Discovery of a spatially and temporally persistent core microbiome of the common bean rhizosphere

Nejc Stopnisek[1,2,3] and Ashley Shade[1,2,3,4,5]

1. Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48840 USA
2. The Plant Resilience Institute, Michigan State University, East Lansing, MI 48840
3. The DOE Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI 48840
4. Program in Ecology, Evolutionary Biology and Behavior, Michigan State University, East Lansing, MI 48840
5. Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI 48840

Corresponding author: shadeash@msu.edu

Keywords: plant-microbiome, Phaseolus vulgaris, plant health, neutral model, co-occurrence network, rhizobia, fertilizer, abundance-occupancy, 16S rRNA, ITS, microbial ecology, soil, hub taxa, connector taxa, plant genotype, sustainable agriculture

Abstract

Plants recruit soil microbes that provide nutrients, promote growth and protect against pathogens[1–3]. However, the full potential of microbial communities for supporting plant health and agriculture is unrealized [4–6], in part because rhizosphere members key for plant health are difficult to prioritize[7]. Microbes that ubiquitously associate with a plant species across large spatial scales and varied soil conditions provide a practical starting point for discovering beneficial members[7]. Here, we quantified the structures of bacterial/archaeal and fungal communities in the common bean rhizosphere (*Phaseolus vulgaris*), and assessed its core membership across space and time. To assess a spatial core, two divergent bean genotypes were grown in field conditions across five major growing regions in the United States, and then also compared to eight genotypes grown in Colombian soil. To assess a temporal core, we conducted a time course of rhizosphere and rhizoplane microbiome members over bean development in the field. Surprisingly, there were 48 persistent bacterial taxa that were detected in all samples, inclusive of U.S. and Colombian-grown beans and over plant development, suggesting cosmopolitan enrichment and time-independence. Neutral models of abundance-occupancy relationships and co-occurrence networks show that many of these core taxa are deterministically selected and likely in intimate relationships with the plant. Many of the core taxa were yet-uncultured and affiliated with Proteobacteria; these taxa are prime targets in support of translational plant-microbiome management. More generally, this work reveals that core members of the plant microbiome can have both broad ranges and temporal persistence with their host, suggesting intimate, albeit possibly opportunistic, interactions.

41 **Main**

42 Common bean (*Phaseolus vulgaris* L.) is the most important food legume grown
43 worldwide, and especially for developing economies in South America, Africa and Asia[8]. The
44 origin of common bean is now central Mexico, and from there it spread to central and south
45 America around 165,000 years ago[9]. This resulted in the development of two major and eco-
46 geographically distinct common bean gene pools with partial reproductive isolation[10–12]. The
47 Mesoamerican gene pool was distributed from northern Mexico to Colombia and the Andean gene
48 pool ranged from southern Peru to northwestern Argentina. Since 8,000 years ago, each pool was
49 separately and selectively bred, leading to further diversification between them[13–15]. Because of pre-
50 existing genetic differences in each gene pool followed by divergent breeding history, common
51 bean presently offers a distinctive opportunity for understanding how the host and the environment
52 contribute to rhizosphere microbiome assembly.

53 The objective of this study was to quantify the influences of space, time and genotype on
54 detecting a core common bean root microbiome in the field, inclusive of bacteria, archaea and
55 fungi. We aimed to identify core microbial lineages that have cosmopolitan associations with
56 common bean and therefore potential as targets towards microbe-improved plant breeding. With
57 the cooperation of U.S. common bean breeding programs, we executed a first-of-its kind study of
58 two divergent bean genotypes grown under field conditions in five major North American bean
59 growing regions, including Michigan, Nebraska, Colorado and Washington. These two genotypes
60 belong to the Mesoamerican (Eclipse genotype) and Andean (California Early Light Red Kidney,
61 CELRK, genotype) gene pools that represent the major divergences from the wild bean ancestor[9].
62 For these two divergent genotypes, we also assessed how core members of the root microbiome
63 changed over plant development. Finally, we performed a comparative analysis of the bacterial
64 microbiome found in these study and with bean genotypes grown in South America[16,17]. From our

effort that was inclusive of both broad biogeography and plant development, we discovered a core bean rhizosphere microbiome of 48 members that persistently associated with this nutritionally, agronomically, and economically important crop. This core was discovered in spite of apparent rhizosphere microbiome differences that were attributable to local soil conditions and management. We did not detect an influence of plant genotype, suggesting that this core membership supersedes it.

71 **Results and Discussion**

72 *Study design, sequencing summary, and diversity*

73 To assess the influence of environment and plant genotype on the common bean
74 rhizosphere microbiome, we designed a biogeography study of two divergent bean genotypes (from
75 Mesoamerican and Andean gene pools[18,19]), both grown in the field at five research and extension
76 farms that represent major U.S. bean production regions (**Table 1**)[20]. From each growing location,
77 triplicate bean plants grown in three to four plots were harvested at flowering (mid to late July).

78 We sequenced 39 rhizosphere (attached to the root, see methods) and root-associated soil
79 samples using 16S rRNA gene and ITS1 region, respectively to assess microbiome community
80 structures. We assessed the bacterial and archaeal communities using 16S rRNA gene amplicon
81 sequencing, and the fungal communities using ITS1 amplicon sequencing. After clustering reads to
82 OTUs at 97% similarity threshold and filtering chloroplast and mitochondria, there was 31,255 to
83 506,166 16S rRNA reads per sample and 22,716 to 252,810 ITS reads. We rarefied samples to
84 31,255 reads for 16S rRNA gene amplicons and to 22,716 for ITS. With these thresholds, we
85 achieved richness asymptotes for both datasets, suggesting that sequencing efforts were sufficient
86 to capture comparative dynamics and diversity (**Supplemental Fig. S1**). The total richness
87 observed at this rarefaction depth was 1,505 fungal and 23,872 bacterial and archaeal OTUs.

88 As reported in other rhizosphere studies, the total fungal diversity was lower than bacterial
89 /archaeal diversity in the rhizosphere of the common bean[21–23]. Richness varied by growing location
90 (ANOVA, F value=12.4, p-value<0.0001 and F value=13.1, p-value<0.0001 for 16S rRNA and ITS
91 data, respectively, **Supplemental Fig. S2**) but was highest at the Montcalm Research Facility
92 (Michigan, US) for both, bacteria/archaea and fungi. An analysis of community beta diversity
93 revealed strong biogeographic patterns in community structure explained by location, soil pH and
94 fertilization, in agreement with other literature[24–27] (**Supplemental Fig. S3**, see **Supplemental**

