Title
Effects of Agricultural Management on Rhizosphere Microbial Structure and Function in Processing Tomato Plants.

Permalink
https://escholarship.org/uc/item/083580tc

Journal
Applied and environmental microbiology, 85(16)

ISSN
0099-2240

Authors
Schmidt, Jennifer E
Vannette, Rachel L
Igwe, Alexandria
et al.

Publication Date
2019-08-01

DOI
10.1128/aem.01064-19

Peer reviewed
Effects of agricultural management on rhizosphere microbial structure and function in processing tomato

Running title: Management effects on tomato rhizosphere

Jennifer E. Schmidt, Rachel L. Vannette, Alexandria Igwe, Rob Blundell, Clare L. Casteel, Amélie C.M. Gaudin

aDepartment of Plant Sciences, University of California at Davis, One Shields Avenue, Davis, CA 95616
bDepartment of Entomology and Nematology, University of California at Davis, One Shields Avenue, Davis, CA 95616
cDepartment of Plant Pathology, University of California at Davis, One Shields Avenue, Davis, CA 95616

#Corresponding author:

Email: agaudin@ucdavis.edu
Phone: 530-752-1212

Mailing address: Department of Plant Sciences
University of California Davis
1 Shields Avenue
Plant and Environmental Science Building
Davis CA 95616
Abstract

Agricultural management practices affect bulk soil microbial communities and the functions they carry out, but it remains unclear how these effects extend to the rhizosphere in different agroecosystem contexts. Given close linkages between rhizosphere processes and plant nutrition and productivity, understanding how management practices impact this critical zone is of great importance to optimize plant-soil interactions for agricultural sustainability. A comparison of six paired conventional-organic processing tomato farms was conducted to investigate relationships between management, soil physicochemical parameters, and rhizosphere microbial community composition and functions. Organically managed fields were higher in soil total N and NO$_3$-N, total and labile C, plant Ca, S, and Cu, and other essential nutrients, while soil pH was higher in conventionally managed fields. Differential abundance, indicator species, and random forest analyses of rhizosphere communities revealed compositional differences between organic and conventional systems and identified management-specific microbial taxa. Phylogeny-based trait prediction showed that these differences translated into more abundant pathogenesis-related gene functions in conventional systems. Structural equation modeling revealed a greater effect of soil biological communities than physicochemical parameters on plant outcomes. These results highlight the importance of rhizosphere-specific studies, as plant selection likely interacts with management in regulating microbial communities and functions that impact agricultural productivity.

Importance

Agriculture relies in part on close linkages between plants and the microorganisms that live in association with plant roots. These rhizosphere bacteria and fungi are distinct from microbial communities found in the rest of the soil and are even more important to plant nutrient uptake and health. Evidence from field studies shows that agricultural management practices such as fertilization and tillage shape microbial communities in bulk soil, but little is known about how these practices affect the rhizosphere. We investigated how agricultural management affects plant-soil-microbe interactions by comparing soil
physical and chemical properties, plant nutrients, and rhizosphere microbial communities from paired fields under organic and conventional management. Our results show that human management effects extend even to microorganisms living in close association with plant roots and highlight the importance of these bacteria and fungi to crop nutrition and productivity.

Introduction

Soil microbial communities mediate the provision of many ecosystem services by soils and are increasingly recognized as fundamental regulators of plant and environmental outcomes of agroecosystems. Agricultural practices such as nutrient inputs and tillage have been shown to shape bulk soil microbial communities and functions across spatial and temporal scales (1–4). Comparisons of bulk soil under different management strategies, i.e. organic (nutrients provided from sources other than synthetic inputs) vs. conventional management have revealed effects on soil properties that in turn drive variation in microbial communities at small and intermediate scales (5–8). Small-scale studies designed to minimize environmental heterogeneity, such as long-term experiments on a single site, show strong effects of management on soil physicochemical parameters (9, 10), microbial biomass (9), and habitat-specific bacterial and fungal taxa (11). At an intermediate spatial scale, such as paired fields within a region, contextual variables such as climate, soil type, and cropping system largely influence the soil physicochemical parameters and microbial processes that differ between conventional and organic fields. Organically managed processing tomato fields in California have higher levels of organic carbon, microbial abundance and diversity, and N mineralization potential compared to conventional, while soils under conventional management have higher inorganic N pools and salinity (7). However, these studies often have not extended to the rhizosphere, and the studies that have done so have not found universal predictors of rhizosphere community assembly across contexts and scales (4, 12–14).
While bulk soil communities affect recruitment and assembly of rhizosphere microbial communities (15), soil under the influence of plant roots represents a unique environment that must be studied separately (16). The rhizosphere is a hotspot of interactions where dynamic relationships between plant roots and soil microbial communities occur, allowing bacteria and fungi to break down and cycle organic matter and release nutrients (17), promote plant growth via direct and indirect mechanisms (18), and suppress pathogens (19). While linking agricultural management to large-scale outcomes such as nutrient fluxes or ecosystem services requires analysis of bulk soil properties and processes, understanding the complex relationship between management practices and plant nutrition and productivity necessitates shifting focus to the rhizosphere (20). Some evidence suggests that management can affect the ecosystem-level functions carried out by bulk soil microbial communities through impacts on microbial diversity (21), but the unique chemistry and microbial communities found in the rhizosphere (22) are more closely linked to plant outcomes of agricultural importance (23). Because rhizosphere soil is shaped by complex interactions between plant and bulk soil processes, the effects of agricultural management on rhizosphere communities and the functional implications are not always easy to predict.

The few studies that have addressed this question have concluded that differences in bulk soil microbial and protist communities do carry over to some extent to rhizosphere communities (22, 24). However, such studies have frequently been conducted on long-term research stations (22, 24), leaving open the questions of scale and context. Do management effects on rhizosphere microbial communities extend to an intermediate scale, such as paired fields within a region? If so, what soil properties are most closely linked to microbial variation, and how do differences in rhizosphere microbial communities influence plant health and productivity?

