Office of Science user facility sponsored by the DOE’s Office of Biological and Environmental Research and located at PNNL. A portion of the research was conducted using PNNL Institutional Computing (PIC) resources and partially supported by the Microbiomes in Transition Initiative under the Laboratory Directed Research and Development Program at PNNL. PNNL is a multiprogram national laboratory operated by Battelle for the U.S. DOE under contract DE-AC05-76RL01830.

We thank visiting research scientist Thulani P. Makhanyane (University of Pretoria, Pretoria, South Africa) for assistance with soil processing; Carrie D. Nicora and Erica M. Zink, both at PNNL, for assistance with the metabolite analysis; and Ruonan Wu at PNNL for help with analysis of the transcript data.

T.R.C. and J.K.J. conceived the project and wrote the manuscript. T.R.C., E.M.B., and S.J.F. performed experiments. J.B., R.A.W. III, J.D.Z., and C.J.B. performed sequence data analysis. T.O.M. and Y.-M.K. performed metabolomics analysis. J.-Y.L. and H.-S.S. devel-

oped a computational algorithm to identify context-specific reaction modules and provided critical inputs for the manuscript. L.M.B. and T.R.C. performed statistical analysis of the data.

The authors declare no competing financial interests. The authors declare no conflicts of interest.

REFERENCES

- Cook BI, Seager R, Miller RL. 2011. Atmospheric circulation anomalies during two persistent North American droughts: 1932–1939 and 1948–1957. *Clim Dyn* 36:2339–2355. <https://doi.org/10.1007/s00382-010-0807-1>.
- Intergovernmental Panel on Climate Change. 2013. The physical science basis: working group I contribution to the fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, New York, NY.
- Taylor C, de Jeu R, Guichard F, Harris P, Dorigo W. 2012. Afternoon rain more likely over drier soils. *Nature* 489:423–426. <https://doi.org/10.1038/nature11377>.
- Porporato A, Vico G, Fay P. 2006. Superstatistics of hydro-climatic fluctuations and interannual ecosystem productivity. *Geophys Res Lett* 33: L15402. <https://doi.org/10.1029/2006GL026412>.
- Guo Y, Gong P, Amundson R, Yu Q. 2006. Analysis of factors controlling soil carbon in the conterminous United States. *Soil Sci Soc Am J* 70: 601–612. <https://doi.org/10.2136/sssaj2005.0163>.
- Owensby C, Ham J, Auen L. 2006. Fluxes of CO₂ from grazed and ungrazed tallgrass prairie. *Rangel Ecol Manag* 59:111–127. <https://doi.org/10.2111/05-116R2.1>.
- Zhang L, Wylie B, Ji L, Gilmanov T, Tieszen L, Howard D. 2011. Upscaling carbon fluxes over the Great Plains grasslands: sinks and sources. *J Geophys Res Biogeosci* 116:G00J03.
- Zhang L, Wylie B, Ji L, Gilmanov T, Tieszen L. 2010. Climate-driven interannual variability in net ecosystem exchange in the northern Great Plains grasslands. *Rangel Ecol Manag* 63:40–50. <https://doi.org/10.2111/08-232.1>.
- Hoover D, Knapp A, Smith M. 2016. The immediate and prolonged effects of climate extremes on soil respiration in a mesic grassland. *J Geophys Res Biogeosci* 121:1034–1044. <https://doi.org/10.1002/2015JG003256>.
- Hoover D, Rogers B. 2016. Not all droughts are created equal: the impacts of interannual drought pattern and magnitude on grassland carbon cycling. *Glob Chang Biol* 22:1809–1820. <https://doi.org/10.1111/gcb.13161>.
- Zeglin LH, Bottomley PJ, Jumpponen A, Rice CW, Arango M, Lindsley A, McGowan A, Mfombep P, Myrold DD. 2013. Altered precipitation regime affects the function and composition of soil microbial communities on multiple time scales. *Ecology* 94:2334–2345. <https://doi.org/10.1890/12-2018.1>.
- Barnard R, Osborne C, Firestone M. 2013. Responses of soil bacterial and fungal communities to extreme desiccation and rewetting. *ISME J* 7:2229–2241. <https://doi.org/10.1038/ismej.2013.104>.
- Welsh D. 2000. Ecological significance of compatible solute accumulation by micro-organisms: from single cells to global climate. *FEMS Microbiol Rev* 24:263–290. <https://doi.org/10.1111/j.1574-6976.2000.tb00542.x>.
- Evans S, Wallenstein M. 2012. Soil microbial community response to drying and rewetting stress: does historical precipitation regime matter? *Biogeochemistry* 109:101–116. <https://doi.org/10.1007/s10533-011-9638-3>.
- Knapp A, Carroll C, Denton E, La Pierre K, Collins S, Smith M. 2015. Differential sensitivity to regional-scale drought in six central US grasslands. *Oecologia* 177:949–957. <https://doi.org/10.1007/s00442-015-3233-6>.