95 **Text**). However, an influence of plant genotype was either week or not detected, which we
96 partially attribute to plant growth in uncontrolled field conditions, which, in other plant species,
97 also show low influence of genotype on rhizosphere structure[24,25,28] (**Supplemental Fig. S2BD**, see
98 **Supplemental Text).**
99
100 *A core rhizosphere microbiome is detected across U.S. bean growing regions*
101 The common bean rhizosphere microbiome included the major expected lineages for both
102 bacteria and fungi (**Supplemental Fig. S4AB**), in agreement with other plant rhizosphere studies[29–
103 34]. Together with previous studies, these data provide more evidence that root-associated microbial
104 taxa are phylogenetically and functionally conserved[35]. Proteobacteria, Acidobacteria,
105 Bacteroidetes, and Actinobacteria collectively comprised on average 73.5% of the bacteria/archaeal
106 community, Ascomycota dominated the fungal community with a mean total relative abundance of
107 53% with notable sample-to-sample variance (range from 16.5% to 84.5%).
108 Despite differences in the rhizosphere microbiomes across growing locations, we noticed a
109 large number of OTUs that were shared among them for the bacterial/archaeal communities (2,173
110 taxa, mean 31.5%, range 29.5% to 34.7%). There was a smaller but notable overlap for the fungal
111 communities (70 taxa or mean 4.5%, range from 0.9% to 17.9%; **Supplemental Fig. S4CD**). These
112 data suggested that, despite measured edaphic differences across growing locations and strong
113 biogeographic signal, the common bean rhizosphere recruited many similar taxa that could be
114 functionally important for the bean. Therefore, we explored occupancy-abundance distributions of
115 taxa ([36,37] and references therein) to infer the core bean microbiome of taxa with an occupancy of 1
116 (i.e., found in all samples and across all growing locations; **Fig. 1AB**). Among bacteria and
117 archaea, 258 phylogenetically diverse taxa were cosmopolitan in the dataset (**Fig. 1C**), including
118 numerous and abundant Proteobacteria (117 OTUs) with a dominant taxon classified as

119 *Arthrobacter* sp. (FM209319.1.1474, mean relative abundance of 1.43%). The bacterial/archaeal
120 core also contained taxa of interest for potential plant benefits (e.g. *Sphingomonas*, *Rhizobium*,
121 *Bacillus, Streptomyces*), as well as some genera that can be associated with disease (e.g. *Ralstonia*).
122 There were 13 taxa in the fungal core (**Fig. 1D**), and these were largely composed of Ascomycota
123 (10 OTUs), with dominating taxon OTU823 from the *Phaeosphaeriaceae* family (mean relative
124 abundance 10.1%). Notably, taxa that were unique to either bean genotype were relatively rare and
125 inconsistently detected (**Fig. 1**, orange and black points). Together, these results suggest that
126 common bean, regardless of the genotype, consistently recruits particular microbiome taxa.
127 The neutral expectation of abundance-occupancy distributions is that very abundant taxa
128 will have high occupancy, while rare taxa will have low[37–40]. Taxa that deviate from the neutral
129 expectation are more strongly influenced by deterministic factors, like environmental conditions,
130 than by stochastic factors, like drift and dispersal. The neutral model fit[38,39] of the abundance-
131 occupancy distribution (solid line, **Fig. 1AB**) identified several taxa that had frequencies either
132 above or below the 95% confidence intervals of the model (dashed lines). Specifically, 13.7% of
133 the bacterial/archaeal and 30.4% of fungal taxa, deviated from the neutral expectation
134 (**Supplemental Dataset S3**). One hundred and seventy-one core taxa were predicted above the
135 neutral model partition; these deterministically-selected taxa are prime candidates for follow-up
136 studies of interactions with the host plant. Overall, the bacteria/archaea community had better fit to
137 the neutral expectation than fungal (R^2 of 0.74 and 0.34, and migration rates (m) of 0.301 and
138 0.003, respectively), suggesting that dispersal was relatively less limiting for the bacteria/archaea
139 than for the fungi. This finding agrees with other work suggesting that fungi are more sensitive to
140 local climate or more dispersal limited than bacteria[41–44].
141
142 *A cross-continental core rhizosphere microbiome is detected for common bean*

143 We wanted to better understand if these U.S. core taxa were associated with the bean
144 rhizosphere across a larger geographical scale, which would suggest the potential for selective plant
145 recruitment and cosmopolitan distribution of core taxa. Therefore, we compared our U.S. data to a
146 recently published study of rhizosphere bacteria and archaea from common beans grown in
147 Colombian agricultural soil[16]. The Colombian study offered a key contrast because it included eight
148 divergent bean lineages, including wild (n=2), landrace (n=1), and cultivated genotypes (n=5),
149 grown in soil from a different continent that has starkly different climate and management from the
150 U.S. growing regions. To enable direct comparison, we re-analyzed raw reads and compared the
151 datasets by matching to either the same taxon identifiers when clustered to SILVA database, or
152 100% identity by BLAST to *de novo* clustered reads (see Materials and Methods). Surprisingly,
153 39.6% (3,359 OTUs) of rhizosphere taxa from the Colombian-grown beans were also shared with
154 the U.S. dataset (**Fig. 2**). Both datasets included taxa that were highly represented in the other: 62%
155 of U.S. core (159 out of 258) were found also in Colombia, and 51% of Colombian core (433 out of
156 848) were shared with the U.S. (**Fig. 2A**). Core taxa were again defined stringently with an
157 occupancy of 1, and 48 taxa were found across all samples, inclusive of both datasets. These core
158 taxa were composed of many Proteobacteria, with *Rhizobiales* showing the most consistent relative
159 abundance between the studies (**Fig. 2B,** e.g. 0.187% and 0.138% in Colombia and U.S. dataset,
160 respectively). Notably, none of this cross-continental core taxa were universally detected in very
161 high abundance, and all but two OTUs (a U.S. *Arthrobacter* sp. and a Colombian *Austrofundulus*
162 *limnaeus* with mean relative abundances of 1.43% and 1.01%, respectively) would be classified as
163 rare by a typical mean relative abundance threshold below 1%, hinting to a potential role of rare
164 taxa in providing key functions. Notably, only 48% of these cross-continental shared core taxa have
165 genus classification, suggesting that most of them are under-described in their functional potential
166 and interactions with plants.

167 We also analyzed reads associated with the global core taxa with the UNOISE3 pipeline[45] to
168 generate predicted biological sequences (zero-radius OTUs – ZOTUs with100% sequence identity)
169 and provide the maximal possible biological resolution. We found that the global core taxa
170 consisted of 422 ZOTUs, and that there was a range of 2 to 35 ZOTUs identified within each OTU
171 (**Supplemental Fig. S5**). With one exception (HQ597858.1.1508), all of core OTUs (clustered at
172 97% sequence identity) contained at least one ZOTUs that also had an occupancy of 1. In addition,
173 all of the ZOTUs with an occupancy of 1 were also the most abundant ZOTU within each OTU
174 (**Supplemental Fig. S5**). This result suggests that the same members constitute the core even with
175 increased taxonomic resolution.