A regional-scale study of paired organic and conventional processing tomato fields in Northern California was conducted to i) characterize impacts of agricultural management on rhizosphere microbial community composition in California processing tomato agroecosystems at an intermediate spatial scale, ii) identify how taxonomic shifts affected predicted metabolic and ecological functions carried out by
these communities, and iii) explore the effects of management-induced microbial variation on crop
nutrition and productivity. To address the first objective of identifying variation in rhizosphere microbial
communities, we employed three complementary approaches: differential abundance, indicator species,
and random forest analyses. Differential abundance analysis of microbial communities adapts RNA-seq
methodology used for gene expression to identify taxa whose abundance varies significantly among
groups of samples (25). Indicator species analysis, an alternative approach, detects taxa preferentially
associated with a given habitat or sample group based on a combination of specificity and fidelity rather
than relative abundance alone (26). Random forest analysis (27), a machine learning method, approaches
the microbe-sample group linkage from the opposite direction than the differential abundance and
indicator species approaches, identifying key taxa whose abundance can be used to assign samples to the
appropriate group.

The second objective, determining whether agricultural management induces shifts in rhizosphere
microbial functions, was addressed using phylogeny-based trait prediction. This method predicts
metagenomic data such as genes involved in key agroecological functions from 16S amplicon sequencing
data (28). Structural equation modeling (SEM), a statistical technique to test hypothesized relationships
among variables (29), was used to address our final objective of exploring linkages between soil
properties, microbial communities, and plant nutrition and productivity. We hypothesized that rhizosphere
community structure and function would differ between conventional and organic systems and that
divergent microbial communities would relate to variation in plant traits within and between fields.

Results

Site and management drive variation in soil and plant variables

Site had a stronger influence on bulk soil and plant variables than management category (organic vs.
conventional) (site $R^2=0.54$, $p=0.001$; management $R^2=0.17$, $p=0.001$) and the site x management
interaction was significant ($R^2=0.09$, $p=0.001$). Two principal components (PCs) explained 41.99% (PC1)
6 and 21.93% (PC2) of variation among samples, respectively (Figure 1). Samples tended to cluster primarily by site along PC1, which was affected by numerous plant and soil nutrients, and secondarily by management within each site. PC2 was primarily influenced by plant Cu, Mg, and Mn as well as soil Mg and NO$_3$-N (Figure 1, Table 1).

Management system significantly affected soil physicochemical variables ($p<0.001$). Soil parameters that were higher in organically managed fields included total N ($p<0.001$), C ($p<0.001$), NO$_3$-N ($p=0.008$), Olsen-extractable P ($p=0.008$), K ($p=0.0087$), Na ($p<0.001$), organic matter (OM) ($p=0.0022$), and permanganate-oxidizable carbon (PoxC) ($p<0.001$), while pH was higher in conventionally managed fields ($p=0.044$). Soil physicochemical properties were highly correlated with one another (Table 2).

Magnesium was correlated with only NO$_3$-N, cation exchange capacity (CEC) and pH, but other macronutrients and key soil properties tended to vary together. Management also affected plant nutrients ($p<0.001$), many of which were correlated with one another (Table 3). Concentrations of Ca ($p=0.004$), S ($p<0.001$), and Cu ($p<0.001$) were all higher in plants from organically managed fields.

**Rhizosphere microbial community composition responds to management practices**

The species composition of both bacterial and fungal rhizosphere communities varied according to site and management (Figure 2), and these effects were also observed when phylogenetic relatedness of bacterial communities was considered (Figure S1 in the supplementary material). Tests of multivariate homogeneity of group dispersions (betadisper function of the vegan package) showed that dispersions did not differ among sites or management types (both $p>0.05$). Management influenced rhizosphere microbial communities, but to different extents depending on the site identity (Figure 2, site x management interaction bacteria $R^2=0.12$, $p<0.01$; fungi $R^2=0.10$, $p<0.01$). Management accounted for the greatest proportion of variation (53%) in bacterial communities at the MR site ($R^2=0.53$, $p=0.02$), slightly more than at the PF site ($R^2=0.43$, $p=0.02$) and nearly three times as much as at the RR site ($R^2=0.19$, $p=0.01$).

Fungal communities were also affected by management, which accounted for 22% of variation at the MR site ($R^2=0.22$, $p=0.02$), 38% at the RR site ($R^2=0.38$, $p=0.01$), and 43% at the PF site ($R^2=0.43$, $p=0.02$).
Bacterial diversity was affected by the site x management interaction (p<0.001). The Shannon index was higher in organically managed fields than conventionally managed fields at all sites except the MR site (Table 4). Fungal diversity was affected by site (p<0.001) and management (p<0.001), but not the interaction. Fungal diversity was higher in organically managed fields at all sites, and higher at the PF site than RR or MR (Table 4).

Forty-eight bacterial amplicon sequence variants (ASVs) differed in abundance between the rhizospheres of conventionally and organically managed plants at the α=0.01 level (Figure 3a). ASVs more abundant in organically managed rhizospheres included two members of the genus *Pseudomonas*, while ASVs more abundant in conventionally managed rhizospheres included six members of the genus *Flavobacterium* and three members of the genera *Devosia* and *Lysobacter*. An ASV belonging to the genus *Pseudomonas* had the highest relative abundance in organically managed fields, and an ASV belonging to the genus *Chryseobacterium* had the highest relative abundance in conventionally managed fields.

Nineteen fungal ASVs differed in abundance between management systems at the p=0.01 level, only one of which was more abundant in conventionally managed fields (Figure 3b). ASVs more abundant in organically managed plant rhizospheres included three members of the genus *Holtermanniella*, three members of the genus *Mucor*, and two members of the genus *Pyrenochaetopsis*. The ASV more abundant in conventionally managed rhizospheres was identified as *Plectosphaerella cucumerina*. *Mucor hiemalis* was most abundant in organic systems relative to conventional.

