scientific reports

OPEN

Characterizing rhizosphere microbiota of peanut (*Arachis hypogaea* L.) from pre-sowing to post-harvest of crop under field conditions

Ankit T. Hinsu[1,2], Ketankumar J. Panchal[2], Ramesh J. Pandit[2], Prakash G. Koringa[2] & Ramesh K. Kothari[1✉]

The rhizosphere, a narrow zone of soil near plant roots, is a hot spot for microbial activity. Rhizosphere microbiota directly or indirectly benefit plants by supplementing nutrients, producing beneficial chemicals, or suppressing pathogens. Plants attract and modulate bacteria within the rhizosphere by releasing exudates. Plants also tend to select the rhizosphere microbiota based on their needs; a phenomenon termed as "rhizosphere effect". In this study, we characterized the rhizosphere microbiota of peanut plants across the crop development cycle from pre-sowing of seeds to post-harvest of crop under field conditions. The rhizosphere and bulk soil samples from different crop developmental stages were also compared. The composition of bulk soil microbiota resembled microbiota of pre-sowing and post-harvest soil and was markedly different from rhizosphere soil samples. Rhizosphere samples were enriched with multiple organisms mostly from the Proteobacteria, Firmicutes and Bacteroidota phyla. Differences in diversity were observed among the rhizosphere samples but not in bulk soil across different crop development stages. *Pseudomonas_M indica* was highly enriched during the germination of seeds. Furthermore, Plant Growth Promoting (PGP) bacteria like *Bacillus* were enriched during the middle stages of crop development but there was a decline in PGP organisms in the matured crop stage. We also observed a significant association of pH and Electrical Conductivity (EC) with the profiles of microbial community. Overall, this study portrayed the changes in rhizosphere microbiota of peanut during different developmental stages of crop and may help to design stage specific bio-strategies such as bio-fertilizer to improve crop yield.

Leonardo da Vinci said that "We know more about the movement of celestial bodies than about the soil underfoot". This remains true in the twenty-first century. Most of these mysteries can be attributed to the microscopic lives in the soil. Microorganisms, especially bacteria, are abundant in soil with concentrations as high as 10^{11} cells per gram of soil[1,2]. Soil microbes are not only involved in major biogeochemical processes but also help plants in various essential functions including nutrient acquisition, stress tolerance and pathogen resistance[3]. Plant-associated microbes can be differentiated into different types based on their locations and vicinity to plants[4]. The rhizosphere is a zone near the vicinity of roots and is a hotspot for microbial activities[3]. The rhizosphere hosts a dynamic microbial community involved in microbe-microbe and microbe-plant communications mediated by plant exudates[5]. Rhizosphere microbes respond to plant exudates and help plants via various plant growth promoting impacts including defence against pathogens[1]. This makes understanding the rhizosphere microbiome an important part of sustainable agriculture. However, the rhizosphere microbiome is very dynamic and changes in response to various internal and external factors making it an exceptionally complex ecosystem[4,6].

Sequence driven metagenomics has emerged as the approach of choice to study microbiota from various habitats including soil. Many of the previous studies have used 16S rRNA gene based community-profiling to study rhizosphere microbiota from diverse plant and crops like Arabidopsis, rice, millet, soybean, corn, barley, wheat,

[1]Department of Biosciences, Saurashtra University, Rajkot 360005, India. [2]Department of Animal Biotechnology, College of Veterinary Sciences & A.H., Anand Agricultural University, Anand 388001, India. ✉email: kothari1971@gmail.com

Sr. No	Date	Plant age from sowing (days)	Sample codes	Crop development stage	Remarks/details
1	07-05-2017	Not applicable	PS(1–5)	PreSowing	15 days prior to sowing and 5 days after addition of DAP
2	21-05-2017	5	R1(1–5), B1(1–5)	Seedling	5 days after sowing of seeds, irrigated with water from borewell during sowing
3	02-07-2017	47	R2(1–5), B2(1–5)	PreNodulating	No rain during the period and hence, no change in the crop
4	23-07-2017	68	R3(1–5), B3(1–5)	PreNodulating	Presence of nodules, last rainfall before 3 days
5	06-08-2017	82	R4(1–5), B4(1–5)	Nodulating	Rainfall during the period, presence of 2–3 groundnut pods
6	20-08-2017	96	R5(1–5), B5(1–5)	Nodulating	5–10 flowers per plant, rainfall during the period, presence of 5–10 groundnut pods
7	03-09-2017	110	R6(1–5), B6(1–5)	Flowering	10–15 flowers per plant, presence of 10–15 groundnut pods
8	17-09-2017	124	R7(1–5), B7(1–5)	Flowering	Matured crop with 15–25 groundnut pods
9	08-10-2017	145	R8(1–5), B8(1–5)	Matured	Crop fully matured with 25–40 groundnut pods. Harvest of crop from 09-10-2017
10	18-10-2017	155	PH(1–5)	PostHarvest	Post-harvest sample

Table 1. Details about the sample collection.

tomato, grapes and many others[3,4,7–12]. While many studies have looked at the effects of soil type, geographic location, crop genotypes and several other factors, very few have looked at the changes in microbiome across the crop developmental cycle[2–4,6–8,13–26]. Also, very few studies have been conducted on the peanut rhizosphere[27,28]. Peanut, also known as groundnut in some parts of world, is a leguminous crop cultivated in more than 100 countries worldwide[29]. The plant is peculiar in the respect that while the flower fertilises above ground, the pod matures in the soil. Furthermore, peanut is a rotation crop which benefits the succeeding crop by enriching soil fertility through atmospheric nitrogen fixation[30].

