



Bacterial communities of the *Salvia lyrata* rhizosphere explained by spatial structure and sampling grain

Jonathan R. Dickey¹ · James A. Fordyce¹ · Sarah L. Lebeis²

Received: 26 October 2019 / Accepted: 31 August 2020 / Published online: 5 September 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Advancements in molecular technology have reduced the constraints that the grain of observation, or the spatial resolution and volume of the sampling unit, has on the characterization of plant-associated microbiomes. With discrete ecological sampling and massive parallel sequencing, we can more precisely portray microbiome community assembly and microbial recruitment to host tissue over space and time. Here, we differentiate rarefied community richness and relative abundance in bacterial microbiomes of *Salvia lyrata* dependent on three spatial depths, which are discrete physical distances from the soil surface within the rhizosphere microhabitat as a proxy for the root system zones. To assess the impact of sampling grain on rarefied community richness and relative abundance, we evaluated the variation of these metrics between samples pooled prior to DNA extraction and samples pooled after sequencing. A distance-based redundancy analysis with the quantitative Jaccard distance revealed that rhizosphere microbiomes vary in richness between rhizosphere soil depths. At all orders of diversity, rarefied microbial richness was consistently lowest at the deepest samples taken (approximately 4 cm from soil surface) in comparison with other rhizosphere soil depths. We additionally show that finer grain sampling (i.e., three samples of equal volume pooled after sequencing) recovers greater microbial richness when using 16S rRNA gene sequencing to describe microbial communities found within the rhizosphere system. In summary, to further elucidate the extent host-specific microbiomes assemble within the rhizosphere, the grain at which bacterial communities are sampled should reflect and encompass fine-scale heterogeneity of the system.

Keywords Rhizosphere microbiomes · 16S rRNA amplicon sequencing · Indicator species analysis · Hill numbers · Alpha diversity · *Salvia*

Introduction

Plants transform the soil that surrounds their roots (i.e., the rhizosphere) with primary metabolites like amino acids, secondary metabolites like alkaloids, and macronutrients through litter deposition [1]. Rhizosphere modifications by plants influence critical relationships with microorganisms, such as fungi and bacteria, which utilize the rhizosphere niche [2].

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00248-020-01594-7>) contains supplementary material, which is available to authorized users.

✉ Jonathan R. Dickey
jdickey9@vols.utk.edu

¹ Department of Ecology and Evolutionary Biology, The University of Tennessee, 569 Dabney Hall, Knoxville, TN 37996, USA

² Department of Microbiology, The University of Tennessee, 307 Ken and Blaire Mossman Bldg., Knoxville, TN 37996, USA

Plant-microbe interactions exist over multiple plant and microbial generations, establishing complex plant-soil feedbacks (PSFs) [3]. Microbiota mediate available carbon and nitrogen directly in the rhizosphere, in turn affecting plant defense strategies, productivity, and fitness [1, 4–6]. Measuring microbial community richness and abundance of rhizosphere microbes through massive parallel sequencing illuminates intricacies of how microbes moderate nutrient cycling and interact with plants [7, 8]. Further, the grain, or the defined individual microbiome sampling unit amongst rhizosphere studies, is often variable and may not consistently capture spatial variation of bacterial communities found within the rhizosphere. Yet within the rhizosphere, spatial heterogeneity of microbial communities is poorly understood [9] necessitating a field-wide standard of resolution for comparison among disparate studies. By incorporating spatial structure into sampling protocols, microbiome richness in the rhizosphere can be captured as a transitional gradient of species distributions responding to root maturity and exudates across plant taxa.

Variation in community richness and abundance of rhizosphere microbiomes can be partially explained by numerous biotic interactions and abiotic conditions. For example, plant species receive microbes from bulk soil, forming distinct rhizosphere communities over time [10, 11]. Plant genotypes may regulate rhizodeposition and decomposition rates of plant-derived organic matter and can serve as strong drivers of rhizosphere microbial richness [12–14]. Plant-derived organic matter affects soil properties such as pH, organic carbon (C) content, moisture, and nitrogen (N) availability, which directly shape how microorganisms colonize unique microhabitats within soils across large geographic scales [15, 16]. Further, N-cycling is considerably more specialized within the rhizosphere than in bulk soils leading to their unique species composition [17]. This contributes to the delicate balance between persistent and labile mineralized forms of C and N concentrations within the rhizosphere, effectively pulling resource-dependent, microbes towards these sources at the root [18].

A contributing factor in the structure of microbial communities of the rhizosphere may be the spatial relationship with root developmental age. The root can be divided into several spatial components based on developmental maturity, also known as root zones. Yang and Crowley [19] showed that eubacterial communities vary in respect to developmental zones along the root by measuring community structure with a PCR-denaturing gel electrophoresis. They also discovered that community richness within each zone is dependent upon differences among soil types. Therefore, an amalgam of host effects and abiotic conditions in the rhizosphere environment are strong determinants of microbial survival and assemblages [20]. With a firm understanding of how host-associated microbial communities respond to these selective agents, it is imperative to further investigate how spatial sampling within the rhizosphere contributes to heterogeneity in bacterial richness recovered with high-resolution 16S rRNA gene sequencing.

Developmental root zones can serve as a means to understand the spatial structure and maintenance of community-level heterogeneity in the context of plant rhizosphere-microbiome interactions [21–23]. The development of root tissue is regulated by plant or bacterial produced auxins and environmental factors like soil aggregates [24, 25]. Each root zone harbors unique microbial reservoirs due to favorable plant compounds, soil nutrients, and limited competition among other microbes [26, 27]. Most studies using a pooled, coarse sampling resolution to characterize plant-associated microbiomes fail to capture this heterogeneity across root zones, or corresponding rhizosphere microhabitats. Thus, combining multiple, fine-grain samples from different root zones or from multiple depths within the rhizosphere might explain scale-dependent variation of the rhizosphere microbiome.

In this study, we sampled from three depths, or distances from the soil surface, within the rhizosphere microhabitat of *Salvia lyrata* to ask: How does rarefied bacterial richness and relative abundance compare between discrete rhizosphere depths and a combined rhizosphere sample from the same plant? Further, we ask how does rarefied bacterial richness, the raw number of reads, and relative abundance in reads compare between differences in sampling grain? To achieve this, we utilized two types of pooling methods to examine the impact of sampling resolution. These were pooling rhizosphere samples from each depth prior to DNA extraction (*in vitro*) and combining community data tables post sequencing and processing (*in silico*) (Fig. 1). Our overarching hypothesis is that rhizosphere bacterial richness and relative abundance correlates with spatially dependent sampling conditions that exist along root depth in the rhizosphere environment. We predict that microbiomes from each rhizosphere depth capture more information on rare species (i.e., rare across all samples) than a single coarse grain representative sample from the same rhizosphere. Additionally, we predict that sampling resolution will impact the characterization of rarefied bacterial community richness via differences in 16S rRNA gene capture, where coarse resolutions favor predominant groups of bacteria found in highly diverse samples. Finally, the following analyses provide a needed spatial characterization of rhizosphere microbiota that associate with the genus *Salvia*, which contains many species utilized for medicinal purposes.

Materials and methods

Study system

Salvia lyrata L., or lyreleaf Sage, is a flowering, biennial mint with a shallow fibrous root ball consisting of roughly 8–12 individual primary roots, and pinnately lobed basal rosette leaves. In general, microbiomes associated with the genus *Salvia* have been underexplored. Recent research efforts have characterized seed and foliar microbiomes associated with the genus *Salvia*. However, little to no attention has been given to rhizosphere microbiomes [28, 29]. Field sampling occurred at Forks of the River Wildlife Management Area (35°56'30.6" N, 83°50'43.3" W) in Knox County, Tennessee during July 1–July 5, 2018. The plant community at this site has been historically disturbed and consists primarily of perennial grasses and mixed hardwoods along the edge of trails where *S. lyrata* is often found. The soil material found here is characterized as both highly weathered, acidic red clay, and brown silty clay loam. Three 25-m transects were drawn, and 30 plants were flagged based on occurrence within 1 m of this line. Plant individuals ($N = 10$) were chosen per transect for rhizosphere microbiome sampling by using a random number generator.