- Knapp A, Ciais P, Smith M. 2017. Reconciling inconsistencies in precipitation-productivity relationships: implications for climate change. *New Phytol* 214:41–47. <https://doi.org/10.1111/nph.14381>.
- Buis G, Blair J, Burkepille D, Burns C, Chamberlain A, Chapman P, Collins S, Fynn R, Govender N, Kirkman K, Smith M, Knapp A. 2009. Controls of aboveground net primary production in mesic savanna grasslands: an inter-hemispheric comparison. *Ecosystems* 12:982–995. <https://doi.org/10.1007/s10021-009-9273-1>.
- Jansson JK, Hofmockel KS. 2018. The soil microbiome—from metagenomics to metaphenomics. *Curr Opin Microbiol* 43:162–168. <https://doi.org/10.1016/j.mib.2018.01.013>.
- Knapp AK, Beier C, Briske DD, Classen AT, Luo Y, Reichstein M, Smith MD, Smith SD, Bell JE, Fay PA, Heisler JL, Leavitt SW, Sherry R, Smith B, Weng E. 2008. Consequences of more extreme precipitation regimes for terrestrial ecosystems. *Bioscience* 58:811–821. <https://doi.org/10.1641/B580908>.
- Swenson T, Karaoz U, Swenson J, Bowen B, Northen T. 2018. Linking soil biology and chemistry in biological soil crust using isolate exometabolomics. *Nat Commun* 9:19. <https://doi.org/10.1038/s41467-017-02356-9>.
- Mackelprang R, Grube A, Lamendella R, Jesus E, Copeland A, Liang C, Jackson R, Rice C, Kapucija S, Parsa B, Tringe S, Tiedje J, Jansson J. 2018. Microbial community structure and functional potential in cultivated and native tallgrass prairie soils of the midwestern United States. *Front Microbiol* 9:1775. <https://doi.org/10.3389/fmicb.2018.01775>.
- Harper C, Blair J, Fay P, Knapp A, Carlisle J. 2005. Increased rainfall variability and reduced rainfall amount decreases soil CO₂ flux in a grassland ecosystem. *Glob Chang Biol* 11:322–334. <https://doi.org/10.1111/j.1365-2486.2005.00899.x>.
- Petrie M, Peters D, Yao J, Blair J, Burruss N, Collins S, Derner J, Gherardi L, Hendrickson J, Sala O, Starks P, Steiner J. 2018. Regional grassland productivity responses to precipitation during multiyear above- and below-average rainfall periods. *Glob Chang Biol* 24:1935–1951. <https://doi.org/10.1111/gcb.14024>.
- Wallenstein MD, Hall EK. 2012. A trait-based framework for predicting when and where microbial adaptation to climate change will affect ecosystem functioning. *Biogeochemistry* 109:35–47. <https://doi.org/10.1007/s10533-011-9641-8>.
- Bouskill N, Lim H, Borglin S, Salve R, Wood T, Silver W, Brodie E. 2013. Pre-exposure to drought increases the resistance of tropical forest soil bacterial communities to extended drought. *ISME J* 7:384–394. <https://doi.org/10.1038/ismej.2012.113>.
- Acosta-Martinez V, Cotton J, Gardner T, Moore-Kucera J, Zak J, Wester D, Cox S. 2014. Predominant bacterial and fungal assemblages in agricultural soils during a record drought/heat wave and linkages to enzyme activities of biogeochemical cycling. *Appl Soil Ecol* 84:69–82. <https://doi.org/10.1016/j.apsoil.2014.06.005>.
- Yuste J, Fernandez-Gonzalez A, Fernandez-Lopez M, Ogaya R, Penuelas J, Sardans J, Lloret F. 2014. Strong functional stability of soil microbial communities under semiarid Mediterranean conditions and subjected to long-term shifts in baseline precipitation. *Soil Biol Biochem* 69:223–233. <https://doi.org/10.1016/j.soilbio.2013.10.045>.
- Naylor D, DeGraaf S, Purdom E, Coleman-Derr D. 2017. Drought and host selection influence bacterial community dynamics in the grass root microbiome. *ISME J* 11:2691–2704. <https://doi.org/10.1038/ismej.2017.118>.
- Santos-Medellin C, Edwards J, Liechty Z, Nguyen B, Sundaresan V. 2017. Drought stress results in a compartment-specific restructuring of the rice root-associated microbiomes. *mBio* 8:e00764-17. <https://doi.org/10.1128/mBio.00764-17>.
- Hartmann A, Barnard R, Marhan S, Niklaus P. 2013. Effects of drought and N-fertilization on N cycling in two grassland soils. *Oecologia* 171: 705–717. <https://doi.org/10.1007/s00442-012-2578-3>.
- Pascual N, Ranjard L, Kaisermann A, Bachar D, Christen R, Terrat S, Mathieu O, Leveque J, Mougé C, Henault C, Lemanceau P, Pean M, Boiry S, Fontaine S, Maron P. 2013. Stimulation of different functional groups of bacteria by various plant residues as a driver of soil priming effect. *Ecosystems* 16:810–822. <https://doi.org/10.1007/s10021-013-9650-7>.
- Battistuzzi F, Hedges S. 2009. A major clade of prokaryotes with ancient adaptations to life on land. *Mol Biol Evol* 26:335–343. <https://doi.org/10.1093/molbev/msn247>.
- Tas N, Prestat E, McFarland J, Wickland K, Knight R, Berhe A, Jorgenson T, Waldrop M, Jansson J. 2014. Impact of fire on active layer and permafrost microbial communities and metagenomes in an upland