In this study, we used a 16S rRNA gene amplicon-based approach to study the microbiota of the rhizosphere associated with peanut and the nearby bulk soil. The study looks at the horizontal profile of microbiota throughout the crop developmental cycle starting from pre-sowing of seeds to post-harvest of the crops. The samples included seedling (S), PreNodulating (PN), Nodulating (N), Flowering (F) and Matured (M) stages of crop development. Although, the microbiota structure changes with the minutest of changes in several factors, we believe that looking at this kind of horizontal structural changes can form a foundation for a sustainable agriculture approach.

Results

In this study, we looked at the structural changes in the community of soil organisms (rhizosphere and nearby bulk soil) during the entire cycle of peanut development starting from pre-sowing of seeds to post-harvest of crop (Figure S1, Table 1). We also checked physical properties and concentrations of important macronutrients and micronutrients in the soil. We started by looking at the changes in the soil properties. Among the four groups of soil samples (PreSowing (PS), PostHarvest (PH), Bulk and Rhizosphere), significant differences (Kruskal–Wallis p-value < 0.05) were observed in pH, Electrical Conductivity (EC), and concentrations of all macro and micro nutrients except phosphate and %Organic Carbon (OC) (Table S1). Furthermore, significant differences (paired Wilcoxon-test p-value < 0.05) were observed only in pH and EC between paired samples of Bulk and Rhizosphere soil (Table S1). Significant differences (Kruskal–Wallis p-value < 0.05) were also observed in pH, EC, OC and the concentrations of Sulphur (S), Manganese (Mn), and Iron (Fe) among different developmental stages in rhizosphere soil samples (Table S1).

Diversity from sequencing data.   The V3-V4 region of 16S rRNA gene was amplified and sequenced on Illumina MiSeq using 250 × 2 paired-end chemistry. Around 6.1 million sequence reads from all 90 samples with an average count of 67,766 reads per sample was generated. The data was processed using DADA2 package in R and the default pipeline inferred 31,742 Amplicon Sequence Variants (ASVs) from 3.8 million reads (63.12%). ASVs with less than 6 supporting reads and samples with less than 10,000 processed reads were filtered out, leaving behind 21,764 ASVs from 88 samples, which were analysed further. Five databases used for taxonomy assignment showed different levels of annotation (Figure S2) however, assignment from GTDB r89 was considered for further analysis (more details in discussion section). All these ASVs were taxonomically classified as Bacteria and assigned uniquely to 38 phyla, 80 classes, 195 orders, 313 families, 664 genera and 90 species.