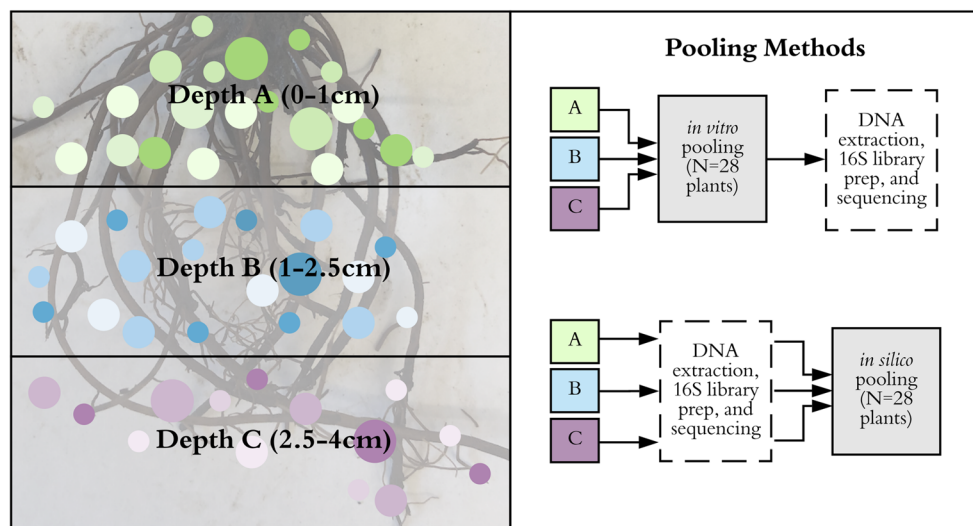


Fig. 1 A conceptual outline describing the spatial structure of the *S. lyrata* rhizosphere microbiome outlined by three depths within the rhizosphere. Colors and variation in circle size illustrate different microbes and relative abundances respectively. Depth A represents microbiomes from the surface of the system from 0 to 1 cm deep. Depth B characterizes microbiomes from 1 to 2.5 cm. Depth C

represents the microbial community found within the rhizosphere at 2.5–4 cm. This figure also illustrates how sample rhizosphere depths from the same plant were pooled via the two methods, *in vitro* (physical samples prior to DNA extraction and 16S library preparation) and *in silico* (community data tables).

All microbial rhizosphere samples were collected from juvenile plant individuals of *S. lyrata* to control for phenological and temporal effects on the colonization of host-specific microbiomes. Juvenile individuals were identified by the absence of a racemose scape.

Rhizosphere microbiome sampling

Here, we defined the rhizosphere of *S. lyrata* as the narrow region of soil spanning from the root epidermis to the soil that is influenced by root exudates (~ 3 mm from epidermal tissue) within the same root system [30, 31]. All plant rhizospheres were sampled in the field at three different depths as a proxy for the root system zones. Sample depths in the rhizosphere were from 0 to 1 cm (depth A), from 1 to 2.5 cm (depth B), and from 2.5 to 4 cm (depth C) aimed to represent the developmental root zones within the root system, accounting for variation in root length and root density across individuals (Fig. 1). For each sampling depth, the entire root system was exposed, and rhizosphere soil was scrapped from the root epidermis to 3–5 mm away from the surface of the root into sterile Nasco Whirl Paks across all visible roots (Nasco, Fort Atkinson, WI, USA). This occurred first at depth A and then at each subsequent sampling soil depth using a sterilized spatula. Sampling tools were sterilized between sample units to avoid contamination. Rhizosphere soils were stored on ice immediately after the sterile Nasco Whirl Paks were sealed throughout the daily sampling period. Rhizosphere microbiome samples were stored at – 80

°C at the University of Tennessee, Knoxville, until further processing. Short-term storage of dry soil microbiome samples in – 80 °C does not affect the microbial community structure [32].

Sampling grain: pooling *in vitro* versus *in silico*

We used two different pooling methods to compare how sampling grain, or the spatial resolution and volume that defines a microbiome sampling unit, affects the microbial community recovered through 16S rRNA gene amplicon sequencing (Fig. 1). The first method combined an equal portion of soil from each rhizosphere depth samples (A, B, and C) into a single sample to represent the whole rhizosphere microbiome of each plant (*in vitro*). To do so, samples were removed from deep freeze and set on ice for an hour to thaw. One gram from each of the three sampling depths was combined into a new sterile Nasco Whirl Pak. Prior to subsampling 0.25 g for DNA extraction, we mixed the pooled samples together by vigorously shaking the samples by hand for 30 s and leaving them at room temperature for an hour. The second pooling method occurred after sequencing and combined read counts for the same amplicon sequence variants (ASVs) for each rhizosphere depth per individual plant. These samples represent the pooled microbiome *in silico*. Prior to sample rarefaction, we accounted for differences in sampling effort between pooling methods by dividing read count tables by three for samples pooled *in silico*. Pooling methods were compared and used to describe the composite microbiome found within the rhizosphere of *S. lyrata*.

16S rRNA gene library preparation for Illumina MiSeq

DNA was extracted from 0.25 g per sample using the DNeasy PowerSoil Kit (MoBio Laboratories, Inc., Carlsbad, CA, USA) with a modification at Step 16 using 50 µL of Solution C6 and an elution time of 5 min before the final centrifuge step. This extraction kit successfully removes all non-DNA organic and inorganic material, such as humic acids at Step 10 and Step 11. PCR reagents were sourced through the KAPA HiFi HotStart Ready Kit (KAPA Biosystems, Wilmington, MA, USA). Within the amplification PCR, the bacterial-specific primer pair 341F and 785R (5 µL at 1 µM each) was used to target the V3–V4 region of the 16S rRNA gene and to prepare our DNA samples for the second index PCR. The sequences for 341F and 785R are 5'-CCTA C G G G N G G C W G C A G - 3' and 5' - G A C T A C H V G G G T A T C T A A T C C - 3' respectively [33, 34]. DNA extractions were thawed on ice, and 2.5 µL of each sample was used for this PCR. For each plate, a PCR negative blank was used to check for PCR contamination. Plates were ran under the following PCR program: an initial denaturing step of 95 °C for 3 min, 25 cycles of 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s; a final elongation cycle of 72 °C for 5 min; and then held at 4 °C. This PCR produced 25 µL of the 16S rRNA amplicon for each sample that was stored at - 20 °C until further use. Samples were then run on a 2% agarose gel with ethidium bromide and Tris-acetate-EDTA (TAE) buffer to confirm amplicon length and purity. All PCR products (25 µL for each sample) were purified with 20 µL of Agencourt AMPure XP magnetic beads (Beckman Coulter, Brea, CA, USA) and two 200 µL washes of freshly prepared 80% ethanol. Purified 16S rRNA gene amplicons for all samples were then resuspended with 50 µL of Tris-HCl and stored at - 20 °C.

The second index PCR attached forward and reverse index primers that are designed specifically for the multiplexing samples and sequencing with the MiSeq instrument (Illumina Corporation, San Diego, CA, USA). Each sample in the library was provided with its own unique primer combination for sample identification in downstream quantitative analyses (Online Resource 1). These primers were from the i5/i7 Nextera XT Index Kit A and D with 5-µL combinations of 8 bases long forward and reverse segments. Further, 25 µL of KAPA HiFi HotStart Ready Mix was used per reaction. Five microliters of purified PCR product from the first reaction was used to reach sufficient concentration of the bacterial 16S rRNA V3–V4 gene region at this step. For this index PCR, 10 µL of PCR grade water was added per sample, bringing the total volume to 50 µL. The index PCR performed had an initial cycle of 95 °C for 3 min followed by an 8-cycle sequence of 95 °C for 3 s, 55 °C for 30 s, and 72 °C for 30 s. The last elongation cycle was held at 72 °C for 5 min.

Following the standard index PCR clean up provided by Illumina, index PCR products (50 µL for each sample) were purified using 56 µL of Agencourt AMPure XP magnetic beads and two 200-µL washes of freshly prepared 80% ethanol. The final 16S rRNA gene amplicon products were then combined with 25 µL of Tris-HCl and stored at - 20 °C until further use. Post two-step PCR, all 16S amplicon products were quantified using a NanoDrop spectrophotometer (NanoDrop Products, Wilmington, DE, USA) and pooled in equal concentrations. Next, a bioanalyzer was used to determine molarity and appropriate adapter ligation by examining product length (Agilent Technologies, Santa Clara, CA, USA). This step behaves as the final quality control step before sequencing. Pooled samples were diluted to 4 pM with Tris and loaded into the MiSeq instrument with 20% PhiX on a V3 flow cell set to read a 2 × 275, or paired end, cycle at the University of Tennessee Genomics Core, Knoxville, Tennessee, USA.

Bioinformatics

Processing of the 16S amplicon data were completed with the open source pipeline DADA2 version 1.8 (<https://github.com/benjjneb/dada2>). With DADA2, we detected sequence anomalies at a single-nucleotide resolution known as ASVs [35]. Amplification primers 341F and 785R were removed from each read by trimming according to their length. Then, all forward and reverse reads were quality filtered and trimmed to 248 and 220 nucleotides in length, respectively. Sequences with sequencing errors were then removed, and forward and reverse reads were merged if there was sufficient overlap for each ASV in each sample. Chimeras were detected within each merged read and removed. Lastly, all merged sequences were matched to SILVA version 132, a taxonomic reference database for 16S rRNA genes [36, 37]. This pipeline was implemented in R version 3.5.2 (<https://www.r-project.org/>). All nucleotide sequence data was made publicly available in GenBank database under the accession number PRJNA575901.