Since system management has a strong impact on multiple soil properties, we conducted redundancy analysis (RDA) with forward selection to identify which soil physicochemical properties have the greatest influence on rhizosphere bacterial and fungal community composition. After site and management, Ca was the most significant driver of both bacterial and fungal community composition. Bacterial community composition also responded to Mg levels, while fungi were significantly influenced by Na and K.
Indicator species of rhizosphere communities differ between systems

Indicator species analysis showed 57 system-specific bacterial ASVs: 35 with the conventional system and 22 with the organic system (Table S3). Members of the genera *Flavobacterium* (8), *Pedobacter* (4), *Lysobacter* (3), and *Pseudomonas* (3) had by the greatest number of sequences in the conventional system and *Pseudomonas* (4) in the organic system. Fewer fungal indicator taxa were discovered, with only four fungal ASVs associated with the conventional system but 17 with the organic system. The four ASVs associated with the conventional system came from different genera, while *Holtermanniella* (6) and *Mucor* (4) were the most represented indicator genera in the organic system. Fifteen of the 78 taxa identified by indicator species analysis were also differentially abundant. Because the IndVal index represents the probability of finding a given species in the environment of interest, taxa with a high relative abundance in the environment will generally score high on the fidelity component of the IndVal index. This was the case for *Flavobacterium* in the conventional system and *Pseudomonas*, *Holtermanniella*, and *Mucor* in the organic system.

Random forest (RF) analysis was used to identify ASVs that could be used to discriminate between management systems. ASVs belonging to the genera *Lysobacter* and *Gibellulopsis* had the greatest impact on the mean decrease in accuracy and mean decrease in Gini coefficient of the random forest model (Figure S2 in the supplementary material). Substantial overlap was observed between the results of RF analysis and differential abundance analysis. Eleven of the twenty most significant ASVs from the RF analysis had also been identified through differential abundance analysis, although ASVs such as *Gibellulopsis* that had a significant impact on the RF model only slightly differed in abundance between systems (Figure 3).

Management induces changes in predicted rhizosphere bacterial functions

Of the total number of genes predicted, 4.8% (169) differed in abundance between the rhizosphere of organic and conventional plants. Of those genes, 79 were more abundant in the organic system and 90
were enriched in the conventional system. Functions corresponding to cellular processes including quorum sensing, biofilm formation, and chemotaxis showed the greatest difference between systems, with only two peroxisome functions upregulated in the organic system and 31 upregulated in the conventional system (Figure 4). Genes with the highest relative abundance in the organic system were distributed across a variety of functions, including ABC transporters (12), two-component systems (8), biosynthesis of siderophores (5), starch and sucrose metabolism (5), and type I polyketide structures (5). A component of the trcR/trcS two-component regulatory system, trcR (K07672), was up-regulated by the greatest ratio in organic systems. Genes with greater relative abundance in the conventional system tended to be associated with biosynthesis of amino acids (19), two-component systems (18), quorum sensing (10), ABC transporters (9), and biofilm formation (9) (Figure 4).

Structural equation modeling identifies key linkages among plant, soil, and microbial variables

Hypothetical links between bulk soil physicochemical parameters, plant nutrition, rhizosphere microbial communities, and plant biomass were tested using structural equation modeling (SEM) across management systems (Figure 5a). Bacterial and fungal communities were represented by two vectors each (PC1B, PC2B, PC1F, PC2F) that were derived from principal components analysis shown in Figure 2. Plant biomass was most strongly positively correlated with plant P, which in turn was most strongly correlated with fungi from the PC2F vector (Figure 5b). The taxa that contributed most to PC2F were Vishniacozyma victoriae and an unidentified Sobicocozyma sp. Neither of these species were identified in the differential abundance analysis (Figure 3b). Fungi from the PC1F vector had a slight positive influence on plant Na and included ASVs classified as Alternaria sp., Cryptococcus aerius, and Plectosphaerella cucumerina. PC1B, the first principal component of bacterial communities, was negatively correlated with shoot C:N ratio; the three ASVs with the greatest contribution to this component were a strain of Pseudomonas and two strains of Stenotrophomonas. The second principal component of bacterial communities (PC2B) was slightly positively correlated with plant biomass, P, Na,
and C:N ratio; four of the five ASVs with the greatest contribution to this component were classified as *Pseudomonas* sp.

The final SEM had a $\chi^2$ test statistic of 1.907 with 3 degrees of freedom, giving a $\chi^2/v$ ratio of 0.64, root mean square error of approximation (RMSEA) of 0.000 (90% confidence interval 0.000≤x≤0.195), comparative fit index (CFI) of 1.000, Tucker Lewis index (TLI) of 1.062, and standardized root mean square residual (SRMR) of 0.016. A low $\chi^2/v$ ratio indicates a good model, although this test statistic does not perform well with small sample sizes (30). The CFI and TLI model indices perform well with small sample sizes and are above acceptable threshold (0.95 for a good model (31)). An SRMR less than 0.08 generally indicates that a model fits the data well (32).