The number of ASVs ranged from 425 (in R1 group) to 3,030 (in B8 group) (Fig. 1). No significant differences were observed among the four groups of samples (Kruskal–Wallis p-value = 0.23). Significant differences were observed among all the collections in rhizosphere samples (Kruskal–Wallis p-value = 0.013) but not in samples from bulk soil (Kruskal–Wallis p-value = 0.094). Shannon diversity index, which accounts for the relative proportion of each ASV, was observed in range of 2.94 (in R1 group) to 7.43 (in B7 group) (Fig. 1). Unlike Observed ASVs, Shannon index differed significantly among the four sample types (Kruskal–Wallis p-value < 0.001) and among all collections of rhizosphere samples (Kruskal–Wallis p-value = 0.0007) but not among all collections of bulk samples (Kruskal–Wallis p-value = 0.095). Furthermore, no significant differences were observed among samples of PS, PH and bulk soil. Significantly higher Shannon diversity (Kruskal–Wallis p-value < 0.05, except Collection-3 where p-value = 0.056) was observed in bulk samples compared to rhizosphere samples of same collection except from samples of Collection-8.
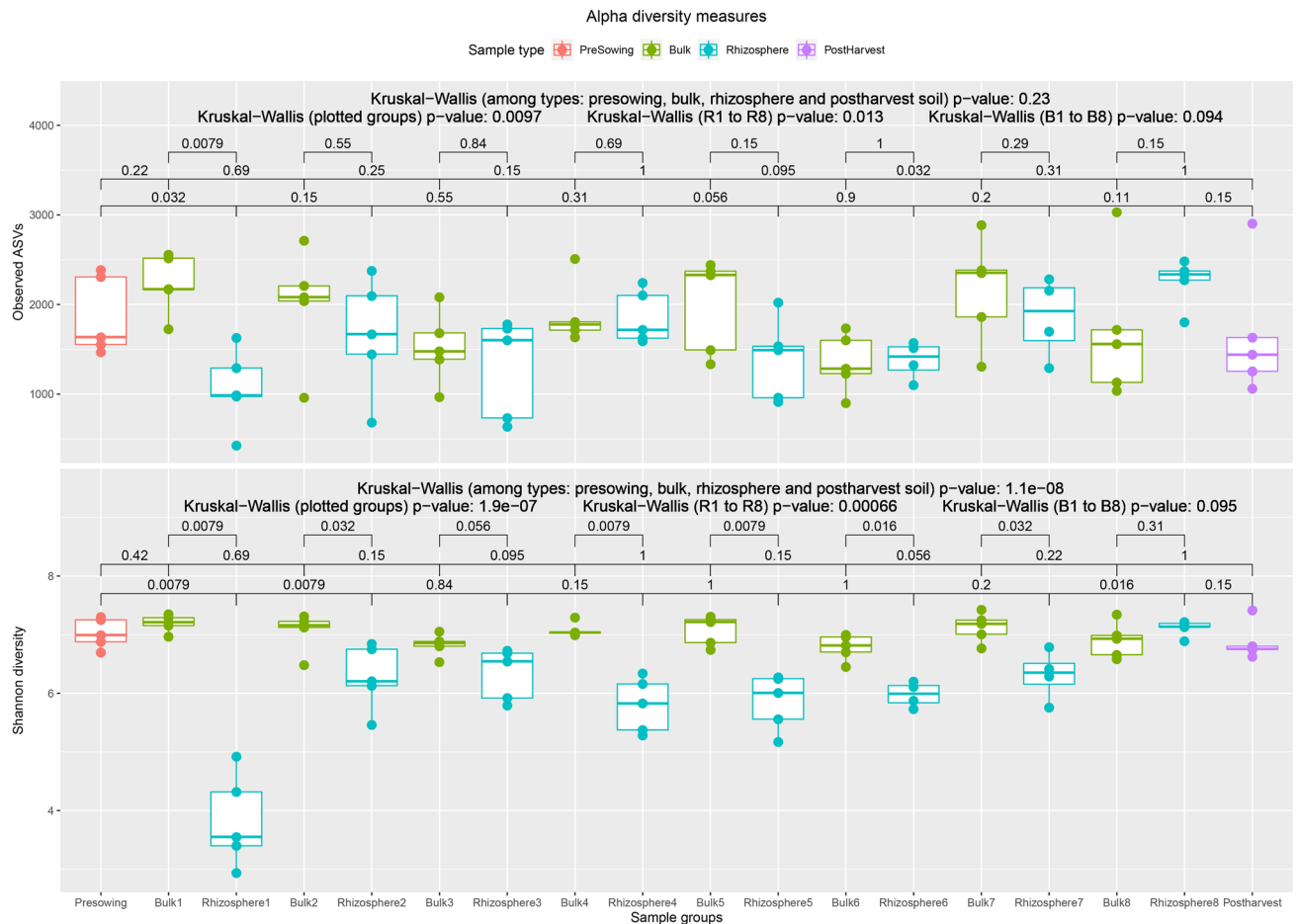
Figure 1. Alpha diversity measures Observed ASVs (top), and Shannon Index (bottom) plotted against sample groups. p-value from Wilcoxon-test between groups are given by bracket pointing to groups. p-value from Kruskal–Wallis test among four types/groups of samples is mentioned on the top. Second line mentions p-value from Kruskal–Wallis test of all groups (as plotted), all rhizosphere sample groups/collections and bulk soil sample groups/collections.

Taxonomic content of microbial communities. All the ASVs were classified as Bacteria and were distributed in 38 phyla and 664 genera. Common soil phyla like Acidobacteriota, Actinobacteriota, Planctomycetota and Proteobacteria were abundant (median relative abundance > 5%) (Fig. 2A). Additionally, Firmicutes and Bacteroidota phyla had higher abundance in rhizosphere samples. At the genus level, *Enterobacter_D, Bacillus_W, Luteitalea, UBA2421, QHWT01, Pseudomonas_M,* unknown member from Sphingomonadaceae family and some unknown bacteria were the most abundant (mean relative abundance > 2%) (Fig. 2B). Clear differences were observed between rhizosphere samples and other samples. R1 group samples were overly abundant with Proteobacteria phyla (~70%—85%) and genera like *Pseudomonas_M, Enterobacter_D* and unclassified genera of *Rhizobiaceae* and *Pseudomonadaceae* family, which explains the lower diversity in the group. Very few highly abundant ASVs were assigned up to species level. However, many of these ASVs were dominant in rhizosphere samples (Figure S3). The most abundant species were *Pseudomonas_M indica, Bacillus_AK korlensis* and *Enterobacter_D cloacae,* all of which belonged to most abundant genera in rhizosphere samples. Other species observed abundantly in rhizosphere samples include *Microvirga vignae, Inquilinus limosus, Flavobacterium anhuiense, Stenotrophomonas maltophilia_J, Chryseobacterium geocarposphaerae, Sphingopyxis macrogoltabida, Exiguobacterium profundum,* and *Cupriavidus alkaliphilus.*