The package phyloseq was then used to combine community data, assigned taxonomy, and metadata into a single, flexible R object [38]. Eukaryotes, chloroplasts, and mitochondria were removed from our dataset. If no assignment was made at the Kingdom level, these ASVs were similarly removed. Lastly, rhizosphere samples that failed during sequencing and contained no ASVs were dropped ($N = 2$). Mean sample abundances were obtained by rarefying to the lowest number of reads (7739 reads) 5000 times with the `rrarefy` function in the `vegan` package when we investigate the difference between ASV richness and relative abundance amongst rhizosphere depths [39]. To address the difference in relative abundance between pooling methods (i.e., sampling grain), pairwise rarefaction to the same number of reads was used per plant.

Hypothesis testing and statistical analyses

An indicator species analysis (R package *indicspecies*) using the group equalized point biserial correlation index for each group was then conducted with 50,000 permutations to elucidate ASVs highly associated with each rhizosphere depth [40]. Two distance-based redundancy analyses (dbRDA) were implemented for hypothesis testing to quantify community variation explained by the sampling depth within the rhizosphere. In these models, we partial out variation explained by individual plants (sample ID) before hypothesis testing in order to account for the non-independence of rhizosphere samples associated with a single plant. Prior to modeling, all sequences were aligned using MAFFT version 7.0 [41]. Next, a phylogenetic tree was built using the Randomized Accelerated Maximum Likelihood (RAxML) program and the general time reversible CAT model of rate heterogeneity for phylogenetic approximation [42]. This tree was made ultrametric using the software treePL, which transformed branch lengths to correspond with relative divergence time using penalized likelihood [43]. This ultrametric tree was then used to create a weighted UniFrac distance matrix. The quantitative Jaccard (Ružička) and the weighted UniFrac distance were then used as response variables in our dbRDAs ($N = 87$ samples). We then evaluated the statistical significance of rhizosphere depth as a predictor of community dissimilarity via 10,000 permutations of our data using the *anova.cca* function in the R package *vegan* 2.5-5 [39].

The effective number of bacterial species (qD) associated with each rhizosphere depth and each pooled sample was described using a Hill numbers approach with the R package *vegan* 1.2 [44–46]. Since the numbers equivalent of ASVs (D) is sensitive to the order of diversity (q), or the balance between rare and common ASVs, we chose to assess alpha and beta diversity from $q = 0$ to 5. There are many frequently used diversity indices that are analogous to qD . For example, at $q = 0$, rare and common ASVs are weighted equally and therefore represent an incident-based measure of species richness. When $q = 1$, ASVs are weighted by their proportional abundance, which is congruent to the exponential of Shannon's entropy index, and at $q = 2$, rare species are down weighted and correspond to the inverse of Simpson's index. As q increases beyond this, rare ASVs are further down weighted from the analyses of alpha and beta diversity.

To understand if alpha diversity significantly varies based on depth within the rhizosphere, fitted linear-mixed effects models with sample ID, or the individual plant itself, as a random effect were constructed at each Hill number using the *lme4* package [47]. A Type II Wald X^2 test of variance was used for hypothesis testing on main effects. For significant results, a Tukey's honest significant difference (HSD) multiple comparisons test was implemented for all pairwise comparisons of rhizosphere depths to investigate where differences in alpha diversity lie.

Two different methods were used to assess beta diversity measures at each Hill number. The first method is similar to a Levene's test in multivariate space, which analyzed the homogeneity of dispersion in community composition between groups and can be used as a proxy for beta diversity [48]. A pairwise distance matrix using the distance from each observation to the group centroid was then fed into a fitted linear mixed-effects model with sample ID as a random effect. Secondly, to address if community variation within rhizosphere depths is similar to one another, a dbRDA on the pairwise distance matrix while conditioning for variation among plants was implemented. A permutational ANOVA was then used to compare this variation between groups (10,000 permutations).

We evaluated the difference in the standardized number of raw reads between samples pooled *in vitro* and samples pooled *in silico* through a paired t test for each plant ($N = 28$ plants). Standardization here means that we controlled for differences in sampling effort between the two pooling methods by dividing the read abundances by three, the number of rhizosphere depths sampled, for each sample pooled *in silico*. Additionally, all samples sequenced held the same fraction of size throughout library preparation. This paired t test used raw read counts that were summed across rows, or plants, in our ASV table. To observe differences in the relative abundance of reads between pooling methods, a paired t test was also conducted using pairwise rarefied data ($N = 28$ plants). Since two rhizosphere depth samples failed during sequencing and equal comparisons across all plants and pooling methods cannot be applied, two plants were excluded from these analyses. For each plant replicate, alpha diversity for $q = 0-5$ was calculated on the mean rarefied community data from 5000 permutations. Then, a paired t test ($N = 28$) was performed to show differences in ASV richness between pooling methods. To evaluate if dispersion between pooling methods is similar for the number of raw reads, relative abundance in reads, and for alpha diversity, a paired Brown-Forsythe test was implemented.

Results

The *Salvia lyrata* rhizosphere microbiome

From the 118 samples sequenced and analyzed, 27,314,720 raw sequence reads were generated from the Illumina MiSeq run. Dereplicated sequences numbered at 3,758,012 reads attributing to 25,123 ASVs with associated abundances for Bacteria and Archaea. For samples that were pooled *in vitro* and represent a homogenous characterization of *S. lyrata*'s rhizosphere, a total of 1,322,876 reads were recovered. These reads correspond to 20 bacterial phyla. The predominant (i.e. more frequent detected, abundant) phyla of Bacteria that associate with this rhizosphere system are *Proteobacteria*

(30%), *Actinobacteria* (24%), *Planctomycetes* (13%), *Verrucomicrobia* (9%), *Bacteroidetes* (8%), and *Acidobacteria* (7%) (Fig. 2). Firmicutes characterized 0.9% of the bacterial microbiome from samples that were pooled *in vitro*. On a finer taxonomic level, the family *WD2101_soil_group*, a Planctomycete, was a highly abundant family, containing 7% of all sequenced reads from pooled samples *in vitro*. The next two predominant bacterial families within this rhizosphere system were *Xanthobacteraceae* (6%) and *Burkholderiaceae* (2%), both in the phylum *Proteobacteria*. On the species level, the most abundant taxon was *Candidatus-Xiphinematobacter*, a verrucomicrobial endosymbiont commonly found in rhizosphere environments.

Microbiomes are spatially structured by depth within the rhizosphere

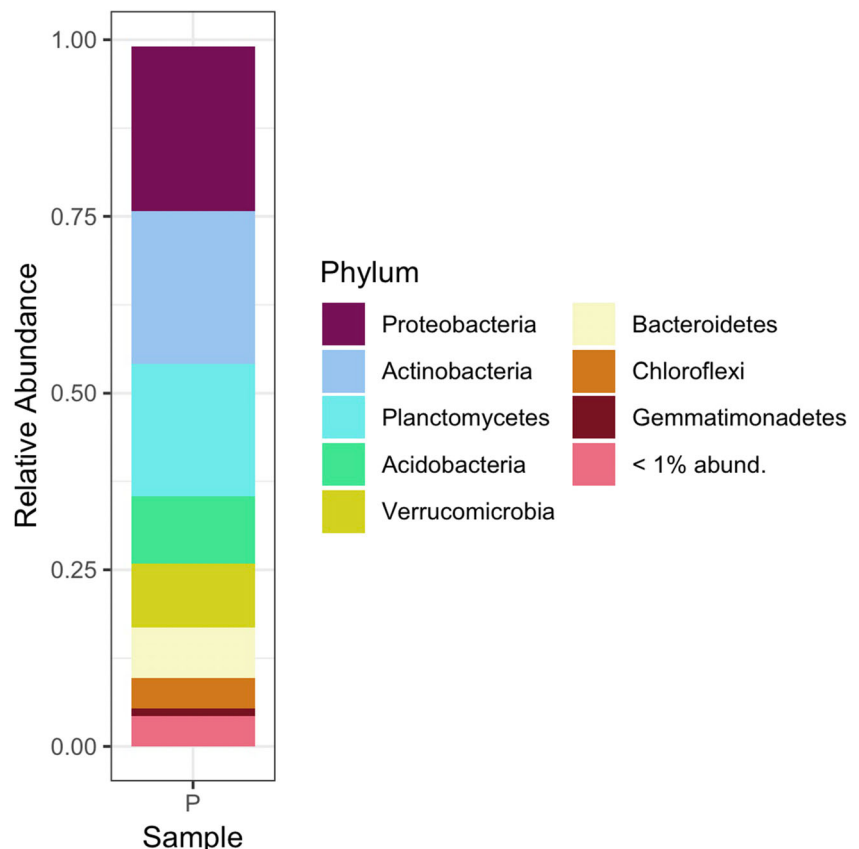
We used two community distance matrices in our multivariate models to evaluate the spatial complexity of bacterial microbiomes as predicted by the rhizosphere sampling depths described in Fig. 1. These models used rarefied and standardized community data to test the strength of depth within the rhizosphere as a predictor variable. With the quantitative Jaccard distance matrix as a response variable, rhizosphere depth explained 3.65% of total variation in bacterial

community dissimilarity. The conditioned variable, sample ID, explained 52.7% of total variation in community dissimilarity. After removing the effect of sample ID, the proportion of remaining variance explained by rhizosphere depth is 7.71% (Fig. 3). Distinctive clustering around the respective centroid occurred within each group on both axes. Richness and relative abundance between depths varied based on the quantitative Jaccard distance matrix provided in our permutational analysis of variance ($p < 0.001$).