**Discussion**

Our objectives were to explore how management practices implemented in organic and conventional tomato production systems shape rhizosphere microbial composition, infer how taxonomic shifts affected microbe-mediated functions, and identify linkages between management-induced shifts in soil physicochemical parameters, rhizosphere microbial communities, and plant nutrition and productivity. In support of our hypotheses, we identified specific taxa that differed in abundance between management systems and predicted the functional implications of those shifts in community composition (Figure 3, Figure 4, Figure S2). Some differentially abundant taxa were confirmed as indicator species that could be used to distinguish communities between management systems. More importantly, phylogeny-based trait prediction showed that management-induced differences in rhizosphere bacterial community composition translated into agriculturally relevant outcomes, particularly with regard to plant nutrition and pathogen-related functions such as quorum sensing and biofilm formation(33, 34) (Figure 4, Figure 5). Although our techniques could not examine the contribution of fungi to predicted function, it is likely that observed compositional shifts in fungal communities increase divergence in functional outcomes between systems.
Bacterial diversity was higher in the rhizospheres of organically managed plants at all sites except MR, and fungal diversity was higher in the organic system across sites, consistent with other studies finding increased microbial diversity under organic management (2, 7, 35–37). Numerous bacterial ASVs belonging to the genus *Pseudomonas*, which contains members known to possess plant-growth-promoting properties (18, 19), had a higher relative abundance in organic systems (Figure 3a). Sixteen of the 17 differentially abundant fungal ASVs were found at higher abundance in the rhizosphere of plants growing in the organic system; these included numerous members of the genera *Holtermanniella* and *Mucor* (Figure 3b). *Holtermanniella* is a small, cold-tolerant genus of potentially parasitic fungi (38) that includes species able to metabolize diverse carbon compounds and generate unique fatty acid profiles (39). *Mucor* are a genus of starch decomposing fungi (40) that are capable of metabolizing a wide range of complex carbohydrates (41). Although a long-term comparison of conventional and organic management found no difference in the relative abundance of *Mucor* sp. in bulk soils (42), potential shifts in the rhizosphere have not been shown. In addition, predicted potential community functions also differed between soils under different management systems. Although our approach relies on predicted potential (DNA-based) functions rather than genomic or transcriptomic information from the strains found at these sites, tax4fun performs well in comparison with shotgun metagenomic data from soils (28), suggesting that broad patterns may be informative. Bacterial community shifts in the rhizospheres of organically managed plants were associated with a higher abundance of predicted genes involved in starch and sucrose metabolism and biosynthesis of siderophores, which can increase the availability of micronutrients such as iron (Figure 4). Other enzymes with high relative abundance in the organic system catalyze reactions involved in the metabolism of tyrosine, carotenoids, and other complex organic compounds (Figure 4).

Rhizosphere diversity was generally lower under conventional management, and community composition and functions were notably different. ASVs belonging to the genera *Flavobacterium*, *Devosia*, and *Lysobacterium* had higher relative abundances in the conventional system. The *Flavobacterium* genus has
been found elsewhere to increase in abundance in response to six years of intensive organic vegetable production (43), suggesting that individual species within the genus may respond differently to conventional and organic management. Members of Lysobacterium have been shown to degrade complex aromatic compounds (44). Plectosphaerella cucumerina, a known pathogen that causes rots on a variety of horticultural species (45), was the only fungal ASV found to be more abundant in the conventional system. Perhaps due to the greater abundance of this pathogen, functions upregulated in the conventional system included genes related to quorum sensing and biofilm formation (Figure 4).

Management practices and sites had strong influence on soil chemical properties, which in turn affected bacterial and fungal community composition. Forward selection revealed that the two kingdoms responded to different sets of soil physicochemical parameters: bacterial community composition was affected by Ca and Mg, while fungal community composition was affected by Ca, Na, and K. These predictors are notably different from variables commonly accepted as important for microbial community composition, such as organic matter (46, 47), pH (48, 49), and N. The failure of organic matter and N to predict microbial community structure is surprising at first glance, given that scarce C and N availability can limit rates of microbial growth and functions such as mineralization, and that the abundance of N-cycling microbial taxa often varies with C and inorganic N species. However, this result is consistent with multiple studies showing no effect of N on microbial community composition (50–52). Agricultural management might outweigh the effects of variation in these parameters, since Ca and Mg were not affected by management. It may also be that low variation in organic matter, pH, and soil N within the context of this study reduced the ability of these parameters to explain variation in community composition (Table S2 in the supplementary material).

Soil Ca and Na have similarly appeared elsewhere as significant predictors of microbial community composition. In another comparison of management systems, soil Ca was higher in soils receiving organic amendments than synthetic amendments and was among the parameters correlated with microbial community composition (53). Ca was also a primary driver of microbial community composition in a
multi-year study of a soil amended with composted tannery sludge (54). Salinity frequently drives variation in microbial community composition, especially in irrigated systems, although most commonly when a stronger salinity gradient is present due to environmental filtering based on salinity tolerance (55–57).

SEM tied together this observed variation in microbial community composition with soil and plant variables and tested a hypothetical model linking plant and soil biological and physicochemical parameters with plant biomass (Figure 5). Management was not retained in the final model, suggesting that management effects were indirect and captured by other included variables at these sites. Other studies have similarly found that soil type and physicochemical parameters affect microbial community composition and catabolic functions more than long-term agricultural management practices (58). Within this study, it appears that rhizosphere microbial communities were more closely linked to differences in bulk soil properties created by management systems than to the management practices themselves.

SEM revealed a greater relative influence of rhizosphere biological communities than bulk soil physicochemical characteristics on plant nutrient content and biomass (Figure 5). A strong indirect linkage was observed between microbial communities and plant biomass: fungal community composition was strongly positively correlated with plant P, which in turn strongly correlated with shoot biomass (Figure 5). The link between plant P and fungal communities is particularly striking given the absence of sequences belonging to the phylum Glomeromycota, which contains mycorrhizal fungi (data not shown). The lack of mycorrhizal sequences may be partly explained by choice of amplicon or primer bias (59).