Differences between four groups of samples. Non-Metric Dimensional Scaling (NMDS) of Bray–Curtis distance was performed to check the patterns of separation between microbial communities across groups. Samples from bulk soil, PS and PH groups formed a separate cluster from rhizosphere soil samples (Fig. 3A). Separate clusters were also observed for R1 and R8. Permutational multivariate analysis of variance (PERMANOVA) corroborates that four groups of samples showed significant differences ($R^2 = 0.271$, p-value < 0.001). Furthermore, this differentiation also associated significantly (p-value < 0.05) with the distribution of pH, EC, Shannon diversity and Observed ASV. These differences were also evident from the observation that almost half of the observed genera (420 of 891, including the most abundant genera) differed significantly (Kruskal–Wallis BH p-value < 0.05) among the four types of samples (Table S2, Figure S4). However, limited significant changes were observed among bulk, PS and PH samples (Table S3).
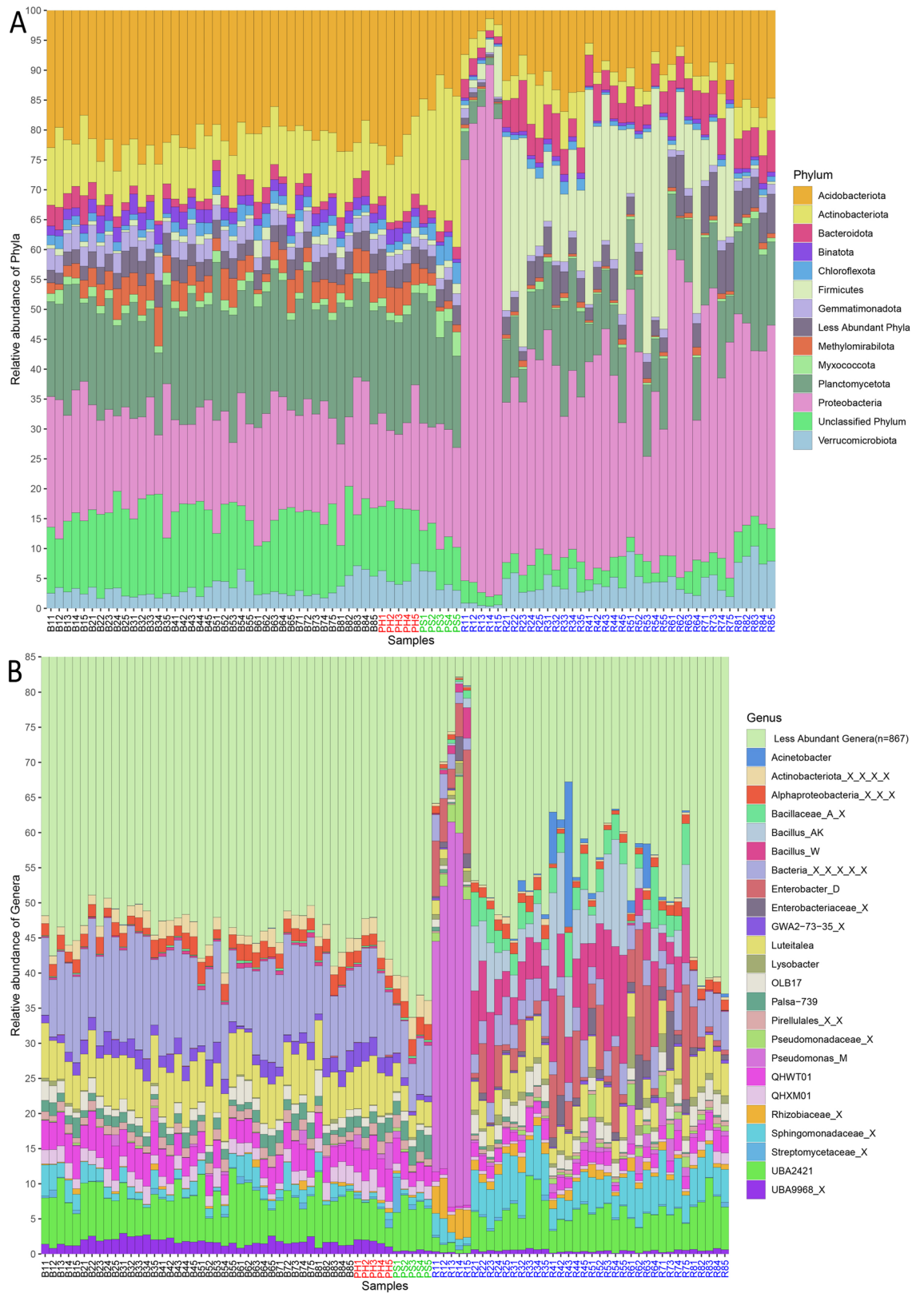
Figure 2. Taxonomic distribution at (**A**) phylum level and (**B**) Genus level. Only the top taxa are plotted for both levels. Sample names are coloured by type of sample.
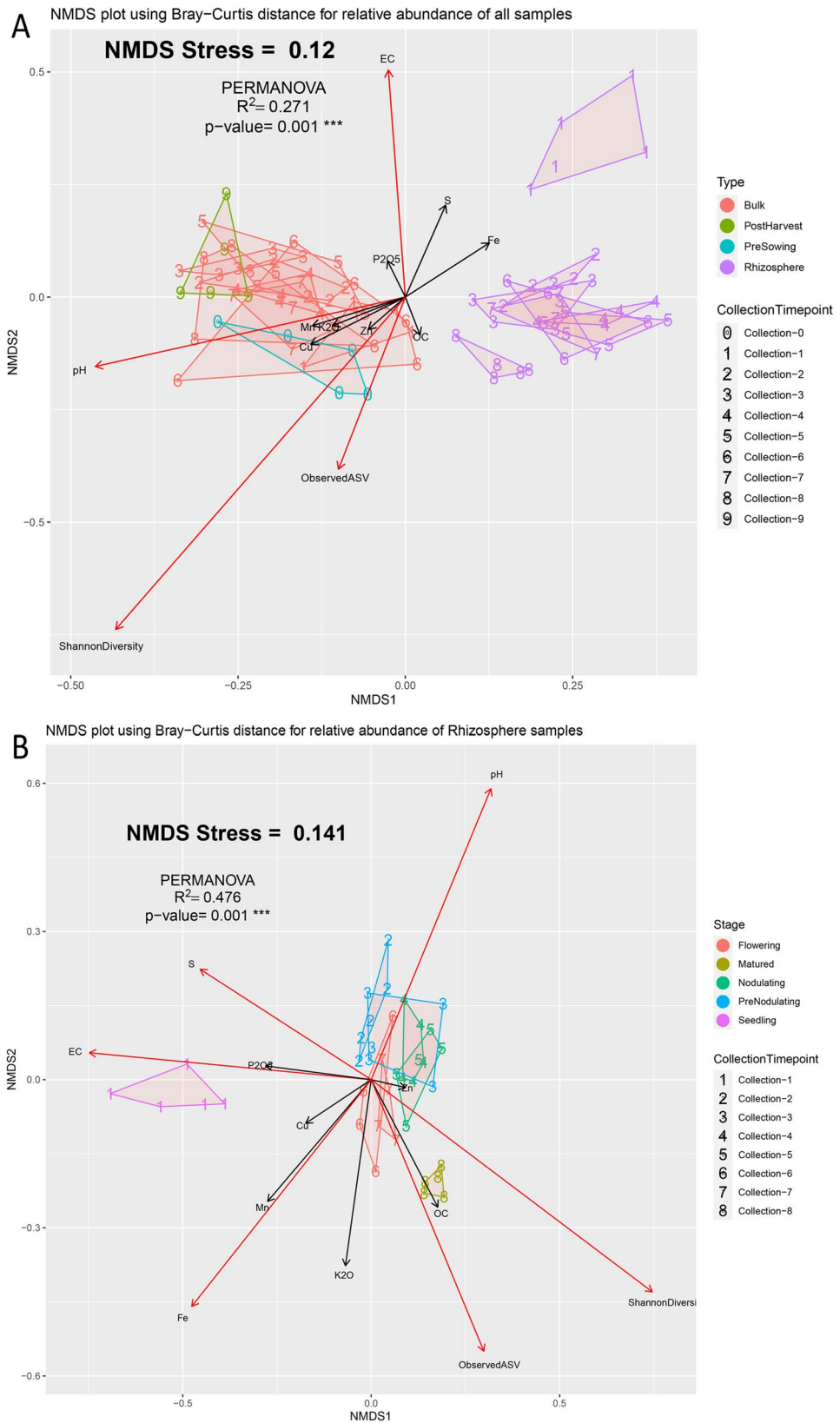
Figure 3. NMDS plot of Bray–Curtis distance calculated from (**A**) all samples and (**B**) Rhizosphere samples only. Collection number of each sample is used as shape to denote the sample. Environment fit vectors representing physical parameters and nutrient concentrations are shown as arrows. Vectors with significant associations are shown in red coloured arrow.