176 A recent study considered the effect of the common bean domestication history on the root
177 microbiome and identified a core set of microbial taxa that are consistently present with these
178 diverse bean genotypes, including bacterial taxa recruited from agricultural and natural soil from
179 Colombia[17]. This core also had high representation of *Proteobacteria*, *Acidobacteria*,
180 *Actinobacteria* and *Verrucomicrobia*, similar to those observed in the present study. We also re-
181 analyzed these data and show that 46 (out of 48) of the cross-continental core taxa identified here
182 have occupancy>0.9 across all three included studies (the present study, Pérez-Jaramillo et al.[16] and
183 Pérez-Jaramillo et al.[17]). Forty-two of these taxa had the highest possible occupancy of 1, but only
184 within rhizosphere samples from agricultural soils (**Supplemental Dataset S4**). When
185 incorporating data from beans grown in forest soils, the average occupancy decreases to 0.77
186 (median 0.75, min = 0.53). However, 14 taxa still had an occupancy of 1 when including beans
187 grown in forest soils (**Supplemental Dataset S4**).

188 In summary, these results show that common bean can associate with a core set of
189 rhizosphere microbiome members at taxon and ecotype levels, across diverse bean genotypes and

190 across continents. Additionally, these core taxa are likely enriched by the host in field agriculture,
191 as suggested by their higher occupancy in agroecosystems.
192
193 *Core taxa are enriched in the rhizoplane and are consistently detected across bean development*
194 To identify the common bean core taxa over space (a biogeographic core) we sampled
195 plants across growing locations at the same growth stage (flowering) and focused on the
196 rhizosphere compartment. However, the question remained whether these core taxa are detected
197 beyond that particular plant development stage. To answer this question, we conducted a field
198 experiment to assess the core taxa over time in plant development stage. In the next growing season
199 (2018), we used the same divergent bean genotypes grown at both Michigan, U.S, locations
200 (Montcalm and Saginaw Valley; see Material and Methods). We harvested root systems at 5
201 different plant development stages including flowering stage. We investigated the relative
202 abundance of the cross-continental core taxa and the U.S.-specific core taxa in both rhizosphere
203 (soil that could be removed from the root after shaking) and rhizoplane (soil adhered to the root
204 tissue and removed via vortex in buffer) compartments to determine their ability to closely
205 associate with the plant tissue. All 48 cross-continental core taxa were detected a year later in these
206 field, and their relative abundances were significantly higher in the rhizoplane as compared to the
207 rhizosphere irrespective of the plant development stage, bean genotype and growing location (**Fig.**
208 **3A**). Interestingly, the remaining core taxa that were found exclusively in the U.S. dataset at an
209 occupancy of 1 were equally abundant in the rhizosphere and rhizoplane at both Michigan growing
210 locations (**Fig. 3A**).
211 We next asked whether there were enrichments of particular core taxa by plant development
212 stage, root compartment, or growing location (**Fig. 3B**). On balance, almost all core taxa showed
213 some growth stage preference, but these trends were specific to each growing location and root

214 compartment. Despite these nuances, all core taxa were consistently found with high occupancy
215 inclusive of the plant development series (**Supplemental Fig. S6**).
216 Together these results suggest that these core taxa are selected by the plant early in the
217 development stage and maintained. Enrichment of the core taxa in the rhizoplane further supports
218 the hypothesis that these core taxa engage closely with the host plant.
219
220 *Core taxa are not hub or connector taxa in an inter-kingdom microbiome network*
221 Network analysis has been proposed to be a useful method to identify important members
222 of the plant microbiome with beneficial traits[46–48]. Hub taxa, identified by their high connectivity
223 with many members of the community, are regarded as the most important part of the community
224 and influence network structure and community stability[5,49]. We applied this method to ask if any of
225 the core taxa that we identified using abundance-occupancy were also key for co-occurrence
226 network structure. Additionally, we were interested in identifying fungal-bacterial co-occurrences
227 because of reports of their potential benefits for the plant[50,51]. To explore these patterns we applied
228 the molecular ecology network analysis pipeline[52] using samples collected across the U.S. where
229 we have both 16S rRNA and ITS amplicon data. We merged rarified 16S and ITS datasets, filtered
230 the datasets to include taxa with occupancy greater than or equal to 50%, and considered only
231 interactions significant at p-value<0.05 and LA score greater or equal to 0.88. The resulting
232 network included 572 taxa (nodes) and 1,857 statistically significant interactions (edges) structured
233 among 52 modules. Most of the modules were relatively small, with only six including more than
234 10 nodes (**Fig. 4**). The network was scale-free and had small-world characteristics as indicated by
235 the node degree distribution fitting to the power law model (R^2= 0.993), and also had significant
236 deviation of the modularity, length and clustering coefficients from those calculated from random
237 network (same number of nodes and edges), respectively (**Supplemental Dataset S5**).

238 The topological role of each taxon within the network was determined by the relationship
239 between their within-module (Zi) and among-module (Pi) connectivity scores (**Fig. 4A**)[53]. Based on
240 this, the majority of taxa were peripheral (potentially, specialists; 563 nodes). There were also 3
241 connectors and 6 module hubs, but no network hubs. Indeed, in agreement with the beta-diversity
242 analysis (**Fig. S3**, Supporting Materials), there was a strong geographic signal in the largest four
243 modules, and these were comprised mostly of bacterial-bacterial (rather than bacterial-fungal or
244 fungal-fungal) hypothesized interactions (**Fig. 4B**). Interestingly, cross-kingdom interactions
245 constituted only a small fraction of all co-occurrences (n=168; bacteria-fungi=156, archaea-
246 bacteria=6 and archaea-fungi=6).
247 The analysis identified 26 co-occurrences between core taxa. However, we were surprised
248 to find that only two bacterial core and no fungal core taxa were also classified as network hubs
249 (taxa that connect to many other taxa within a modules) or connectors (taxa that connect across
250 modules). As exceptions, core *Chitinophagaceae* taxon FR749720.1.1492 was a module hub node
251 and a *Nitrobacter* sp. GDHX01215817.4.1477 was a connector. Our results, inclusive of a dataset
252 of divergent plant genotypes and broad biogeography, suggest that while hub and connector taxa
253 may be important for the maintenance of the root microbiome, these taxa are not consistently
254 detected in the common bean rhizosphere and, by deduction, could not be of universal importance
255 for the host plant. Our study cannot speak to the potential for functional redundancy among hub or
256 connector taxa, which could ultimately suggest a functional core among phylogenetically diverse
257 taxa[54]. Taken together, these results suggest that core taxa likely are important for the plant, while
258 hub and connector taxa are important for the integrity of the soil microbial community and its
259 responses to the local environment.
260
261 *Prospectus*