- Alaskan boreal forest. *ISME J* 8:1904–1919. <https://doi.org/10.1038/ismej.2014.36>.
34. Wilhelm R, Cardenas E, Maas K, Leung H, McNeil L, Berch S, Chapman W, Hope G, Kranabetter J, Dube S, Busse M, Fleming R, Hazlett P, Webster K, Morris D, Scott D, Mohn W. 2017. Biogeography and organic matter removal shape long-term effects of timber harvesting on forest soil microbial communities. *ISME J* 11:2552–2568. <https://doi.org/10.1038/ismej.2017.109>.
 35. Mohammadipanah F, Wink J. 2016. Actinobacteria from arid and desert habitats: diversity and biological activity. *Front Microbiol* 6:1541. <https://doi.org/10.3389/fmicb.2015.01541>.
 36. Jumpponen A, Jones K, Blair J. 2010. Vertical distribution of fungal communities in tallgrass prairie soil. *Mycologia* 102:1027–1041. <https://doi.org/10.3852/09-316>.
 37. Fay P, Carlisle J, Danner B, Lett M, McCarron J, Stewart C, Knapp A, Blair J, Collins S. 2002. Altered rainfall patterns, gas exchange, and growth in grasses and forbs. *Int J Plant Sci* 163:549–557. <https://doi.org/10.1086/339718>.
 38. Fay P, Carlisle J, Knapp A, Blair J, Collins S. 2000. Altering rainfall timing and quantity in a mesic grassland ecosystem: design and performance of rainfall manipulation shelters. *Ecosystems* 3:308–319. <https://doi.org/10.1007/s100210000028>.
 39. Linn D, Doran J. 1984. Effect of water-filled pore-space on carbon dioxide and nitrous oxide production in tilled and non-tilled soils. *Soil Sci Soc Am J* 48:1267–1272. <https://doi.org/10.2136/sssaj1984.03615995004800060013x>.
 40. Thompson LR, Sanders J, McDonald D, Amir A, Ladau J, Locey KJ, Prill RJ, Tripathi A, Gibbons S, Ackermann G, Navas-Molina J, Janssen S, Koplova E, Vazquez-Baeza Y, Gonzalez A, Morton J, Mirarab S, Zech XZ, Jiang L, Haroon M, Kanbar J, Zhu Q, Song SJ, Kosciulek T, Bokulich NA, Lefler J, Brislawn CJ, Humphrey G, Owens SM, Hampton-Marcell J, Berg-Lyons D, McKenzie V, Fierer N, Fuhrman JA, Clauser A, Stevens RL, Shade A, Pollard KS, Goodwin KD, Jansson JK, Gilbert JA, Knight R, Earth Microbiome Project Consortium. 2017. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 551:457–463. <https://doi.org/10.1038/nature24621>.
 41. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6:1621–1624. <https://doi.org/10.1038/ismej.2012.8>.
 42. Aronesty E. 2013. Comparison of sequencing utility programs. *Open Bioinforma J* 7:1–8. <https://doi.org/10.2174/1875036201307010001>.
 43. Bushnell B. 2014. BBMap: a fast, accurate, splice-aware aligner.
 44. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–2461. <https://doi.org/10.1093/bioinformatics/btq461>.
 45. Cole JR, Wang Q, Fish JA, Chai BL, McGarrell DM, Sun YN, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM. 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* 42:D633–D642. <https://doi.org/10.1093/nar/gkt1244>.
 46. Voskoboinik A, Neff NF, Sahoo D, Newman AM, Pushkarev D, Koh W, Passarelli B, Fan HC, Mantalas GL, Palmeri KJ, Ishizuka KJ, Gissi C, Griggio F, Ben-Shlomo R, Corey DM, Penland L, White RA, Weissman IL, Quake SR. 2013. The genome sequence of the colonial chordate, *Botryllus schlosseri*. *Elife* 2:e00569. <https://doi.org/10.7554/eLife.00569>.
 47. White RA, III, Bottos EM, Roy Chowdhury T, Zucker JD, Brislawn CJ, Nicora CD, Fansler SJ, Glaesemann KR, Glass K, Jansson JK. 2016. Molecule long-read sequencing facilitates assembly and genomic binning from complex soil metagenomes. *mSystems* 1:e00045-16. <https://doi.org/10.1128/mSystems.00045-16>.