The differences between types of samples were also observed in terms of significant (adjusted p-value < 0.01) log2FoldChange (lFC) of ASVs between pairs of groups as checked using DESeq2 (Figure S5). There were more significant differences in lFC between ASVs in rhizosphere samples compared to bulk (n = 827), PH (n = 254) and PS (n = 110) with lFC as high as 30. Significantly differing ASVs from phyla Firmicutes, Firmicutes_A, Firmicutes_I and Patescibacteria were more enriched in rhizosphere samples. Specifically, genera like *Pseudosphingobacterium, Chryseobacterium, Neorhizobium, Haloferula, Verrucomicrobium, Sphingosinicella*, and members of *Bacillaceae_A, Pseudomonadaceae, Enterobacteriaceae, Alteromonadaceae, Rhizobiaceae* and *Burkholderiaceae* family were present with lFC more than 20 in rhizosphere samples. While very few ASVs differed significantly between PS and bulk (n = 3); PS and PH (n = 13); bulk and PH (n = 15).

Differences between pairwise bulk and rhizosphere samples.

The intriguing differences between bulk and rhizosphere samples compelled us to look at the differences between these groups for each collection. In total, 446 genera out of 891 differed significantly (Wilcoxon test p-value < 0.05) between bulk and rhizosphere samples (Table S3). Furthermore, 41, 97, 126, 93, 86, 51, 47 and 56 ASVs were observed to have significant (adjusted p-value < 0.01) lFC between bulk and rhizosphere samples from Collection 1 to 8, respectively (Figure S6). ASVs belonging to Proteobacteria, especially multiple ASVs of *Enterobacter_D, Pseudoxanthomonas_A, Cupriavidus* and some unknown members of *Pseudomonadaceae* and *Enterobacteriaceae* family, were more abundant in rhizosphere (with lFC > 10) across all collections. Additionally, ASVs from Firmicutes phylum (*Bacillus_AK* and some unknown *Bacillaceae_A* family) were also abundant in R1 to R7, but were significantly enriched (lFC > 10, p-value < 0.01) in B8. Some changes were also observed confirming the developmental stage-like patterns. For example, members of *Rhizobia* (*Neorhizobium* and multiple ASVs assigned to unknown *Rhizobiaceae* family), are nitrogen fixers associated with root nodules of legumes, and were observed to be significantly enriched with a lFC as high as 40 in Collection 1 to 5, corresponding to the early S, PN and N developmental stages, but were not significantly different thereafter.

Changes among rhizosphere samples associated with crop development stages.