262 Legumes like the common bean that form symbiotic relationships with nitrogen fixing bacteria are
263 a crucial component of sustainable agriculture. Growing common bean can both improve soil
264 fertility[55,56] and contribute to food security, as bean production is >31 million metric tons
265 worldwide (http://www.fao.org/faostat/, accessing on July 12, 2019)). However, common bean
266 yield estimates, production areas, and even nutritional value, are predicted to decrease in this
267 century due to the impacts of global change[57–59]. Thus, breeding or other improvements to common
268 bean is needed to make this important crop resistant to stressors associated with climate change
269 while increasing yields to meet higher demands. To date, only traditional plant breeding
270 approaches have been used for the common bean, exploiting the genetic diversity of the plant, but
271 the latest knowledge of the plant microbiome in supporting host well-being has opened new
272 avenues for breeding the next generation of crops with plant microbiome manipulation[4].
273 Here, we show that common bean genotype has marginal effect on the rhizosphere
274 microbiome grown in field conditions. We describe for the first time the common bean core
275 microbiome at multiple spatial and temporal scales. A majority of the 48 core bacterial taxa
276 cosmopolitan among these common bean rhizospheres are under-described or largely unknown, as
277 only 23 out of 48 have genus classification. But, due to their ubiquitous association with the bean
278 over space and time, we hypothesize that these taxa provide functions that are crucial for common
279 bean health. A next step is to understand functions associated with these taxa and to determine if
280 and how they contribute to plant health and productivity under different growth conditions, such as
281 drought or with particular management strategies[28,60]. Additionally, targeting persistent microbiota
282 that are part of a native assemblage for management and manipulation can be advantageous, as
283 these are expected to be both host and environment adapted and thus may be more persistent *in situ*
284 than external or non-native biocontrols[3]. In conclusion, this work provides insights for prioritizing

285 plant-microbiome members that will advance goals in microbiome management and microbe-
286 improved crops.
287

288 **Material and Methods**

289 An expanded description of materials and methods is included in the Supporting Text.

290 *Sampling and soil physicochemical analysis*

291 We selected two common bean cultivars with very distinct genotypes: Eclipse[18] with
292 Mesoamerican origin and California Early Light Red Kidney (CELRK) and old kidney bean
293 landrace with Andean origin[19]. The plants were grown at 5 different locations on research extension
294 farms: Saginaw Valley (SVERC), Montcalm county (MRF) (both in Michigan), Scott Bluff county
295 (Nebraska) (NE), Fort Collins (Colorado) (CO) and Othello (Washington) (WA), each representing
296 an important U.S. common bean growing region[20]. Both bean genotypes were grown in the same
297 locations except in Michigan, where CELRK was grown at Saginaw Valley and Eclipse at
298 Montcalm county. Plants ($n \geq 3$) were harvested at flowering from replicated plots ($n \geq 3$) at each
299 location in summer 2017. The flowering time point was selected because we wanted to analyze
300 mature microbial communities as influenced by the plant and not by plant development. Plant roots
301 with attached soil were packed into bags, stored on ice and shipped immediately to Michigan State
302 University for processing.

303 From each sample, root attached soil was collected and used as rhizosphere soil and
304 remaining soil was collected and used as bulk soil. Both soils were sieved through 4 mm sieves to
305 remove any organic, plant debris and larger soil minerals before soil analysis and DNA isolation
306 and soil physical-chemical analysis (see Supporting Text). Soil analysis was done on bulk soil
307 samples pooled by plant genotype and plot.

308 We conducted a second study to assess the dynamics of the core taxa over plant
309 development. The bean genotypes, CELRK and Eclipse were grown at the same sites in Michigan,
310 U.S. as previously (Montcalm county (MRF) and Saginaw Valley (SVERC)) in the summer of
311 2018. Plants root systems were harvested at 5 growth stages, including stage 1: V2 (appearance of
312 second trifoliate), stage 2: V5 (appearance of fifth trifoliate), stage 3: flowering, stage 4: pod
313 filling, and stage 5: senescence/drying). At each sampling time, roots were collected and
314 transported on ice to the laboratory for immediate processing. Rhizosphere soil was collected by
315 gently shaking roots. Remaining attached soil (rhizoplane soil) was collected by first vortexing
316 roots in 1x phosphate buffer solution (PBS) for 4 min and finally removing the supernatant after
317 centrifugation for 10 min at 8000 *g* and 4°C. Collected soil was immediately frozen in liquid
318 nitrogen and stored at -80°C.

319 16S rRNA gene amplicon and ITS amplicon sequencing was performed at the Michigan
320 State Genomics Core Research Support Facility and analyzed using UPARSE (version 11)[61]. We
321 processed the 16S rRNA gene amplicons using an open-reference clustering first against the
322 SILVA v128 database[62], [63]. ITS OTU representative sequences were taxonomically annotated
323 using the CONSTAX tool[64] with the UNITE database version 7.2[65]. OTUs with unassigned
324 taxonomy at the domain level and OTUs annotated as mitochondria or chloroplasts were removed.
325 Statistical analysis and visualization were performed in R (see Supporting Text). The co-
326 occurrence network and global network properties were calculated using the Molecular Ecological
327 Network Analysis Pipeline (MENAP) [52] and visualized with Cytoscape v.3.5.1 [66]. Comparative
328 analyses with published datasets was performed by downloading and processing study raw reads
329 from NCBI to analyze them identically. Despite differences in our processing pipeline and the
330 dataset's originally published pipeline, we generated very similar numbers of OTUs (12,209
331 compared to 12,293[16]).

Acknowledgments

This work was supported by the Plant Resilience Institute at Michigan State University. We are grateful to Dr. James Kelly, Dr. Phillip Miklas, Dr. James J. Heitholt, Dr. Carlos Urrea Florez, Dr. Mark A. Brick, Dr. Juan M. Osorno, Dr. Thomas H Smith and other supporting staff at the research extension farms for their partnership in this study. We would also like to thank the Michigan State Genomics Core Research and Technology Support Facility and the Institute for Cyber-Enabled Research High Performance Computing Center for excellent support and service.

Footnotes

Author contributions: N.S. and A.S. designed research, analyzed data and wrote the paper; and N.S. performed research.

The authors declare no conflict of interest.

Data deposition: The raw sequence data have been deposited in the NCBI Sequence Read Archive (BioProject accession no. PRJNA524532). All read processing steps, bioinformatic workflows, and custom scripts are available on GitHub (https://github.com/ShadeLab/PAPER_Stopnisek_2019_BeanBiogeography).

349 **References**

350 1. Berendsen, R. L. *et al*. Disease-induced assemblage of a plant-beneficial bacterial
351 consortium. *ISME J*. 1 (2018). doi:10.1038/s41396-018-0093-1
352 2. Castrillo, G. *et al*. Root microbiota drive direct integration of phosphate stress and
353 immunity. *Nature* **543**, 513–518 (2017).
354 3. Bell, T. H. *et al*. Manipulating Wild and Tamed Phytobiomes: Challenges and Opportunities.
355 *Phytobiomes J*. **3**, 3–21 (2019).
356 4. Wallenstein, M. D. Managing and manipulating the rhizosphere microbiome for plant health:
357 A systems approach. *Rhizosphere* (2017). doi:10.1016/j.rhisph.2017.04.004
358 5. Toju, H. *et al*. Core microbiomes for sustainable agroecosystems. *Nat. Plants* **4**, 247–257
359 (2018).
360 6. Hawkes, C. V. & Connor, E. W. Translating Phytobiomes from Theory to Practice:
361 Ecological and Evolutionary Considerations. *Phytobiomes* **1**, 57–69 (2017).
362 7. Shade, A. & Stopnisek, N. Abundance-occupancy distributions to prioritize plant core
363 microbiome membership. *Current Opinion in Microbiology* (2019).
364 doi:10.1016/j.mib.2019.09.008
365 8. Broughton, W. J. *et al*. Beans (Phaseolus spp.) – model food legumes. *Plant Soil* **252**, 55–
366 128 (2003).
367 9. Schmutz, J. *et al*. A reference genome for common bean and genome-wide analysis of dual
368 domestications. *Nat. Genet*. (2014). doi:10.1038/ng.3008
369 10. Rendón-Anaya, M. *et al*. Genomic history of the origin and domestication of common bean
370 unveils its closest sister species. *Genome Biol*. **18**, 60 (2017).
371 11. Bitocchi, E. *et al*. Beans (Phaseolus ssp.) as a Model for Understanding Crop Evolution.
372 *Front. Plant Sci*. **8**, 722 (2017).