Our second distance-based redundancy analysis took into account phylogenetic autocorrelation by using a weighted UniFrac distance as the response variable. This model showed rhizosphere depth to explain 5.51% of the total variation in bacterial microbiome dissimilarity. Sample ID in this model explained 12.2% of total variation in community dissimilarity. Like our previous model, we also partial out this effect before hypothesis testing, and rhizosphere depth then explained 6.29% of the remaining variation in community dissimilarity. The permutational ANOVA on our model with a weighted UniFrac distance showed that each rhizosphere depth is significantly different from one another in terms of rarefied bacterial richness ($p < 0.001$). Additional information on these model summaries and subsequent ANOVAs are shown in Online Resources 2 and 3.

Alpha diversity was examined using a Hill numbers approach to gather and compare the effective number of “types”

Fig. 2 Bacterial Phyla that associate with *Salvia lyrata* for samples pooled *in vitro*. On average, Proteobacteria represent 30% of all bacteria recovered. Actinobacteria, Planctomycetes, Verrucomicrobia, and Bacteroidetes cover 24%, 13%, 9%, and 8% respectively. Phyla represented less than 1% in this system have been combined together.



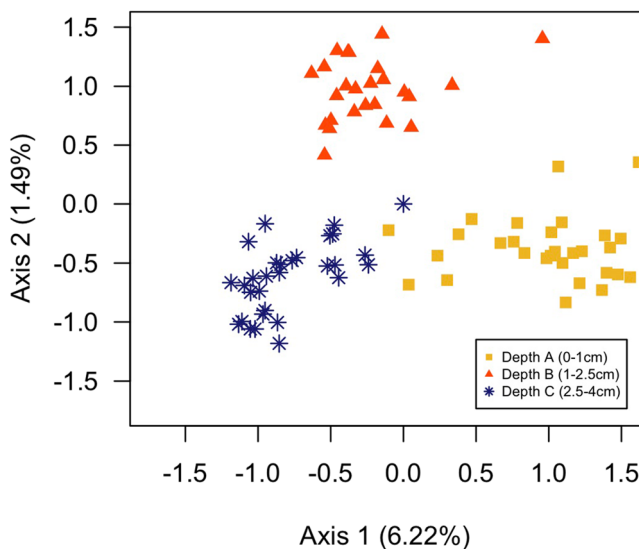


Fig. 3 Principal coordinate analysis on the reduced dimensional analysis based on a quantitative Jaccard distance with rhizosphere depth as a predictor variable. Constrained axis together explains 3.532% of total variation in multivariate space

or ASVs between each rhizosphere depth. A Type II Wald's X^2 test on fitted linear mixed-effects model at $q = 4, 5$ showed that alpha diversity was distinct and varied across rhizosphere depth ($q = 4$: $X^2 = 6.04$, $p = 0.0487$, $df = 2$; $q = 5$: $X^2 = 6.84$, $p = 0.0327$, $df = 2$). Table 1 shows a full summary of each analysis of variance on all orders of alpha diversity. Tukey's HSD revealed that at these orders, depths B and C are driving the differential pattern over other comparisons between rhizosphere depths. Further, alpha diversity across $q = 0-3$ on average is greatest at depth A (Fig. 4). At $q = 4-5$, depth B has the greatest alpha diversity out of all rhizosphere depths. Samples pooled *in vitro* contain on average the greatest effective number of ASVs when at $q = 0-3$, however, while continually down-weighting rare species from our analysis pooled samples lie between depths B and C. Figure 4 also shows that depth C consistently has the least effective number of ASVs than any other group.

Analysis of multivariate dispersion showed greater heterogeneity in rarefied community richness and relative

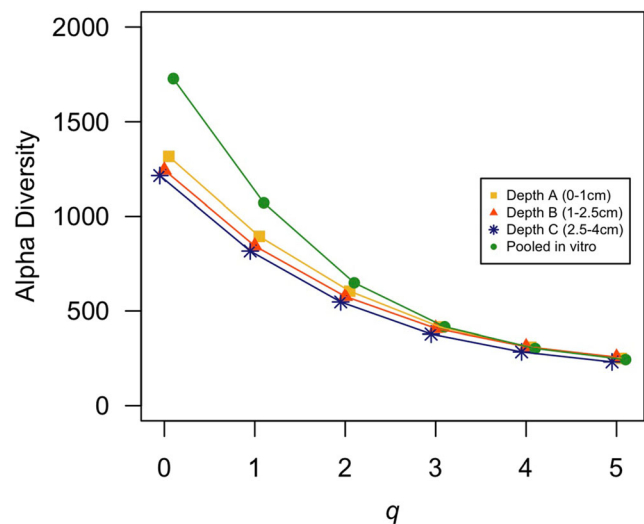


Fig. 4 Alpha diversity profile plot showing intercepts from each fitted linear model a function of q . As rare ASVs are removed from the analysis at higher orders, rhizosphere depths become less distinct in terms of the effective number of species and converge

abundance between depths A, B, C at $q = 0$ ($F_2 = 3.29$, $p < 0.05$). As rare species are down weighted across increasing q , multivariate dispersion, or the average distance from each individual point to the multivariate centroid (median), was not different between rhizosphere depths. Thus, at lower orders of diversity, rare ASVs contributed more to the differences in community turnover between rhizosphere depths than common ASVs. Similarly, a dbRDA on community variation within groups while accounting for the effect of individual plants showed that each rhizosphere depth is significantly different in terms of turnover at $q = 0$ ($F_{2,55} = 3.25$, $p < 0.05$). Lastly, these analyses illuminated that depth A is the richest in ASVs while depth C is consistently the least rich.

An indicator species analysis was performed across our dataset to characterize which taxa are more strongly associated with the communities at each spatial depth within the rhizosphere of a plant root system described by Fig. 1. Depth A harbors the most diverse and numerous indicator taxa with 311 indicator ASVs. Commonly detected phyla that were represented as indicators for depth A are *Acidobacteria*, *Actinobacteria*, *Bacteroides*, *Chloroflexi*, *Gemmatimonadetes*, *Planctomycetes*, *Proteobacteria*, and *Verrucomicrobia*. As seen in Fig. 5, the most represented known families in this indicator analysis for depth A were *Gemmataceae* and *WD2101_soil_group*, both with an abundance of 17 in this analysis. Depth A has numerous rare families that are distinct indicators, with *Phycisphaeraceae* and *Xiphinematobacteraceae* identified as significant indicators ($p < 0.001$). Species identified as indicators for this rhizosphere depth are *Agromyces ramosus* (*Actinobacteria*), *Lechevalieria aerocolonigenes* (*Actinobacteria*), and *Arenimonas daejeonensis* (*Proteobacteria*). At the ASV level, 61 unique taxa were associated with depth B. Associated with this

Table 1 Summary of the Type II Wald's X^2 test performed on alpha diversity across depths within the rhizosphere at $q = 0$ to 5

Order of Diversity (q)	Chisq	Pr(> Chisq)
0	1.4384	0.4871
1	2.303	0.3162
2	3.9319	0.14
3	5.1613	0.07573
4	6.0449	0.04868*
5	6.8364	0.03277*