Lastly, we looked at the changes in the rhizosphere samples with respect to crop developmental stages. NMDS ordination on Bray–Curtis distance revealed a separate and distinct cluster for R1 (S stage) and R8 (M stage) (Fig. 3B). Except for 2 samples from R3, all other samples from R2 to R7 were grouped according to developmental stage (PN, N and F). However, all these collections formed clusters quite close to each other. The difference among developmental stages were further confirmed through PERMANOVA where significant difference was observed ($R^2 = 0.476$, p-value < 0.001). A pairwise-adonis between all pairs of collections were significant except for R4–R5 (p-value = 0.215), R6–R7 (p-value = 0.557) and R2–R3 (p-value = 0.057), all of which are the collections from the same developmental stage (Table S4). An environmental fit of all variables also revealed a significant association of pH, EC, concentrations of S and Fe, Observed ASVs and Shannon diversity. Further, increase in OC, Shannon diversity and Observed ASVs with crop development and decrease in EC, phosphate, S concentration with crop development was observed from ordisurf plots (Figure S7).

At the phylum level, 28 out of 30 phyla (with mean relative abundance > 0.0001) were significantly different (Kruskal–Wallis, BH p-value < 0.05) among developmental stages (Table S5, Figure S8). We believed that this could be happening because of the hugely different and less diverse samples in R1 group. Removing the R1 group samples decreased the number of significantly differing phyla (Kruskal–Wallis, BH p-value < 0.05) to 22 out of 30. Similarly, at genus level, 306 out of 611 and 236 out of 611 genera (mean relative abundance > 0.00001) differed significantly (Kruskal–Wallis, BH p-value < 0.05) among developmental stages with and without R1 group, respectively (Table S6, Figure S9). Many of the higher abundance genera like *Acinetobacter, Bacillus_AK, Bacillus_W*, unknown genera from Bacillales order and *Bacillaceae* family seemed to increase with crop development stage till N stage and then decrease, forming a bell-like shape. While other abundant genera like *Chthoniobacter, QHWT01, UBA2421* and unknown genera from *Burkholderiaceae* family seemed to increase with crop developmental stages throughout the life cycle. However, very few ASVs (less than 25) were observed to differ significantly (adjusted p-value < 0.01) as per DESeq2 (Figure S10). Around 53 ASVs differed significantly (p-value < 0.01) between PS and S stage, most of which were PGP bacteria from the Proteobacteria. Further, 23, 22 and 3 ASVs differed significantly between S, PN, N and F stage, respectively. Nineteen ASVs, most of which were from Bacillales order, reduced (lFC > 5) significantly (p-value < 0.01) in M stage compared to F stage. Similarly, 139 ASVs, most of which reduced with lFC > 5, differed significantly between M and PH stage.