373 12. Bitocchi, E. *et al*. Mesoamerican origin of the common bean (Phaseolus vulgaris L.) is
374 revealed by sequence data. *Proc. Natl. Acad. Sci.* **109**, 788–796 (2012).
375 13. Kwak, M. & Gepts, P. Structure of genetic diversity in the two major gene pools of common
376 bean (Phaseolus vulgaris L., Fabaceae). *Theor. Appl. Genet.* **118**, 979–992 (2009).
377 14. Bitocchi, E. *et al*. Molecular analysis of the parallel domestication of the common bean
378 (Phaseolus vulgaris) in Mesoamerica and the Andes. *New Phytol.* **197**, 300–313 (2013).
379 15. Schmutz, J. *et al*. A reference genome for common bean and genome-wide analysis of dual
380 domestications. *Nat. Publ. Gr.* **46**, (2014).
381 16. Perez-Jaramillo, J. E. *et al*. Linking rhizosphere microbiome composition of wild and
382 domesticated Phaseolus vulgaris to genotypic and root phenotypic traits. *ISME J.* 1–14
383 (2017). doi:10.1038/ismej.2017.85
384 17. Pérez-Jaramillo, J. E. *et al*. Deciphering rhizosphere microbiome assembly of wild and
385 modern common bean (Phaseolus vulgaris) in native and agricultural soils from Colombia.
386 *Microbiome* **7**, 114 (2019).
387 18. NDSU Agriculture. *Eclipse Black Bean*. (2004).
388 19. McClean, P. & Myers, J. Pedigrees of dry bean cultivars, lines and PIs. *Annu. Rep. Bean*
389 *Improv. Coop* **33**, 23–25 (1990).
390 20. USDA. *United States Department of Agriculture National Agricultural Statistics Service*
391 *Crop Production 2016 Summary*. (2017).
392 21. Singer, E., Bonnette, J., Kenaley, S. C., Woyke, T. & Juenger, T. E. Plant compartment and
393 genetic variation drive microbiome composition in switchgrass roots. *Environ. Microbiol.*
394 *Rep.* **11**, 185–195 (2019).
395 22. Coleman-Derr, D. *et al*. Plant compartment and biogeography affect microbiome
396 composition in cultivated and native *Agave* species. *New Phytol.* **209**, 798–811 (2016).

397 23. De Souza, R. S. C. *et al*. Unlocking the bacterial and fungal communities assemblages of
398 sugarcane microbiome. *Sci. Rep*. **6**, 1–15 (2016).
399 24. Walters, W. A. *et al*. Large-scale replicated field study of maize rhizosphere identifies
400 heritable microbes. *Proc. Natl. Acad. Sci*. **115**, 7368–7373 (2018).
401 25. Wagner, M. R. *et al*. Host genotype and age shape the leaf and root microbiomes of a wild
402 perennial plant. *Nat. Commun*. **7**, 12151 (2016).
403 26. Xu, J. *et al*. The structure and function of the global citrus rhizosphere microbiome. *Nat*.
404 *Commun*. **9**, 4894 (2018).
405 27. Edwards, J. *et al*. Structure, variation, and assembly of the root-associated microbiomes of
406 rice. *Proc. Natl. Acad. Sci*. (2015). doi:10.1073/pnas.1414592112
407 28. Edwards, J. A. *et al*. Compositional shifts in root-associated bacterial and archaeal
408 microbiota track the plant life cycle in field-grown rice. *PLOS Biol*. **16**, e2003862 (2018).
409 29. Guttman, D. S. *et al*. Assembly and ecological function of the root microbiome across
410 angiosperm plant species. *Proc. Natl. Acad. Sci*. (2018). doi:10.1073/pnas.1717617115
411 30. Hamonts, K. *et al*. Field study reveals core plant microbiota and relative importance of their
412 drivers. *Environ. Microbiol*. **20**, 124–140 (2018).
413 31. Gomes, E. A., De Sousa, S. M. & Oliveira, C. A. Root-Associated Microbiome of Maize
414 Genotypes with Contrasting Phosphorus Use Efficiency. **2**, 129 (2018).
415 32. Estendorfer, J. *et al*. The Influence of Land Use Intensity on the Plant-Associated
416 Microbiome of Dactylis glomerata L. *Front. Plant Sci*. **8**, 930 (2017).
417 33. Leff, J. W., Lynch, R. C., Kane, N. C. & Fierer, N. Plant domestication and the assembly of
418 bacterial and fungal communities associated with strains of the common sunflower,
419 Helianthus annuus. *New Phytol*. **214**, 412–423 (2017).
420 34. Lundberg, D. S. *et al*. Defining the core Arabidopsis thaliana root microbiome. *Nature* **488**,

(2012).

35. Levy, A. *et al*. Genomic features of bacterial adaptation to plants. *Nat. Genet*. (2018). doi:10.1038/s41588-017-0012-9

36. Grady, K. L., Sorensen, J. W., Stopnisek, N., Guittar, J. & Shade, A. Assembly and seasonality of core phyllosphere microbiota on perennial biofuel crops. *Nat. Commun*. **19**, (2019).

37. Shade, A. *et al*. Macroecology to Unite All Life, Large and Small. *Trends Ecol. Evol*. **33**, 731–744 (2018).

38. Woodcock, S. *et al*. Neutral assembly of bacterial communities. *FEMS Microbiol. Ecol*. **62**, 171–180 (2007).

39. Burns, A. R. *et al*. Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME J*. **10**, 655–664 (2016).

40. Gaston, K. J. *et al*. Abundance-occupancy relationships. *J. Appl. Ecol*. **37**, 39–59 (2000).

41. Martiny, J. B. H., Eisen, J. A., Penn, K., Allison, S. D. & Horner-Devine, M. C. Drivers of bacterial beta-diversity depend on spatial scale. *Proc. Natl. Acad. Sci. U. S. A*. **108**, 7850–4 (2011).

42. Gumiere, T., Durrer, A., Bohannan, B. J. M. & Andreote, F. D. Biogeographical patterns in fungal communities from soils cultivated with sugarcane. *J. Biogeogr*. **43**, 2016–2026 (2016).

43. Talbot, J. M. *et al*. Endemism and functional convergence across the North American soil mycobiome. *Proc. Natl. Acad. Sci*. **111**, 6341–6346 (2014).