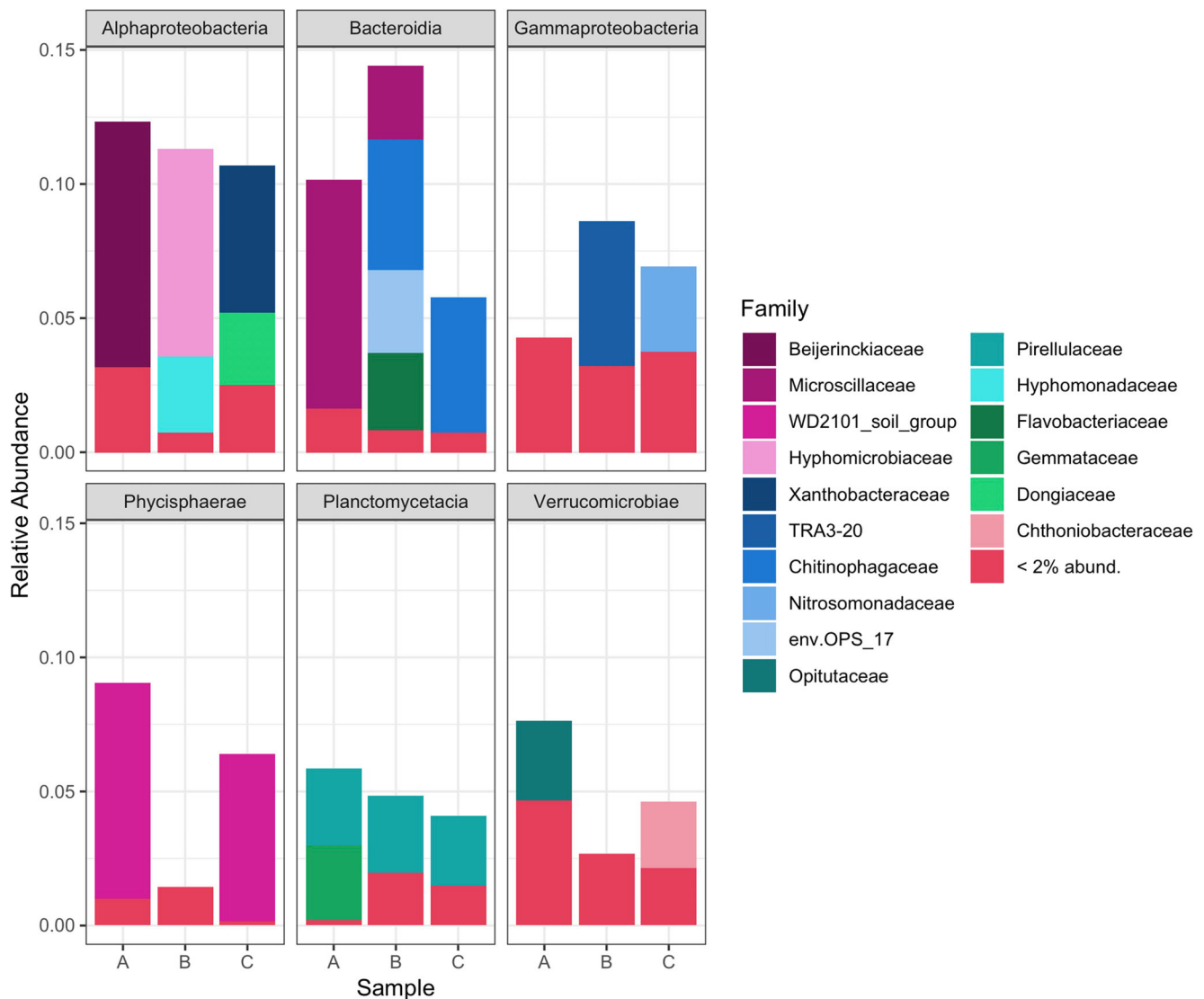


Fig. 5 Indicator species analysis visualized at the family level for each rhizosphere depth. Each depth is illustrated here as sample A (0–1 cm), sample B (1–2.5 cm), and sample C (2.5–4 cm). At the class level, *Alphaproteobacteria*, *Bacteroidia*, *Gammaproteobacteria*,

Phycisphaerae, *Planctomycetacia*, and *Verrucomicrobia* are represented. For simplicity, this figure combines all families that occur less than 2% in our indicator species analysis

rhizosphere depth, there are 10 known phyla, with *Bacteroidetes*, *Planctomycetes*, and *Proteobacteria* as the most abundant. The most common and known family as an indicator for depth B was *Hyphomicrobiaceae* (Fig. 5). Rare phyla uncovered as indicators for depth B are *Cyanobacteria*, *Latescibacteria*, and *Patescibacteria*. The strongest correlated indicator ASV for depth B came from the family *Desulfarculaceae*, a *Deltaproteobacteria* (r.g. corr = 0.242, $p < 0.01$). Depth C had 14 phyla as indicators representing 152 ASVs. Common phyla identified as indicators were *Actinobacteria* and *Proteobacteria*. The family *Xanthobacteraceae* (*Alphaproteobacteria*) was the most abundant family as an indicator representing 4.6% of all indicator ASVs for depth C. Additionally, depth C has many significant indicators at the Genus level. No indicator ASVs

were identifiable to the species level for depth B or C. Online Resource 4 further depicts r.g. correlations and relative significance for all indicator taxa that came from this analysis. This analysis shows that strong indicators for depth A come from the phylum *Planctomycetes*, while depth B and depth C are best predicted by the presence of *Deltaproteobacteria* and *Alphaproteobacteria* respectively.

Rarefied richness is greater when rhizosphere depths are pooled *in silico*

Pooling samples across rhizosphere depths per individual plant after sequencing captured a larger, more comprehensive image of bacterial community composition and diversity in comparison with pooling physical samples prior to DNA

extraction. To illuminate if there were differences in the rarefied richness and relative abundance of reads between the two pooling methods, paired t tests were implemented on the standardized number of raw reads, rarefied and standardized number of reads, and alpha diversity (Fig. 6). On average, the standardized raw number of reads for each rhizosphere *in vitro* on average were 17,353.8 reads greater than samples pooled *in silico* ($t = 3.46$, $df = 27$, $p < 0.001$). A paired Brown-Forsythe test confirmed that dispersion around the median between the two pooling methods was significantly different for the raw number of reads ($t = 6.65$, $df = 27$, $p < 0.001$). Similarly, the pairwise rarefied abundance in total reads was greater on average by 9073.23 for samples pooled *in vitro* ($t = 4.59$, $df = 27$, $p < 0.001$). Pairwise Brown-Forsythe tests illustrated that there was heterogeneity in dispersion of rarefied abundances in reads between pooling methods ($t = 9.27$, $df = 27$, $p < 0.001$). For $q = 0$ and $q = 4-5$, paired t tests showed significant differences ($\alpha = 0.05$) in the effective number of ASVs between pooling methods with samples pooled *in silico* having the greatest alpha diversity. The largest difference in alpha diversity between pooling methods occurred at $q = 0$, with a mean difference of 679.85 effective number of ASVs (Table 2). Figure 6 b shows the relationship between alpha diversity at $q = 0$ and each pooling method, where samples that were pooled *in silico* have on average a greater alpha diversity than samples pooled *in vitro*. Paired Brown-Forsythe tests on the dispersion of the effective number of ASVs (alpha diversity) were significantly different for $q = 0-2$ between the two pooling methods (Table 2). As rare species are further down weighted beyond $q = 2$, we failed to detect a difference in the dispersion between sampling grains.

Discussion

Summary

In this study, we demonstrate that (1) rarefied bacterial richness and relative abundance varies with rhizosphere depths. Alpha diversity among rhizosphere depths differs the most when rare ASVs are down weighted. Depth C is characterized as the least diverse amongst rhizosphere depths, while depth A is shown to be the richest. (2) By evaluating the spatial structure of bacterial communities within the rhizosphere of *Salvia lyrata*, we were able to obtain a more holistic characterization of the rhizosphere microbiome and show that samples from a coarser grain (pooled samples *in vitro*) have higher rarefied abundance in reads than samples from individual rhizosphere depths. (3) With increased fine grain sampling (i.e., pooling *in silico*), more rare ASVs are captured, and thus, a greater amount of spatial heterogeneity in microbial richness is observed. Lastly, samples pooled *in silico* have less dispersion in rarefied ASV richness and relative abundance in the number of reads than pooling *in vitro*.