There were 118 genera identified as part of core microbiome of rhizosphere samples with minimum abundance 0.1% across 40% samples (Fig. 4). These taxa were further correlated among each other to see the co-occurrence pattern. Six different clusters of genera could be made out from the significantly (p-value < 0.05) correlating genera (Fig. 5). Two clusters were observed to be showing a negative correlation with all the other genera. These contained genera like *Neorhizobium, Pseudomonas_M, Xanthomonadaceae_X, Rhizobiaceae_X, Pseudomonadaceae_X, Enterobacter_D* and *Enterobacteriaceae_X*, many of which were overly abundant in S stage. A separate cluster could also be made out containing mostly Bacilli class members like *Bacillus_AK, Bacillus_W, Bacillus_AA, Bacillus_Y* and unknown members of *Bacillaceae_A* family, Bacillales order and Bacilli class. Many of these genera, especially higher abundant ones showed negative correlation with all other genera probably explaining their increase during PN, N and F stages. Apart from these, two other clusters were observed which included majority of the genera but showed no specific pattern of correlation.
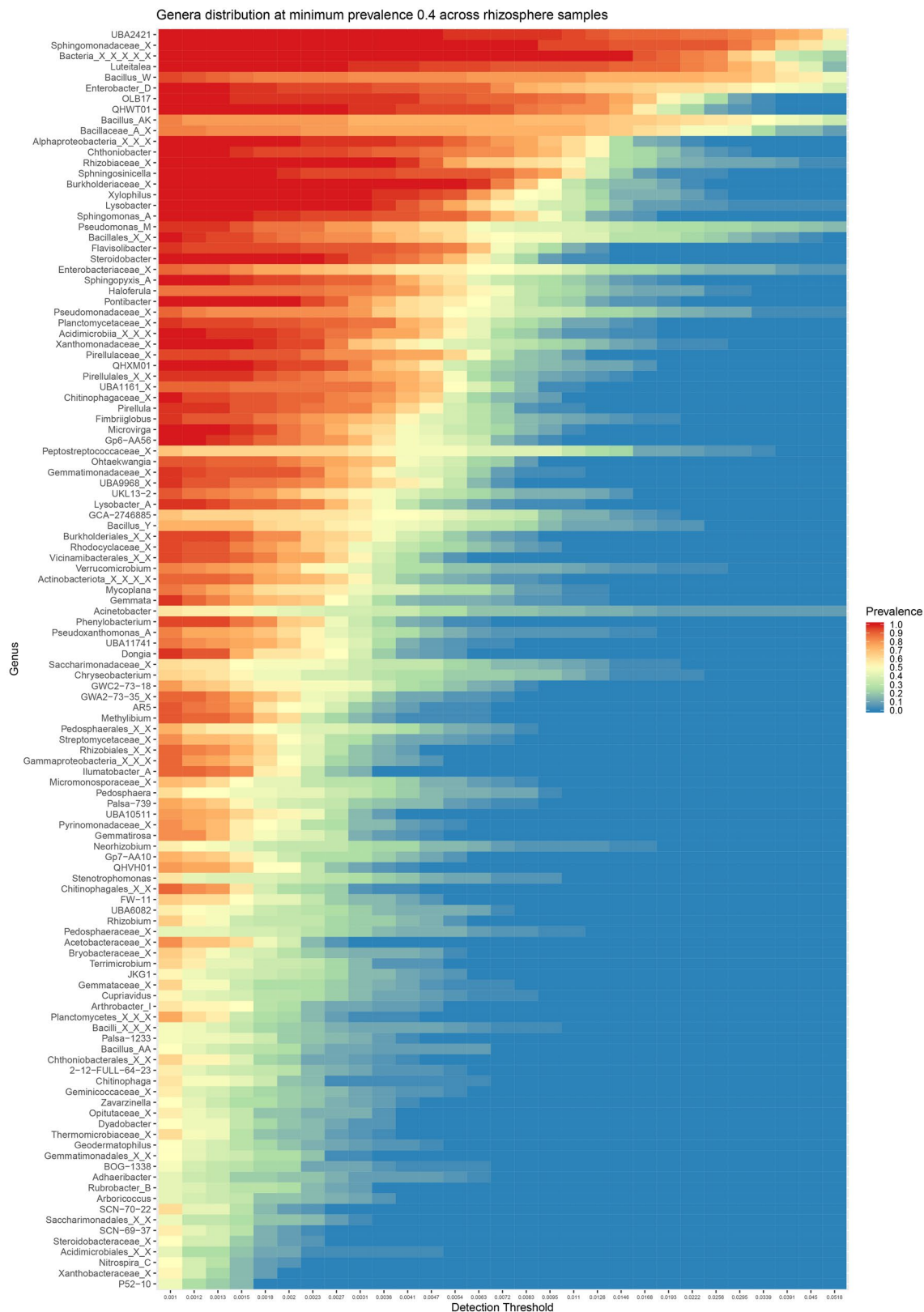
Figure 4. Plot representing core microbiome from rhizosphere samples. The plot compares prevalence of genus in samples across varying levels of abundance. Only the genera with minimum prevalence of 0.4 at 0.001 abundance are plotted.
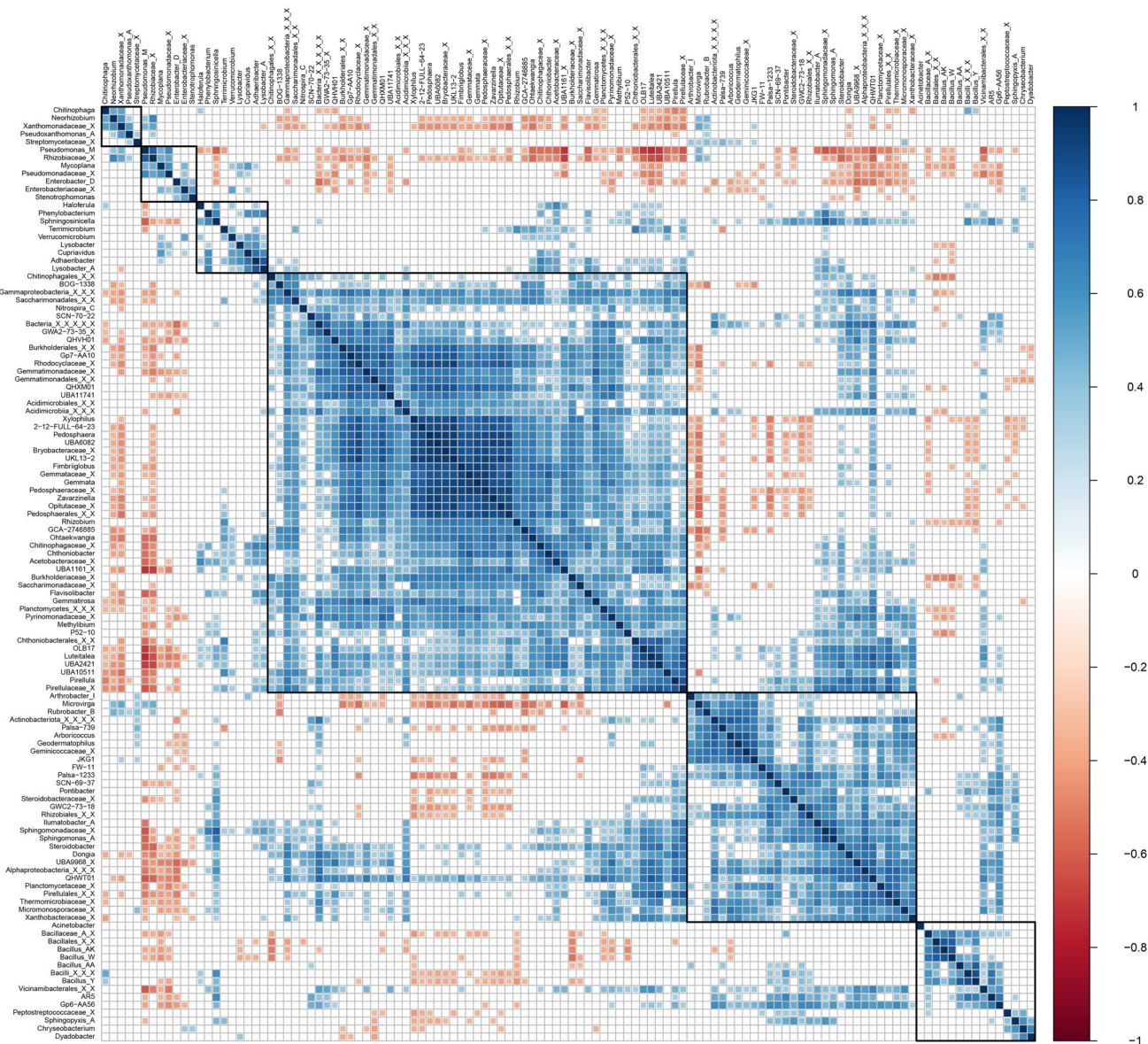
Figure 5. Correlation plot among genera from core microbiome. Only the significant (p-value < 0.05) correlations are plotted.