44. Tedersoo, L. *et al*. Global diversity and geography of soil fungi. *Science (80-.)*. **346**, (2014).

45. Edgar, R. C. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv* 81257 (2016). doi:10.1101/081257

445 46. Agler, M. T. *et al*. Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome
446 Variation. *PLOS Biol*. **14**, e1002352 (2016).
447 47. Trivedi, P. *et al*. Keystone microbial taxa regulate the invasion of a fungal pathogen in agro-
448 ecosystems. *Soil Biol. Biochem*. **111**, 10–14 (2017).
449 48. Shi, S. *et al*. The interconnected rhizosphere: High network complexity dominates
450 rhizosphere assemblages. *Ecology Letters* **19**, (2016).
451 49. Carlström, C. I. *et al*. Synthetic microbiota reveal priority effects and keystone strains in the
452 Arabidopsis phyllosphere. *Nat. Ecol. Evol*. **3**, 1445–1454 (2019).
453 50. Durá, P. *et al*. Microbial Interkingdom Interactions in Roots Promote Arabidopsis Survival.
454 *Cell* **175**, 973–983 (2018).
455 51. Wagg, C., Schlaeppi, K., Banerjee, S., Kuramae, E. E. & van der Heijden, M. G. A. Fungal-
456 bacterial diversity and microbiome complexity predict ecosystem functioning. *Nat*.
457 *Commun*. **10**, 4841 (2019).
458 52. Deng, Y. *et al*. Molecular ecological network analyses. *BMC Bioinformatics* **13**, (2012).
459 53. Guimerà, R. & Nunes Amaral, L. A. Functional cartography of complex metabolic networks.
460 *Nature* **433**, 895–900 (2005).
461 54. Lemanceau, P., Blouin, M., Muller, D. & Moënne-Loccoz, Y. Let the Core Microbiota Be
462 Functional. *Trends in Plant Science* (2017). doi:10.1016/j.tplants.2017.04.008
463 55. Aguilar, O. M., Riva, O. & Peltzer, E. Analysis of Rhizobium etli and of its symbiosis with
464 wild Phaseolus vulgaris supports coevolution in centers of host diversification. *PNAS* **101**,
465 13548–13553 (2004).
466 56. Martínez-Romero, E. Diversity of Rhizobium-Phaseolus vulgaris symbiosis: overview and
467 perspectives. *Plant Soil* **252**, 11–23 (2003).
468 57. Jaggard, K. W., Qi, A. & Ober, E. S. Possible changes to arable crop yields by 2050. *Philos*.

469 *Trans. R. Soc. Lond. B. Biol. Sci.* **365**, 2835–51 (2010).

470 58. Beebe, S., Rao, I., Mukankusi, C. & Buruchara, R. Improving Resource Use Efficiency and
471 Reducing Risk of Common Bean Production in Africa, Latin America, and the Caribbean. in
472 *CIAT Book Chapter* (ed. Hershey, C. H.) 1–18 (Centro Internacional de Agricultura Tropical
473 (CIAT), 2012).

474 59. Hummel, M. *et al*. Reduction in nutritional quality and growing area suitability of common
475 bean under climate change induced drought stress in Africa. *Sci. Rep.* **8**, 16187 (2018).

476 60. Xu, L. *et al*. Drought delays development of the sorghum root microbiome and enriches for
477 monoderm bacteria. *Proc. Natl. Acad. Sci.* **115**, E4284–E4293 (2018).

478 61. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat.*
479 *Methods* **10**, 996–998 (2013).

480 62. Rideout, J. R. *et al*. Subsampled open-reference clustering creates consistent, comprehensive
481 OTU definitions and scales to billions of sequences. *PeerJ* **2**, e545 (2014).

482 63. Quast, C. *et al*. The SILVA ribosomal RNA gene database project: Improved data
483 processing and web-based tools. *Nucleic Acids Res.* **41**, (2013).

484 64. Gdanetz, K., Benucci, G. M. N., Vande Pol, N. & Bonito, G. CONSTAX: A tool for
485 improved taxonomic resolution of environmental fungal ITS sequences. *BMC*
486 *Bioinformatics* **18**, (2017).

487 65. UNITE Community. UNITE USEARCH/UTAX release. (2017).

488 66. Shannon, P. *et al*. Cytoscape: a software environment for integrated models of biomolecular
489 interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

490

491

492 **Table and Figure Legends**

493 **Table 1**: Description of geographical, managemental and soil properties of the common bean

494 growing locations included in the study.

Sample same	Growing location	Elevation	Bean genotype	Weather (T, precipitation)*	Rotation history	Fertilization (per acre)	Irrigation	pH	Nitrogen (%)	Organic matter (%)
SVREC	Saginaw Valley Research and Extension Center, MI, US	190 m	CELRK	24.5ºC, 5.7 mm	Common bean, wheat, maize	Synthetic (400lbs of 15N-5 P-13K)	None	7.7 (± 0.1)	0.13 (± 0.00)	2.33 (± 0.06)
MRC	Montcalm Research Center, MI, US	280 m	Eclipse	24ºC, 7.4 mm	Common bean, maize, potato	Synthetic (200lbs of 19N-0P-19K)	Yes	5.9 (± 0.4)	0.10 (± 0.01)	2.13 (± 0.06)
NE	Scottsbluff, NE, US	1200 m	CELRK and Eclipse	27ºC, 2.18 mm	Common bean, maize, common bean	None	Yes	7.9 (± 0.1)	0.07 (± 0.01)	1.39 (± 0.18)
CO	ARDEC, Fort Collins, CO, US	1536 m	CELRK and Eclipse	31ºC, 1.4 mm **	Common bean, maize, barley	Organic (220lbs urea (2016), 3 tons manure (2015))	Yes	8.1 (± 0.2)	0.11 (± 0.00)	1.7 (± 0.15)
WA	WSU, Othello, WA, US	320 m	CELRK and Eclipse	26.3ºC, 0.2 mm	Common bean, maize, wheat	Synthetic (40lbs N/20lbs P/10lbs ZN2O5)	Yes	6.1 (± 0.4)	0.08 (± 0.01)	1.78 (± 0.10)

495 * Temperature (as max air temperature) and precipitation are averaged for the period between May

496 1st and July 31st 2017.