Since the length of the amplified region differs for mycorrhizae compared to the more abundant Ascomycota and Basidiomycota (60); it is unlikely that mycorrhizae were truly absent from all samples. Nonetheless, even non-mycorrhizal fungi can improve plant P status through solubilization, mineralization, and direct transfer of phosphate (61). Genera such as *Aspergillus* and *Penicillium* release organic acids that can solubilize phosphate, potentially rendering it available for direct uptake by plants or mycorrhizae (62).
PC2B was slightly positively correlated with plant Na and PC1B was negatively correlated with C:N ratio. The correlation between bacterial community composition and plant Na could be the direct effect of microbial interference in plant metabolism, or changes in soil parameters could foster unique microbial communities and also increase plant Na. While limitations of the measured data do not allow us to distinguish between these explanations in this context, microbial influence on plant Na has been reported elsewhere; certain bacterial strains are capable of plant tissue-specific regulation of sodium transporters that increases salt tolerance in Arabidopsis (63), while other bacterial strains reduce salt accumulation in salinity-stressed plants (64). A negative correlation with C:N ratio indicates that the bacterial populations improved plant N content, a result that could be due to increased N availability via N fixation or mineralization of organic matter.

This study identified rhizosphere microbial taxa and functions affected by agricultural management and illuminated unexpected linkages between soil, microbes, and crop nutrition and productivity, but compelling questions remain. Organic certification encompasses a diverse set of management practices and variation in cover crop species, green manure inputs, or crop rotation complexity and duration likely lead to diverse effects on soil microbes. To translate the broad, extensive conventional-organic literature into tangible recommendations, future studies should focus on causal relationships between specific inputs or techniques and key soil physicochemical parameters. This could be achieved in part by employing SEM with a much larger dataset (a sample size of at least 200 (65) and data satisfying the requirement of multivariate normality (66)) to allow the incorporation of additional variables (e.g. crop genotype, N fertility source and rate, tillage) and improve the predictive power of the model. Such analysis would add nuance to the results of this study and enable the development of management systems that foster agricultural productivity by maximizing beneficial plant-soil-microbe interactions in the rhizosphere.

Our results add an additional layer of complexity to previous investigations of the effects of agricultural management on microbial communities. Others have noted the importance of scale in determining how
soil properties relate to microbial community composition or function, as geographic scale alters the relative importance of factors such as environmental heterogeneity and distance that influence microbial distribution (67, 68). We emphasize the importance of integrating plants and rhizosphere processes into these discussions of microbial biogeography, particularly at intermediate scales, as plants exert strong influence on rhizosphere communities and may modulate management effects on rhizosphere communities. Management of plant-microbe-soil interactions in the rhizosphere is a critical step toward building more resource-efficient and resilient agricultural systems, and our study indicates that soil management has strong and consistent effects on landscape-level variation in the rhizosphere composition and predicted function.

Materials and Methods

Sample collection

Samples were collected from 6 paired fields under conventional and organic management on Yolo Silt Loam during the 2017 growing season (details of sites and management practices can be found in Table S1 in the supplementary material). Plant and soil samples were collected ~ 6 weeks after transplanting on the same date at paired fields. Samples were taken from six locations per field (two on the exterior margins of the field and four internal). At each location, two entire plants were excavated and shoot and root samples were separated by clipping at the base of the shoot. A bulk soil sample was collected from the upper 10 cm of soil immediately adjacent to each plant. Roots were separated from bulk soil, stored in paper bags and transported to the lab on ice. Twelve root fragments from each plot (6 from each individual plant) were pooled and rhizosphere soil was collected using a shaking wash in an 0.9% NaCl/0.01% Tween 80 (v/v) solution followed by centrifugation. Because this volume of soil was insufficient for full textural and nutrient analysis, we assumed that rhizosphere soil characteristics such as texture, organic matter, etc. would be similar to the parameters measured for the corresponding bulk soil. Shovels and other sampling implements were cleaned thoroughly between samples. The remaining roots and shoots were dried at 60°C and weighed.
Plant and soil analysis

Dried bulk soil samples and aboveground dried biomass were homogenized and analyzed for total nitrogen (N) and carbon (C) via combustion analysis (69). Soil nitrate was measured using a flow injection analyzer (70), soil extractable phosphorus (P) was determined according to Olsen and Sommers (71), and other soil nutrients were measured using ICP-AES (72). Soil organic matter content was determined via the loss-on-ignition method (73). Soil pH was measured on a saturated paste extract. Bulk soil properties can be found in Table S2 in the supplementary material.

Dried aboveground biomass was ground thoroughly to pass a 2 mm sieve. Plant leaf samples were analyzed for N, P, K, Ca, Mg, Mn, Fe, Cu, B, and Zn at the Agricultural Analytical Services Lab of Pennsylvania State University. Total N was analyzed via combustion (74), and concentrations of the remaining elements were determined via hot block acid digestion (75).

Microbial community analysis

DNA was extracted from rhizosphere samples using the MoBio PowerSoil Kit (Qiagen). At least 5 ng of DNA from each sample was sent for library prep and sequencing using MiSeq at Dalhousie IMR facility. The V4-V5 region of the 16S rRNA region was sequenced to characterize bacterial communities and the ITS region of the rRNA gene was sequenced to characterize fungal communities (76, 77). Negative controls were also extracted and submitted, but no reads were recovered. All statistical analyses were carried out using R software (78). Reads were error-corrected and assembled into amplicon sequence variants (ASVs) using DADA2 v.1.8 (79), with taxonomy assigned using SILVA v.128 for bacteria (80) and UNITE database (2017 release) for fungi (81). Taxa without a taxonomic assignment, or assigned to Archaea, mitochondria, or chloroplasts were removed from this dataset. Those not assigned to the kingdom Fungi were removed from the fungal dataset. Sequence abundance was rarefied to 15,310 sequences per sample for bacteria and 13,000 per sample for fungi and all samples approached saturation.
Sequencing data is available in the NCBI SRA data repository under the project accession number PRJNA539989.

Non-metric multidimensional scaling (NMDS) was used to ordain samples in two-dimensional space (ordinate function of the phyloseq package using method = “NMDS”). Two outliers were removed from this and subsequent analyses in order to minimize the stress function. A second NMDS ordination was performed based on weighted UniFrac distances (distance function of phyloseq package with “wunifrac” command) to determine whether phylogenetic distance among samples was affected by site and management. Shannon diversity was calculated for each sample using the estimate_richness function (measures = “Shannon”) of the phyloseq package.