 48. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
 49. Konwar KM, Hanson NW, Bhatia MP, Kim D, Wu SJ, Hahn AS, Morgan-Lang C, Cheung HK, Hallam SJ. 2015. MetaPathways v2.5: quantitative functional, taxonomic and usability improvements. *Bioinformatics* 31:3345–3347. <https://doi.org/10.1093/bioinformatics/btv361>.
 50. Nakayasu ES, Nicora CD, Sims AC, Burnum-Johnson KE, Kim YM, Kyle JE, Matzke MM, Shukla AK, Chu RK, Schepmoes AA, Jacobs JM, Baric RS, Webb-Robertson BJ, Smith RD, Metz TO. 2016. MPlex: a robust and universal protocol for single-sample integrative proteomic, metabolomic, and lipidomic analyses. *mSystems* 1:e00043-16. <https://doi.org/10.1128/mSystems.00043-16>.
 51. Canelas AB, Ras C, ten Pierick A, van Dam JC, Heijnen JJ, Van Gulik WM. 2008. Leakage-free rapid quenching technique for yeast metabolomics. *Metabolomics* 4:226–239. <https://doi.org/10.1007/s11306-008-0116-4>.
 52. Kim YM, Nowack S, Olsen MT, Becraft ED, Wood JM, Thiel V, Klapper I, Kuhl M, Fredrickson JK, Bryant DA, Ward DM, Metz TO. 2015. Diel metabolomics analysis of a hot spring chlorophototrophic microbial mat leads to new hypotheses of community member metabolisms. *Front Microbiol* 6:209. <https://doi.org/10.3389/fmicb.2015.00209>.
 53. Hiller K, Hangebrauk J, Jager C, Spura J, Schreiber K, Schomburg D. 2009. MetaboliteDetector: comprehensive analysis tool for targeted and non-targeted GC/MS based metabolome analysis. *Anal Chem* 81:3429–3439. <https://doi.org/10.1021/ac802689c>.
 54. Kind T, Wohlgemuth G, Lee DY, Lu Y, Palazoglu M, Shahbaz S, Fiehn O. 2009. FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Anal Chem* 81:10038–10048. <https://doi.org/10.1021/ac901952z>.
 55. McMurdie PJ, Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. <https://doi.org/10.1371/journal.pone.0061217>.
 56. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. <https://doi.org/10.1186/s13059-014-0550-8>.
 57. Luo WJ, Friedman MS, Shedden K, Hankenson KD, Woolf PJ. 2009. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics* 10:161. <https://doi.org/10.1186/1471-2105-10-161>.
 58. Luo WJ, Brouwer C. 2013. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* 29:1830–1831. <https://doi.org/10.1093/bioinformatics/btt285>.
 59. R Development Core. 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
 60. Matzke MM, Waters KM, Metz TO, Jacobs JM, Sims AC, Baric RS, Pounds JG, Webb-Robertson B. 2011. Improved quality control processing of peptide-centric LC-MS proteomics data. *Bioinformatics* 27:2866–2872. <https://doi.org/10.1093/bioinformatics/btr479>.
 61. Roweis S. 1998. EM algorithms for PCA and SPCA, p 626–632. In Jordan MI, Kearns MJ, Solla SA (ed), *Advances in neural information processing systems* 10. Morgan Kaufmann Publishers, Burlington, MA.
 62. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J R Stat Soc Series B Methodol* 57:289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
 63. Darzi Y, Letunic I, Bork P, Yamada T. 2018. iPath3.0: interactive pathways explorer v3. *Nucleic Acids Res* 46:W510–W513. <https://doi.org/10.1093/nar/gky299>.
 64. Wang L, Qu L, Zhang L, Hu J, Tang F, Lu M. 2016. Metabolic responses of poplar to *Apriona germari* (Hope) as revealed by metabolite profiling. *17:923*. <https://doi.org/10.3390/ijms17060923>.