Bacterial communities are spatial structured across rhizosphere depths

The rhizosphere is a highly context-dependent microhabitat with transient, diurnal cycles of rhizosphere soil temperature and moisture driving microbial community composition [49, 50]. Further, the structure of bacterial communities within the rhizosphere is partially explained by plant species, genotype, and developmental age [10, 12, 16]. Previous studies

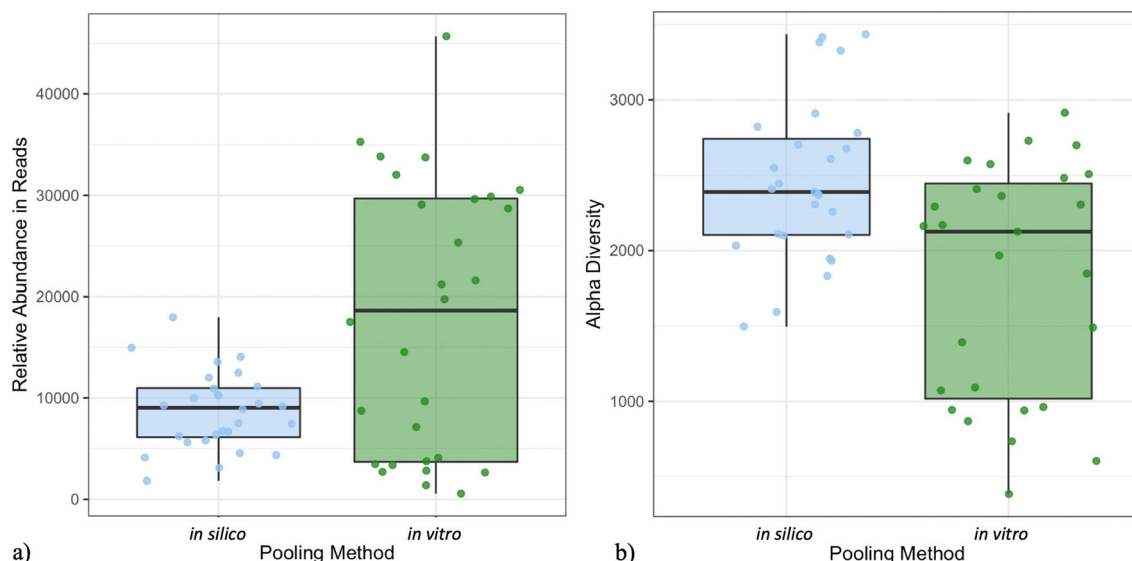


Fig. 6 **a** The total relative abundance of reads is greater for samples that were pooled *in vitro* compared with those *in silico*. **b** Alpha diversity at $q = 1$ for pooled samples *in silico* and *in vitro*. On average, samples pooled

in silico have greater alpha diversity than samples pooled *in vitro*. Further, *in vitro* have greater variance around the median than *in silico*.

Table 2 Paired results comparing alpha diversity, or the effective number of species, between samples pooled *in vitro* and *in silico* at $q = 0$ to 5

Order of diversity (q)	Paired t -tests		Paired Brown-Forsythe test	
	t stat	p value	F-stat	p value
0	-3.4266	0.002**	4.8314	0.032*
1	-1.7288	0.095	10.82	0.001**
2	-1.2332	0.228	8.0142	0.006**
3	-1.9203	0.065	2.1882	0.145
4	-2.3941	0.024*	0.4312	0.514
5	-2.4718	0.020*	0.1216	0.728

characterized spatial differences in microbial community along the root length based on bacterial growth strategies and soil niche occupancy in nutrient-limited environments [51, 52]. We can similarly infer that bacterial community richness and abundance are spatially distinct at our three tested rhizosphere depths (Fig. 3). We see each of the three rhizosphere depths grouping apart from one another, illustrating their unique microbial composition. Further, on Axis 2, depth B is grouped away from the other two depths, indicating that some microbiota within depth B might experience more or less diurnal habitat fluctuations within the rhizosphere environment than the other two spatial depths which. For example, depth A harbors the greatest heterogeneity in community structure in comparison with other rhizosphere depths possibly due to increased contact with ambient oxygen and then influence of temperatures near the surface of rhizosphere soil [14]. Once phylogenetic relatedness was accounted for, the percent of total variation in bacterial richness and abundance explained increased 1.86%. Since weighted UniFrac calculates the proportion of unique branches in our phylogeny against the total number of branches, we were able to interpret this as each rhizosphere depth containing phylogenetically distinct microbial communities ($p < 0.001$). However, our multivariate models show a large amount of variation in rhizosphere community structure explained by the plant itself (52.7% and 82.2% for models using the quantitative Jaccard and weighted UniFrac distances respectively). Since bacterial communities were sampled from plants of the same developmental age, it is possible that this variation could be explained by genotypic variation within *S. lyrata* and heterogeneity of abiotic conditions [53].

Complex organic root exudates directly affect microbial community structure within the rhizosphere and can indirectly mediate plant growth through elaborate plant-soil feedbacks [1, 54]. Carbon released through rhizodeposition can be beneficial or antagonistic towards groups of bacteria dependent on a myriad of abiotic conditions like rhizosphere temperature, moisture, pH, and oxygen availability [14, 55, 56].

Survival in this heterogeneous environment is maintained through many different mechanisms. For example, chemotaxis or the preferential motility of bacteria away from harmful substances and towards beneficial conditions may be largely responsible for the spatial structure we see in rhizosphere microbiomes [57, 58]. Microbe-microbe competition facilitates the predominance of bacterial groups at specific depths within the rhizosphere, primarily in locations with increased carbon-based resources [15, 59]. Thus, interspecific competition might lead to lower alpha diversity like we see in depth C in comparison with other rhizosphere depths. However, given our sampling methods, we can only illustrate the spatial structure of bacterial communities as it varies with rhizosphere depth and cannot tease apart the effects of host-associated processes and environmental conditions that might shape these communities.

Our indicator species analysis revealed which ASVs are more strongly associated with each depth within the rhizosphere microhabitat. However, due to high microbial diversity found in the rhizosphere and the high throughput sequencing resolution of this study, more meaningful interpretations are given at higher taxonomic levels that encompass known ecological factors. In terms of possible indicators, depth A contains the most numerous and diverse indicator ASVs, which might be due to increased heterogeneity in abiotic and biotic conditions. For example, *Xiphinematobacteraceae*, a *Verrucomicrobia*, is strongly correlated with depth A and might be predicted as an indicator at this rhizosphere depth due to the presence of other microbiota like arbuscular mycorrhizal fungi (AMF), which have a larger impact on carbon cycling than bacteria in zones of maturing root hairs [60] present in our system at depth A, although no AMF data were observed in this study. AMF presence can have positive or negative effects for certain bacterial groups. AMF can be growth promoting for *Xiphinematobacteraceae* due to increased exuded carbon and phosphorous content [61, 62]. *Agromyces ramosus* is a common predatory bacteria of gram-negative bacteria in soil and utilizes the rhizosphere niche because of the high microbial richness and abundance of gram-negative prey [63]. This might explain the presence of *A. ramosus* as a strong indicator in depth A. *Hyphomicrobiaceae*, the most abundant indicator of depth B (Fig. 5), are chemoheterotrophs in the class *Alphaproteobacteria* and therefore heavily rely on carbon provided by other organisms [64]. *Xanthobacteraceae* (*Alphaproteobacteria*) was consistently associated within depth C as an indicator species and is known as chemolithoautotrophs that utilize inorganic compounds from external sources, such as sulfur from bedrock or soil particulate [65]. The identified indicator ASVs in this analysis and the defined spatial structure of this study illustrates how microbial communities are environmentally filtered in heterogeneous habitats like the rhizosphere [15, 66].

Sampling grain affects rarefied bacterial richness and relative abundance

High-resolution molecular techniques and a comparable sampling grain are both integral components towards understanding the spatial structure of bacterial microbiomes within rhizosphere systems. By definition, the grain size during sampling procedures can impact one's ability to detect macroecological patterns regarding species richness [67, 68]. This similarly holds true when investigating plant-microbiome interactions within the rhizosphere. After standardization, comparisons between sequencing resolutions demonstrated that pooled samples *in vitro* contain on average more raw reads. Similarly, after pairwise rarefaction, samples that were pooled *in vitro* not only have on average a greater relative abundance in reads, but they also have greater dispersion around the median in comparison to samples pooled *in silico* (Fig. 6a). This illustrates that pooling prior to DNA extraction captures larger variance in the richness of 16S genes recovered from rhizosphere microbiomes. Samples pooled *in silico* captured a larger quantity of rare, or less common across all samples, ASVs after sequencing than did coarse grain samples from the same rhizosphere (Fig. 6b). This pattern is the clearest at $q = 1$; however, it is still relevant when rare ASVs are down weighted from our analyses suggesting that pooling post sequencing can reduce the bias effects of library preparation and sequencing on richness estimates. At several steps throughout the gene sequencing procedure, selective and stochastic biases can affect the ratio of 16S genes from a complex, unknown microbiome. For instance, DNA extractions can be incomplete for soil microbiomes, where mechanical lysis and other purification steps favor particular groups of bacteria. Additionally, PCR amplification can cause biases during the annealing stage by selecting for templates that separate into single-strand molecules with minimal effectiveness [69, 70]. Pooling rhizosphere samples prior to DNA sequencing has been previously proposed as one method to reduce variation in richness estimates [71]. However, by combining multiple samples after sequencing, it is possible to illustrate the spatial dependence of rhizosphere-associated microbiota while also reducing variation around richness estimates. Thus, the sampling grain chosen during library preparation and sequencing can have an effect on our interpretations of rarefied richness of microbiomes.