Differential abundance of microbial taxa

Differential abundance of bacterial and fungal in the rhizosphere of plants grown in organically and conventionally managed systems was carried out using the DESeq2 package (25). Although applying this analysis to compositional datasets obtained from sequencing microbial communities has been critiqued (82), the method has been shown to be effective when library sizes are similar across groups and sample size is small (<50 samples per group) (83), as was the case here. Sequences occurring in fewer than three samples were filtered out prior to the analysis to avoid bias due to rare taxa (filter_taxa function of phyloseq package). Dispersions were fit to the mean intensity using a gamma-family GLM by setting the parameter fitType= “parametric” and significance was assessed using the Wald test with a significance threshold of \( \alpha = 0.01 \).

Indicator species analysis

Indicator species analysis was conducted to identify specific rhizosphere microbial taxa that were associated with the conventional or organic system using the indicspecies package (84). Briefly, the Indicator Value (IndVal) index was calculated for each ASV-system combination as the product of specificity and fidelity indices (84). The highest IndVal index for each ASV was tested for significance.
with 999 permutations (multipatt function of indicspecies package using duleg = TRUE and control = how(nperm=999). A Bonferroni correction was used to control the family-wise error rate at α = 0.01.

Random forest analysis

We complemented the indicator species analysis with a random forest approach, which identifies bacterial and fungal ASVs that could be used to classify samples by management system through a machine learning algorithm. Random forest analysis was conducted using the randomForest package (27). The dataset was split into subsets for training (70% of observations) and validation (30% of observations). Model parameters were adjusted to minimize the error rate, but the default parameters for ntree (ntree = 500) and mtry (mtry = √p, with p representing the number of model parameters) resulted in the lowest error rate (6.52%). The classification accuracy was calculated to be 95%, indicating high prediction accuracy. ASVs with the greatest contribution to the classification algorithm were identified according to the highest scores for mean decrease in accuracy or mean decrease in the Gini coefficient (importance function of randomForest package).

Phylogeny-based functional trait prediction

We determined potential shifts in rhizosphere microbial functions with management and soil properties using functional trait prediction of 16S communities with the themetagenomics package (85). Briefly, this package implements Tax4Fun (28) to predict functions from the KEGG Orthology database that are associated with provided abundance tables, sample metadata, and phylogenetic information. Phylogeny is assigned according to the SILVA rRNA database project (80). To identify functions that differed in abundance between systems, predicted functions were subjected to differential abundance analysis using the DESeq2 package. Parameters were identical to those described previously and the significance threshold was set at α=0.01.

Principal component analysis of plant and soil variables
Principal component analysis (PCA) was used to reduce the dimensions of the multivariate dataset containing scaled soil and plant variables, visualize samples in two-dimensional space, and calculate factor loadings (prcomp function of stats package). Outliers for individual soil and plant variables were identified with Grubb’s test (grubbs.test function of outliers package) and removed from the dataset prior to PCA. The multivariate homogeneity of group dispersions (betadisper function of vegan package) was tested to determine whether variances differed among sampling sites. The effect of management on soil and plant variables was tested with multivariate analysis of variance (MANOVA) using the manova function of the stats package (78).

Permutational multivariate analysis of variance

Permutational multivariate analysis of variance was used to test the effect of the interaction between site and management on microbial community composition (adonis function of vegan package), separately for bacteria and fungi. If the interaction was significant, the magnitude of the management effect was then tested within each site. If the interaction was not significant, PERMANOVA was used to test the relative magnitude of site and management effects. Redundancy analysis (RDA) was conducted to identify soil physicochemical properties with the greatest influence on rhizosphere microbial community composition. Parameters that significantly explained variation in bacterial or fungal community composition were identified using forward selection (ordistep function of vegan package).

Structural equation modeling (SEM)

SEM was used to test a hypothetical model linking soil, plant, and microbial variables that affect shoot biomass (Figure 5a). Parameters included in the model were chosen using forward selection of a linear model with shoot biomass as the response variable and all other soil, microbial, and plant parameters as independent variables (step function of stats package) (78). The model was established using the sem function of the lavaan package (86) and visualized with the semPlot package (87). The model was then refined by sequentially removing variables with poor explanatory power (R²<0.50). Management (organic
vs. conventional) was originally included as a variable but was ultimately removed because management significantly and consistently decreased the fit statistics for the model, perhaps because the variables retained in the model were good indicators of management differences.

Although the variables identified by forward selection (soil Na, soil Ca, plant P, plant C:N, plant Na) were not consistent with a hypothesis of multivariate normality, sample size was too small to permit the exclusion of outliers. The first two principal components of microbial species composition, which accounted for 31% and 15% of bacterial variation (PC1B and PC2B respectively) and 26% and 21% of fungal variation (PC1F, PC2F respectively), were used to represent microbial communities in the model (Figure 5). The maximum likelihood (ML) method was used to estimate model fit test statistics. The goodness of fit of the model was tested using standard model fit indices: the ratio of the chi-square statistic to degrees of freedom ($\chi^2/v$), Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Standardized Root Mean Square Residual (SRMR) (88).

Acknowledgements

The authors would like to thank Samuel Tookey, Griffin Hall, and Russell Ranch staff and growers for participating in this study, assisting with sampling, and providing valuable comments on the manuscript. This work was supported by the California Tomato Research Institute to AG, CC and RV and the USDA-NIFA, Agricultural Experiment Station Project #CA-D-PLS-2332-H to AG. The funding sources had no role in the study design, data collection and interpretation, or manuscript preparation.

References


Krebsbach CM. 2014. Bootstrapping with small samples in structural equation modeling: Goodness of fit and confidence intervals. M.A., University of Rhode Island.