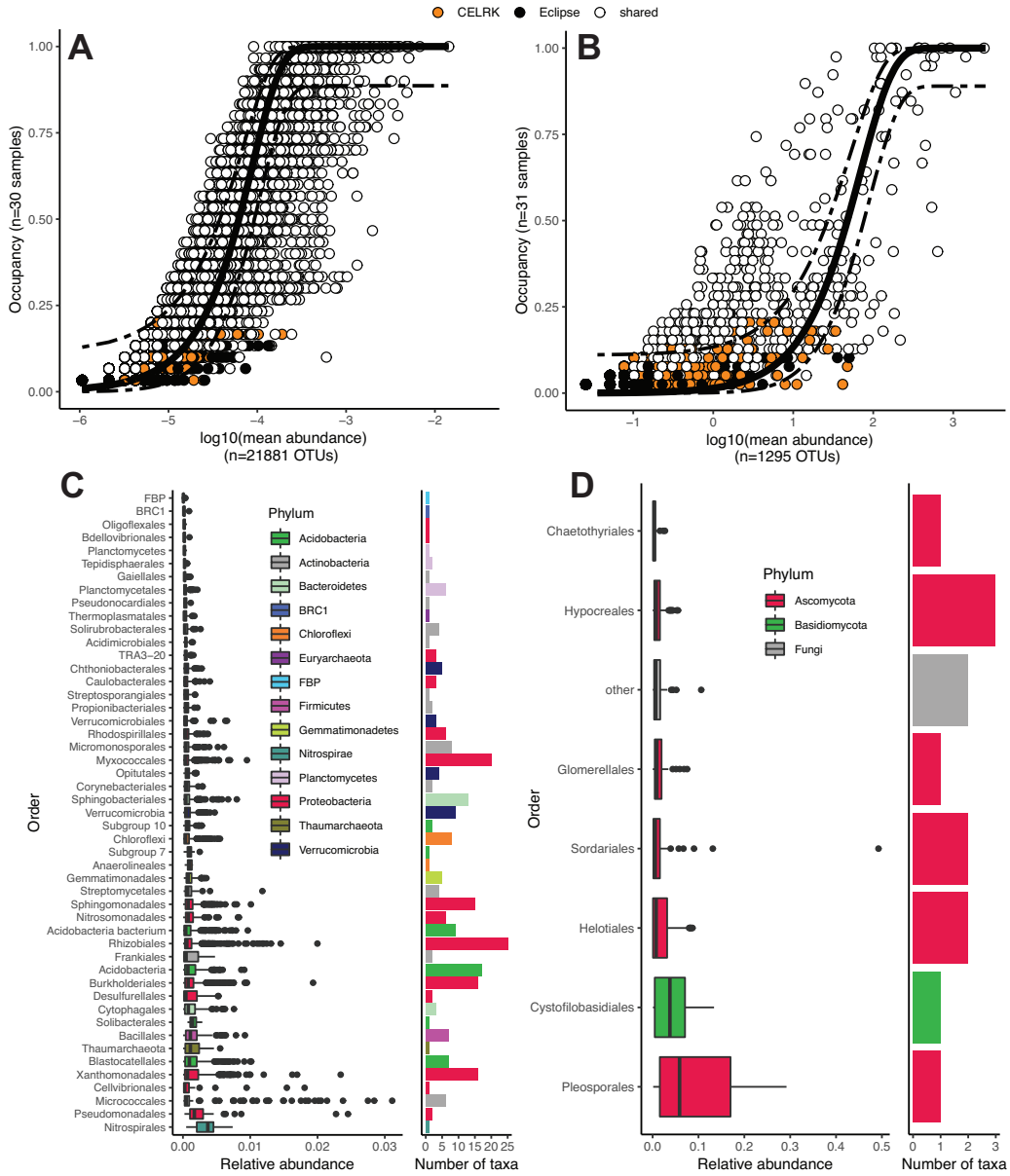
497 ** Data available only for the period between Jun 28th and Jul 31st 2017.

498

499

500

501 **Fig. 1**: Abundance-occupancy distributions were used to identify core members of the rhizosphere
502 microbiome for bacteria/archaea (A) and fungi (B). Taxa exclusive to a genotype are indicated in
503 orange (CELRK) or black (Eclipse), and taxa shared across both genotypes are white. The solid
504 line represents the fit of the neutral model, and the dashed line is 95% confidence around the model
505 prediction. Taxa with an occupancy of 1 (i.e., detected in all samples) were considered members of
506 the core. Relative abundance of these taxa is represented as boxplots, grouped by order and number
507 of taxa therein (C, D). Panels C and D are color-coded by phyla.



508

Fig. 2: A cross-continental core rhizosphere microbiome. There were 3361 shared bacterial/archaeal taxa across U.S. and Colombia rhizosphere samples, suggesting highly similar recruitment across continental scales. Forty-eight taxa were detected in all samples of both datasets, as depicted by the Venn diagram (A), and included many Proteobacteria (B). Relative abundance of the 48 U.S.-Colombia core taxa is represented as boxplots (left panel), grouped by order and dataset (Colombia/U.S.). Number of taxa per order is represented as bars (right panel). Boxplots and bars are color-coded by phylum.

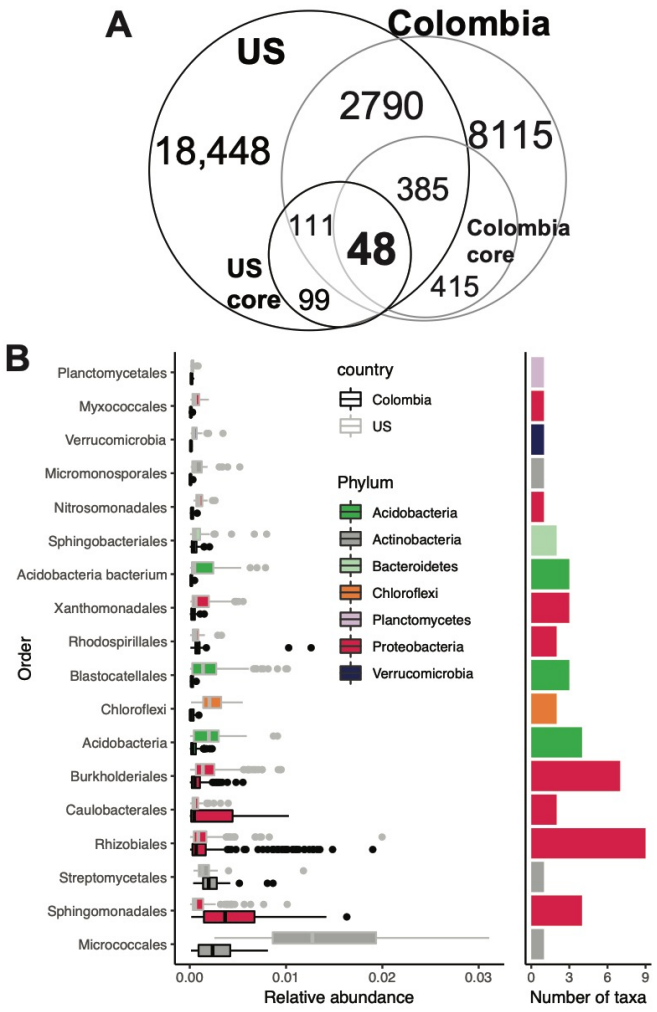


Fig. 3: Relative abundance of core taxa in the root system of the common bean during plant development by root compartment and growing location. The combined relative abundance of 48 core taxa is significantly higher in the rhizoplane (green) compared to the rhizosphere (orange) and does not show plant development dependence (A – upper panel). The combined relative abundance of remaining core taxa specific to the U.S. (n=210), tend to be high throughout the plant development but are equally abundant in both compartments (A - lower panel). Stars above box plots represent statistical significances as determined by Wilcoxon test (**** ≤ 0.0001, *** ≤ 0.001, ** ≤ 0.01, * ≤ 0.05). Z-score normalized relative abundance of the 48 core taxa across plant development, compartments and growing location (B). Taxa on the y-axis are arranged by their classification at the phylum level.