In summary, we conclude that sampling method contributes to differences in the relative abundance of reads and richness of rhizosphere-associated bacteria, where higher sampling resolution captures more reliable estimates of relative abundance while also recovering higher alpha diversity estimates (i.e., more rare ASVs) on average than a bulk sampling approach. Our study explicitly shows the spatial organization of bacterial communities within the rhizosphere environment; however, given our sampling design, we cannot tease apart host effects from environmental effects or the interaction between the two. Yet, this information proves useful when developing future rhizosphere

microbiome-related questions with the aim of having reproducible and comparable experimental designs in field settings. For example, to address how root-associated microbial communities change in response to applied drought treatments, taking multiple fine grain samples from different points within the rhizosphere, pooling these *in silico*, and testing against an indicator analysis can likely clarify which ASVs distinctly correlate to drought conditions. Better understanding how bacterial communities are spatially structured along plant host tissue and which grain size is appropriate for sampling will allow us to more precisely infer which abiotic and biotic conditions shape and maintain rhizosphere microbiomes.

Acknowledgements We personally would like to thank Stephanie Kivlin, Jen Schweitzer, Susan Kalisz, and Elizabeth Derryberry for aiding in the intellectual and writing discussion centered around this study. Field sampling could not have been possible without permits approved by the Tennessee Wildlife Resource Agency. We would additionally like to thank Veronica Brown and the University of Tennessee Genomics Core, Katherine Moccia, and Andrew Willems for their help preparing and sequencing amplicon libraries.

Author contributions All the authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Jonathan Dickey. The first draft of the manuscript was written by Jonathan Dickey, and all the authors commented on the previous versions of the manuscript. All the authors read and approved the final manuscript.

Funding This work was supported by The National Science Foundation [DEB1638922 to J.A.F. and S.L.L.].

Data availability The datasets generated during and/or analyzed during the current study are available in the NCBI Sequence Read Archive (SRA) repository. <https://www.ncbi.nlm.nih.gov/sra/PRJNA575901>

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval This article does not contain any studies with human participants or animals performed by any of the authors.

Code availability RStudio code can be made available upon request to the corresponding author.

References

1. Hu L, Robert CAM, Cadot S, Zhang X, Ye M, Li B, Manzo D, Chervet N, Steinger T, van der Heijden MGA, Schlaeppi K, Erb M (2018) Root exudate metabolites drive plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota. *Nat Commun* 9:1–13. <https://doi.org/10.1038/s41467-018-05122-7>
2. Wardle DA, Bardgett RD, Klironomos JN et al (2004) Ecological linkages between aboveground and belowground biota. *Science* (80-) 304:1629–1633. <https://doi.org/10.1126/science.1094875>

3. Bever JD, Westover KM, Antonovics J (1997) Incorporating the soil community into plant population dynamics : the utility of the feedback approach. *J Ecol* 85:561–573
4. Van Der Heijden MGA, Bardgett RD, Van Straalen NM (2008) The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol Lett* 11:296–310. <https://doi.org/10.1111/j.1461-0248.2007.01139.x>
5. Lau JA, Lennon JT (2011) Evolutionary ecology of plant-microbe interactions: soil microbial structure alters selection on plant traits. *New Phytol* 192:215–224. <https://doi.org/10.1111/j.1469-8137.2011.03790.x>
6. Van Nuland ME, Wooliver RC, Pfennigwerth AA et al (2016) Plant–soil feedbacks: connecting ecosystem ecology and evolution. *Funct Ecol* 30:1032–1042. <https://doi.org/10.1111/1365-2435.12690>
7. Hobbie SE (1992) Effects of plant species nutrient cycling. *TREE* 7:336–339
8. Creer S, Deiner K, Frey S, Porazinska D, Taberlet P, Thomas WK, Potter C, Bik HM (2016) The ecologist's field guide to sequence-based identification of biodiversity. *Methods Ecol Evol* 7:1008–1018. <https://doi.org/10.1111/2041-210X.12574>
9. Žifčáková L, Větrovský T, Howe A, Baldrian P (2016) Microbial activity in forest soil reflects the changes in ecosystem properties between summer and winter. *Environ Microbiol* 18:288–301. <https://doi.org/10.1111/1462-2920.13026>
10. Berg G, Smalla K (2009) Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiol Ecol* 68:1–13. <https://doi.org/10.1111/j.1574-6941.2009.00654.x>
11. Marschner P, Yang C-H, Lieberei R, Crowley DE (2001) Soil and plant species effects on bacterial community composition in the rhizosphere. *Soil Biol Biochem* 33:1437–1445. [https://doi.org/10.1016/S0038-0717\(01\)00052-9](https://doi.org/10.1016/S0038-0717(01)00052-9)
12. Schweitzer JA, Bailey JK, Hart SC, Wimp GM, Chapman SK, Whitham TG (2005) The interaction of plant genotype and herbivory decelerate leaf litter decomposition and alter nutrient dynamics. *Oikos* 110:133–145. <https://doi.org/10.1111/j.0030-1299.2005.13650.x>
13. Hénault C, English LC, Halpin C et al (2006) Microbial community structure in soils with decomposing residues from plants with genetic modifications to lignin biosynthesis. *FEMS Microbiol Lett* 263:68–75. <https://doi.org/10.1111/j.1574-6968.2006.00416.x>
14. Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH (2013) Going back to the roots: the microbial ecology of the rhizosphere. *Nat Rev Microbiol* 11:789–799. <https://doi.org/10.1038/nrmicro3109>
15. Fierer N, Jackson RB (2006) The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci U S A* 103:626–631. <https://doi.org/10.1073/pnas.0507535103>
16. Fierer N (2017) Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat Rev Microbiol* 15:579–590. <https://doi.org/10.1038/nrmicro.2017.87>
17. Nielsen TH, Bonde TA, Sørensen J (1998) Significance of microbial urea turnover in N cycling of three Danish agricultural soils. *FEMS Microbiol Ecol* 25:147–157. [https://doi.org/10.1016/S0168-6496\(97\)00091-3](https://doi.org/10.1016/S0168-6496(97)00091-3)
18. Yao J, Allen C (2006) Chemotaxis is required for virulence and competitive fitness of the bacterial wilt pathogen *Ralstonia solanacearum*. *J Bacteriol* 188:3697–3708. <https://doi.org/10.1128/JB.188.10.3697-3708.2006>
19. Yang C, Crowley DE (2000) Rhizosphere microbial community structure in relation to root location and plant iron nutritional status. Downloaded from <http://aem.asm.org/> on January 11, 2018 by NATIONAL CENTRE FOR CELL SCIENCE (NCCS). *Appl Environ Microbiol* 66:345–351. <https://doi.org/10.1128/AEM.66.1.345-351.2000>. Updated, Rhizosphere Microbial Community Structure in Relation to Root Location and Plant Iron Nutritional Status
20. Dean SL, Farrer EC, Porras-Alfaro A, Suding KN, Sinsabaugh RL (2015) Assembly of root-associated bacteria communities: Interactions between abiotic and biotic factors. *Environ Microbiol Rep* 7:102–110. <https://doi.org/10.1111/1758-2229.12194>
21. Borcard D, Legendre P, Drapeau P (1992) Partialling out the spatial component of ecological variation. *Ecology* 73:1045–1055
22. Legendre P, Borcard D, Peres-Neto PR (2005) Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecol Monogr* 75:435–450
23. Agrawal AA, Ackerly DD, Adler F, Arnold AE, Cáceres C, Doak DF, Post E, Hudson PJ, Maron J, Mooney KA, Power M, Schemske D, Stachowicz J, Strauss S, Turner MG, Werner E (2007) Filling key gaps in population and community ecology. *Front Ecol* 5:145–152
24. Perilli S, Di Mambro R, Sabatini S (2012) Growth and development of the root apical meristem. *Curr Opin Plant Biol* 15:17–23. <https://doi.org/10.1016/j.pbi.2011.10.006>
25. Smucker AJM (1993) Soil environmental modifications of root dynamics and measurement. *Annu Rev Phytopathol* 31:191–216
26. Zhalnina K, Louie KB, Hao Z, Mansoori N, da Rocha UN, Shi S, Cho H, Karaoz U, Loqué D, Bowen BP, Firestone MK, Northen TR, Brodie EL (2018) Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nat Microbiol* 3:470–480. <https://doi.org/10.1038/s41564-018-0129-3>
27. Lugtenberg B (2015) Life of microbes in the rhizosphere
28. Chen H, Wu H, Yan B, Zhao H, Liu F, Zhang H, Sheng Q, Miao F, Liang Z (2018) Core microbiome of medicinal plant *Salvia miltiorrhiza* seed: a rich reservoir of beneficial microbes for secondary metabolism? *Int J Mol Sci* 19. <https://doi.org/10.3390/ijms19030672>
29. Huang W, Long C, Lam E (2018) Roles of plant-associated microbiota in traditional herbal medicine. *Trends Plant Sci* 23:559–562. <https://doi.org/10.1016/j.tplants.2018.05.003>
30. Micallef SA, Shiaris MP, Colón-Carmona A (2009) Influence of *Arabidopsis thaliana* accessions on rhizobacterial communities and natural variation in root exudates. *J Exp Bot* 60:1729–1742. <https://doi.org/10.1093/jxb/erp053>
31. Hafner S, Wiesenberger GLB, Stolnikova E, Merz K, Kuzyakov Y (2014) Spatial distribution and turnover of root-derived carbon in alfalfa rhizosphere depending on top- and subsoil properties and mycorrhization. *Plant Soil* 380:101–115. <https://doi.org/10.1007/s11040-014-2059-z>
32. Lauber CL, Zhou N, Gordon JJ, Knight R, Fierer N (2010) Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. *FEMS Microbiol Lett* 307:80–86. <https://doi.org/10.1111/j.1574-6968.2010.01965.x>
33. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 41:1–11. <https://doi.org/10.1093/nar/gks808>
34. Thijs S, De Beeck MO, Beckers B et al (2017) Comparative evaluation of four bacteria-specific primer pairs for 16S rRNA gene surveys. *Front Microbiol* 8:1–15. <https://doi.org/10.3389/fmicb.2017.00494>
35. Callahan BJ, McMurdie PJ, Rosen MJ et al (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 13:581–583. <https://doi.org/10.1038/nmeth.3869>
36. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:590–596. <https://doi.org/10.1093/nar/gks1219>

37. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Priesse E, Quast C, Schweer T, Peplies J, Ludwig W, Glöckner FO (2014) The SILVA and “all-species Living Tree Project (LTP)” taxonomic frameworks. *Nucleic Acids Res* 42:643–648. <https://doi.org/10.1093/nar/gkt1209>
38. 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>
39. Oksanen AJ, Blanchet FG, Friendly M, et al (2019) Package ‘vegan’. <https://cran.r-project.org>, <https://github.com/veg>
40. De Caceres M, Legendre P (2009) Associations between species and groups of sites: indices and statistical inference. *Ecology*, URL <http://sites.google.com/site/miqueldecaceres/>. *Ecology* 90:3566–3574
41. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>
42. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
43. Smith SA, O’Meara BC (2012) TreePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* 28:2689–2690. <https://doi.org/10.1093/bioinformatics/bts492>
44. Jost L (2007) Partitioning diversity into independent alpha and beta components. *Ecology* 88:2427–2439. <https://doi.org/10.1002/ecy.2446>
45. Hill MO (1973) Diversity and evenness: a unifying notation and its consequences. *Ecology* 54:427–432. <https://doi.org/10.2307/1934352>
46. Charney N, Record S (2012) vegetarian: Jost Diversity Measures for Community Data, version 1.2
47. Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67: 10.18637/jss.v067.i01
48. Anderson MJ, Ellingsen KE, McArdle BH (2006) Multivariate dispersion as a measure of beta diversity. *Ecol Lett* 9:683–693. <https://doi.org/10.1111/j.1461-0248.2006.00926.x>
49. Bobille H, Fustec J, Robins RJ, Cukier C, Limami AM (2019) Effect of water availability on changes in root amino acids and associated rhizosphere on root exudation of amino acids in *Pisum sativum* L. *Phytochemistry* 161:75–85. <https://doi.org/10.1016/j.phytochem.2019.01.015>
50. Liang T, Yang G, Ma Y, et al (2019) Seasonal dynamics of microbial diversity in the rhizosphere of *Ulmus pumila* L. var. *sabulosa* in a steppe desert area of Northern China. *PeerJ* 7:e7526. <https://doi.org/10.7717/peerj.7526>
51. De Leij FAAM, Whipps JM, Lynch JM (1994) The use of colony development for the characterization of bacterial communities in soil and on roots. *Microb Ecol* 27:81–97
52. Bever JD, Platt TG, Morton ER (2012) Microbial population and community dynamics on plant roots and their feedbacks on plant communities. *Annu Rev Microbiol* 66:265–283. <https://doi.org/10.1146/annurev-micro-092611-150107>
53. Andreote FD, Da Rocha UN, Araújo WL et al (2010) Effect of bacterial inoculation, plant genotype and developmental stage on root-associated and endophytic bacterial communities in potato (*Solanum tuberosum*). *Antonie van Leeuwenhoek, Int J Gen Mol Microbiol* 97:389–399. <https://doi.org/10.1007/s10482-010-9421-9>
54. Kozdrój J, Van Elsas JD (2000) Response of the bacterial community to root exudates in soil polluted with heavy metals assessed by molecular and cultural approaches. *Soil Biol Biochem* 32:1405–1417. [https://doi.org/10.1016/S0038-0717\(00\)00058-4](https://doi.org/10.1016/S0038-0717(00)00058-4)
55. Aulakh MS, Wassmann R, Bueno C, Rennenberg H (2001) Impact of root exudates of different cultivars and plant development stages of rice (*Oryza sativa* L.) on methane production in a paddy soil. *Plant Soil* 230:77–86. <https://doi.org/10.1023/A:1004817212321>
56. Grayston SJ, Campbell CD (1996) Functional biodiversity of microbial communities in the rhizospheres of hybrid larch (*Larix eurolepis*) and Sitka spruce (*Picea sitchensis*). *Tree Physiol* 16: 1031–1038
57. Wadhams GH, Armitage JP (2004) Making sense of it all: bacterial chemotaxis. *Nat Rev Mol Cell Biol* 5:1024–1037. <https://doi.org/10.1038/nrm1524>
58. O’Banion BS, O’Neal L, Alexandre G, Lebeis SL (2019) Bridging the gap between single-strain and community-level plant-microbe chemical interactions. *Mol Plant-Microbe Interact* 33:MPMI-04-19-0115. <https://doi.org/10.1094/mpmi-04-19-0115-cr>
59. Walker TS, Bais HP, Grotewold E, Vivanco JM (2003) Root exudation and rhizosphere biology. *Plant Physiol* 132:44–51. <https://doi.org/10.1104/pp.102.019661.Although>
60. Hamel C (2004) Impact of arbuscular mycorrhizal fungi on N and P cycling in the root zone. *Can J Soil Sci* 84:383–395. <https://doi.org/10.4141/s04-004>
61. Zhang L, Xu M, Liu Y, Zhang F, Hodge A, Feng G (2016) Carbon and phosphorus exchange may enable cooperation between an arbuscular mycorrhizal fungus and a phosphate-solubilizing bacterium. *New Phytol* 210:1022–1032. <https://doi.org/10.1111/nph.13838>
62. Chiang E, Schmidt ML, Berry MA, Biddanda BA, Burtner A, Johengen TH, Palladino D, Deneff VJ (2018) Verrucomicrobia are prevalent in north-temperate freshwater lakes and display class-level preferences between lake habitats. *PLoS One* 13:1–20. <https://doi.org/10.1371/journal.pone.0195112>
63. Korp J, Vela Gurovic MS, Nett M (2016) Antibiotics from predatory bacteria. *Beilstein J Org Chem* 12:594–607. <https://doi.org/10.3762/bjoc.12.58>
64. Rosenberg E, DeLong EF (2014) Lory S, et al. The prokaryotes, Actinobacteria
65. Oren A (2014) The Family Xanthobacteraceae BT - the Prokaryotes: Alphaproteobacteria and Betaproteobacteria. In: Rosenberg E, DeLong EF, Lory S et al (eds) Springer. Berlin Heidelberg, Berlin, Heidelberg, pp 709–726
66. Horner-Devine MC, Lage M, Hughes JB, Bohannan BJM (2004) A taxa-area relationship for bacteria. *Nature* 432:750–753. <https://doi.org/10.1038/nature03073>
67. Wiens JA (2016) Spatial scaling in ecology. *Br Ecol Soc* 3:385–397
68. Hortal J, Borges PAV, Gaspar C (2006) Evaluating the performance of species richness estimators: sensitivity to sample grain size. *J Anim Ecol* 75:274–287. <https://doi.org/10.1111/j.1365-2656.2006.01048.x>
69. Suzuki MT, Giovannoni SJ (1996) Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl Environ Microbiol* 62:625–630. <https://doi.org/10.1128/aem.62.2.625-630.1996>
70. Kennedy K, Hall MW, Lynch MDJ, Moreno-Hagelsieb G, Neufeld JD (2014) Evaluating bias of Illumina-based bacterial 16S rRNA gene profiles. *Appl Environ Microbiol* 80:5717–5722. <https://doi.org/10.1128/AEM.01451-14>
71. Cahill JF, Cale JA, Karst J et al (2016) No silver bullet: different soil handling techniques are useful for different research questions, exhibit differential type I and II error rates, and are sensitive to sampling intensity. *New Phytol* 216:11–14. <https://doi.org/10.1111/nph.14141>