Figure Legends

Figure 1: PCA of soil and plant variables measured in six processing tomato fields. Soil physicochemical parameters and plant variables separated primarily by site along PC1, which explained 42% of variation. Samples separated secondarily by management within site, and a significant site x management interaction was observed.

Figure 2: NMDS ordination of microbial communities sampled from the rhizosphere of processing tomatoes. Non-metric dimensional scaling based on Bray-Curtis dissimilarity matrices revealed that A) bacterial and B) fungal communities separated primarily by site and secondarily by management.

Figure 3: Differentially abundant microbial taxa. A) 48 bacterial and B) 19 fungal taxa differed in abundance between conventional and organic management systems at the \( \alpha=0.01 \) level. Colored bars represent the natural logarithm of abundance of each taxa and gray bars represent the ratio of abundance in the organic system to abundance in the conventional system. Multiple strains or species within genus are shown. NA indicates that sequences could not be identified at the genus level.

Figure 4: Differentially abundant functions. Phylogeny-based trait prediction revealed 169 functional genes that differed in abundance between the two systems at the \( \alpha=0.01 \) level, 79 of which were more abundant in the organic system and 90 in the conventional system.

Figure 5: Structural equation model linking soil, plant, and microbial variables. A) A hypothetical model linking soil, microbial, and plant parameters was tested using structural equation modeling. B) The final SEM showed that microbial communities had a strong but indirect effect on plant biomass through a positive correlation between fungal community composition and plant P. Soil Ca and Na affected fungal communities more strongly than bacterial communities. Red represents soil variables, blue represents microbial variables (principal components 1 and 2 extracted from PCA of bacterial and fungal communities, respectively), and green represents plant variables. Dashed lines represent fixed parameters.

Table 1. Factor loadings of scaled soil and plant variables contributing to PC1 and PC2.
<table>
<thead>
<tr>
<th></th>
<th>PC1</th>
<th>PC2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soil N (total)</td>
<td>-0.243</td>
<td>-0.159</td>
</tr>
<tr>
<td>Soil C</td>
<td>-0.254</td>
<td>-0.159</td>
</tr>
<tr>
<td>Soil NO$_3$-N</td>
<td>-0.063</td>
<td>-0.202</td>
</tr>
<tr>
<td>Soil P (Olsen)</td>
<td>-0.186</td>
<td>-0.024</td>
</tr>
<tr>
<td>Soil K</td>
<td>-0.260</td>
<td>-0.132</td>
</tr>
<tr>
<td>Soil Na</td>
<td>-0.190</td>
<td>-0.161</td>
</tr>
<tr>
<td>Soil Ca</td>
<td>-0.215</td>
<td>-0.166</td>
</tr>
<tr>
<td>Soil Mg</td>
<td>-0.108</td>
<td>0.361</td>
</tr>
<tr>
<td>CEC</td>
<td>-0.227</td>
<td>0.219</td>
</tr>
<tr>
<td>SOM</td>
<td>-0.264</td>
<td>-0.050</td>
</tr>
<tr>
<td>pH</td>
<td>-0.186</td>
<td>0.120</td>
</tr>
<tr>
<td>PoxC</td>
<td>-0.221</td>
<td>-0.173</td>
</tr>
<tr>
<td>Shoot mass</td>
<td>-0.211</td>
<td>0.199</td>
</tr>
<tr>
<td>Root mass</td>
<td>0.025</td>
<td>0.025</td>
</tr>
<tr>
<td>Shoot:root ratio</td>
<td>-0.111</td>
<td>0.031</td>
</tr>
<tr>
<td>Plant C</td>
<td>-0.217</td>
<td>0.230</td>
</tr>
<tr>
<td>Plant N</td>
<td>-0.118</td>
<td>0.145</td>
</tr>
<tr>
<td>Plant C:N</td>
<td>-0.097</td>
<td>0.034</td>
</tr>
<tr>
<td>Plant P</td>
<td>-0.190</td>
<td>0.250</td>
</tr>
<tr>
<td>Plant K</td>
<td>-0.253</td>
<td>0.055</td>
</tr>
<tr>
<td>Plant Ca</td>
<td>-0.251</td>
<td>-0.074</td>
</tr>
<tr>
<td>Plant Mg</td>
<td>-0.112</td>
<td>0.347</td>
</tr>
<tr>
<td>Plant S</td>
<td>-0.213</td>
<td>-0.117</td>
</tr>
<tr>
<td>Plant Mn</td>
<td>0.165</td>
<td>0.284</td>
</tr>
<tr>
<td>Plant Cu</td>
<td>0.012</td>
<td>-0.356</td>
</tr>
<tr>
<td>Plant B</td>
<td>-0.185</td>
<td>-0.114</td>
</tr>
<tr>
<td>Plant Zn</td>
<td>0.248</td>
<td>-0.124</td>
</tr>
<tr>
<td>Plant Na</td>
<td>0.006</td>
<td>0.234</td>
</tr>
</tbody>
</table>
Table 2. Correlations among soil physicochemical properties based on the Pearson correlation coefficient.
* = p<0.05, ** = p< 0.01, *** = p<0.001.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>C</th>
<th>NO₃-N</th>
<th>P</th>
<th>K</th>
<th>Na</th>
<th>Ca</th>
<th>Mg</th>
<th>CEC</th>
<th>OM</th>
<th>pH</th>
<th>PoxC</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1.00</td>
<td>0.96***</td>
<td>0.46***</td>
<td>0.66***</td>
<td>0.84***</td>
<td>0.79***</td>
<td>0.64***</td>
<td>-0.04</td>
<td>0.37**</td>
<td>0.82***</td>
<td>0.25</td>
<td>0.84***</td>
</tr>
<tr>
<td>C</td>
<td>1.00</td>
<td>0.26</td>
<td>0.50***</td>
<td>0.90***</td>
<td>0.71***</td>
<td>0.76***</td>
<td>-0.02</td>
<td>0.45***</td>
<td>0.90***</td>
<td>0.41**</td>
<td>0.89***</td>
<td></td>
</tr>
<tr>
<td>NO₃-N</td>
<td>1.00</td>
<td>0.62***</td>
<td>0.32*</td>
<td>0.56***</td>
<td>0.22</td>
<td>-0.44***</td>
<td>-0.24</td>
<td>0.06</td>
<td>-0.29*</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>1.00</td>
<td>0.57***</td>
<td>0.55***</td>
<td>0.36**</td>
<td>0.13</td>
<td>0.35**</td>
<td>0.43**</td>
<td>0.12</td>
<td>0.35**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>1.00</td>
<td>0.63***</td>
<td>0.88***</td>
<td>0.02</td>
<td>0.55***</td>
<td>0.85***</td>
<td>0.61***</td>
<td>0.70***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Na</td>
<td>1.00</td>
<td>0.53***</td>
<td>-0.13</td>
<td>0.23</td>
<td>0.55***</td>
<td>0.22</td>
<td>0.59***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ca</td>
<td>1.00</td>
<td>-0.11</td>
<td>0.48***</td>
<td>0.76***</td>
<td>0.60***</td>
<td>0.67***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mg</td>
<td>1.00</td>
<td>0.81***</td>
<td>0.25</td>
<td>0.51***</td>
<td>-0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CEC</td>
<td>1.00</td>
<td>0.68***</td>
<td>0.79***</td>
<td>0.36*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OM</td>
<td>1.00</td>
<td>0.55***</td>
<td>0.76***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pH</td>
<td>1.00</td>
<td>0.34**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PoxC</td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Correlations among plant variables based on the Pearson correlation coefficient. * = p<0.05, ** = p<0.01, *** = p<0.001.