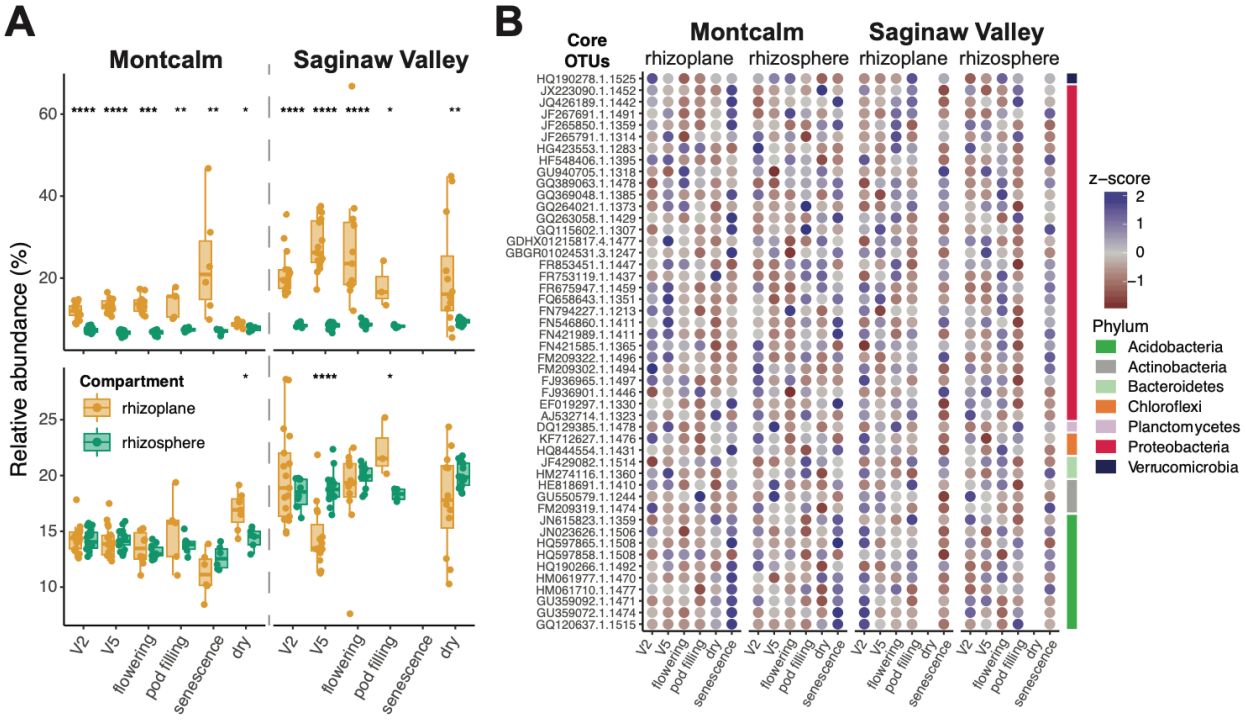
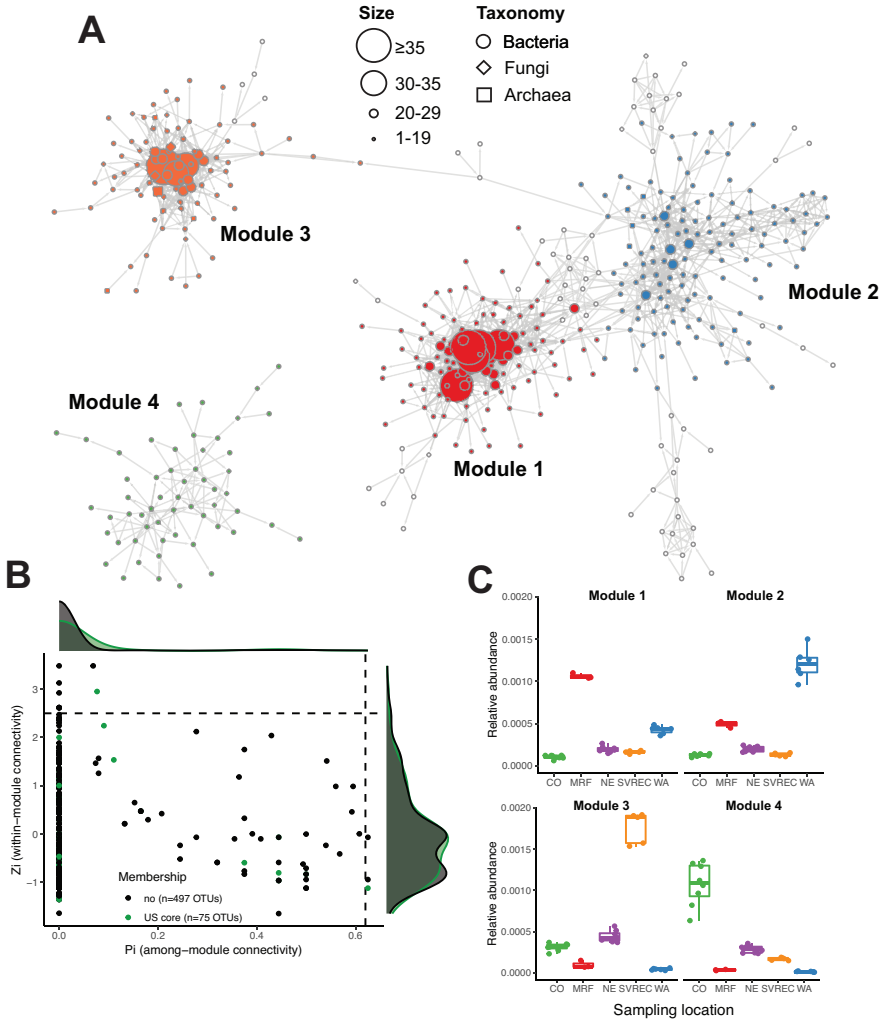



Fig. 4: Network co-occurrence analysis shows that core rhizosphere microbiome members are predominantly classified as peripheral taxa that are weakly connected, and generally clustering by the growing location. The network depicts only clusters of modules that were connected by more than 6 nodes (A). Nodes shape is representing domain associations (archaeal, bacterial or fungal) and node size is proportional to its total number of connecting edges (A). The four largest modules are generally reflective of community biogeography and distinguished by color (A). The within- (Zi) and among- (Pi) module connectivity plot was used to identify module (Pi<0.62, Zi>0.2.5) or network hub taxa (Pi>0.62, Zi>2.5), as well as connector (Pi>0.62, Zi<2.5) and peripheral taxa (B). The density plots surrounding the Zi-Pi plot represent core (green) and non-core taxa (black) (B). The relative abundances of taxa within each module is represented in box plots (C).