<table>
<thead>
<tr>
<th></th>
<th>Shoot biomass</th>
<th>C</th>
<th>N</th>
<th>P</th>
<th>K</th>
<th>Mg</th>
<th>S</th>
<th>Mn</th>
<th>Ca</th>
<th>B</th>
<th>Zn</th>
<th>Na</th>
<th>C:N</th>
<th>Root biomass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoot biomass</td>
<td>1.00</td>
<td>0.75**</td>
<td>0.39**</td>
<td>0.46**</td>
<td>0.69**</td>
<td>0.53**</td>
<td>0.31**</td>
<td>-0.01</td>
<td>-0.38**</td>
<td>0.27*</td>
<td>-0.20</td>
<td>0.38**</td>
<td>-0.01</td>
<td>0.33**</td>
</tr>
<tr>
<td>C</td>
<td>1.00</td>
<td>0.58**</td>
<td>0.80**</td>
<td>0.73**</td>
<td>0.54**</td>
<td>0.71**</td>
<td>0.32*</td>
<td>-0.04</td>
<td>-0.57***</td>
<td>0.55***</td>
<td>-0.20</td>
<td>0.31*</td>
<td>-0.21</td>
<td>0.36**</td>
</tr>
<tr>
<td>N</td>
<td>1.00</td>
<td>0.53**</td>
<td>0.52**</td>
<td>0.33*</td>
<td>0.40**</td>
<td>0.06</td>
<td>0.03</td>
<td>-0.44**</td>
<td>0.13</td>
<td>-0.58**</td>
<td>0.40**</td>
<td>-</td>
<td>0.59**</td>
<td>0.30*</td>
</tr>
<tr>
<td>P</td>
<td>1.00</td>
<td>0.72**</td>
<td>0.53**</td>
<td>0.84**</td>
<td>0.42**</td>
<td>0.10</td>
<td>-0.47***</td>
<td>0.17</td>
<td>0.39**</td>
<td>0.17</td>
<td>-0.05</td>
<td>0.33**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>1.00</td>
<td>0.78**</td>
<td>0.55**</td>
<td>0.42**</td>
<td>0.69***</td>
<td>-</td>
<td>0.13</td>
<td>0.60***</td>
<td>-</td>
<td>0.17</td>
<td>-0.09</td>
<td>0.38**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mg</td>
<td>1.00</td>
<td>0.20</td>
<td>0.81**</td>
<td>0.55**</td>
<td>-0.43**</td>
<td>-0.11</td>
<td>0.04</td>
<td>0.52***</td>
<td>-0.14</td>
<td></td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>1.00</td>
<td>0.59***</td>
<td>0.45***</td>
<td>0.25</td>
<td>-0.08</td>
<td>0.31*</td>
<td>-0.19</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mn</td>
<td>1.00</td>
<td>-0.58**</td>
<td></td>
<td>0.53***</td>
<td>0.20</td>
<td>0.37**</td>
<td>-0.16</td>
<td>0.07</td>
<td>-0.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cu</td>
<td>1.00</td>
<td>0.15</td>
<td>0.35**</td>
<td>-</td>
<td>0.00*</td>
<td>-0.10</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>1.00</td>
<td>0.00</td>
<td>-</td>
<td>0.49***</td>
<td>0.26*</td>
<td>-0.01</td>
<td>0.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zn</td>
<td>1.00</td>
<td>0.08</td>
<td>-</td>
<td>0.46***</td>
<td>0.18</td>
<td>-0.37**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Na</td>
<td>1.00</td>
<td>-0.31*</td>
<td>0.11</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C:N</td>
<td>1.00</td>
<td>0.00</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Alpha diversity of bacterial and fungal communities by site and management. Values reported are Shannon index ± standard error.

<table>
<thead>
<tr>
<th></th>
<th>MR</th>
<th>RR</th>
<th>PF</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S</td>
<td>3.50 ± 0.14</td>
<td>2.90 ± 0.09</td>
<td>2.72 ± 0.05</td>
</tr>
<tr>
<td>ITS</td>
<td>2.60 ± 0.08</td>
<td>2.62 ± 0.18</td>
<td>2.68 ± 0.08</td>
</tr>
</tbody>
</table>