Discussion

The 16S rRNA gene sequencing-based approach is commonly used to characterise various microbiomes including soil and rhizosphere under various conditions and stresses. In this study, we analysed the bulk soil and rhizosphere microbiota using Illumina based 16S rRNA gene sequencing. To provide more resolution to the analysis, we opted for the DADA2 approach for data analysis[31]. DADA2 is a denoising algorithm developed specifically for Illumina data and infers Amplicon Sequence Variants (ASVs) with differences of a single nucleotide and hence, retains strain-level information. In our study, we could observe several ASVs distinguished to species level (Figure S3) including some ASVs with higher abundances. However, the power of analysis also depends on the choice of database for taxonomy assignment. We used SILVA v132, RDP trainset 16, RDP trainset 16 + RefSeq (sequences from NCBI accessed on 14-05-2018) and GTDB versions 86 and 89 for taxonomy assignment and decided to use GTDB v89 for the entire analysis. Our decision to use GTDB was based on the observation that higher number of variants were annotated at genus level in GTDB (Figure S2) and also because of the taxonomy lineage assign-ment method in GTDB[32,33]. GTDB is a curated database with comprehensive genome-based taxonomy based on monophyly and relative evolutionary divergence of taxa, which is an added advantage while annotating ASVs. The reclassifications by GTDB works well by distributing/reclassifying popular genera into several novel ones[32]. For example, *Pseudomonas* genus was divided into 15 genera, 1 retaining the name *Pseudomonas* consisting of *P. aeruginosa* while, 14 other labelled with a letter from A to O (*Pseudomonas_A* to *Pseudomonas_O*)[34]. This gives higher resolution to the observed organisms in the study. For example, we observed higher abundances of *Pseudomonas_M*, which is represented by a single species *P. indica*, and not any other *Pseudomonas* genera suggesting the differences in levels among all *Pseudomonas* genera. Similar observations were also made with

Bacillus genus where ASVs classified as *Bacillus_AK, Bacillus_W* and some unclassified *Bacillaceae_A* (which includes *Bacillus_AK, Bacillus_W* and other genera) were the most abundant.

Previous studies have focussed on the microbiome from different compartments of soil[4]. In this study, we compared the rhizosphere microbiota with nearby bulk soil microbiota. In our study, we observed significantly higher Shannon diversity in bulk soil samples compared to rhizosphere, except for samples from M stage. This type of pattern is reported in several earlier studies[3,17,23,24], while many other studies have reported mixed results[7,14]. While the rhizosphere effect (i.e., compositional and functional difference of microbiome between rhizosphere and bulk soil) attracts beneficial microbes and increases the microbial activity in the rhizosphere, the bacterial diversity remains lower compared to bulk soil apparently[3]. We also observed the significant differences in rhizosphere microbiota diversity across crop cycle but not in bulk soil which suggests that the development stage of crop is one of the major drivers of the rhizosphere effect. Although samples of PS and PH were collected from barren farm, we were expecting more similarity between PH and rhizosphere samples. However, PH samples were very similar to bulk samples. This shows that bacteria community changed from rhizosphere-like to bulk-like within span of 10 days in the absence of plants.

In accordance with previous studies, we observed higher abundances of Proteobacteria, Acidobacteria (named as Acidobacteriota in GTDB), Actinobacteria (named as Actinobacteriota in GTDB) and Planctomycetes (named as Planctomycetota in GTDB) in all samples[3,22]. Additionally, higher abundance of Firmicutes, Bacteroidetes (named as Bacteroidota in GTDB) and Verrucomicrobia (named as Verrucomicrobiota in GTDB) were also observed in the rhizosphere samples. A higher abundance of Firmicutes has been reported in the rhizosphere in a few earlier studies[16,19,25,35] including a study on peanut microbiome[27], however this was not observed in other studies on the peanut rhizosphere[28,36]. The majority of DESeq2 based enriched ASVs in rhizosphere samples also belonged to Proteobacteria, Firmicutes and Bacteroidota phyla while ASVs enriched in bulk soil belonged to Acidobacteriota, Actinobacteriota and Planctomycetota (Figure S5). Similar findings have been reported in earlier studies where plant growth promoting bacteria from Proteobacteria and Firmicutes phyla were enriched in rhizosphere[3,26,37]. Overall, *Pseudomonas_M, Acinetobacter, Enterobacter, Bacillus_W, Bacillus_AK, UBA2421* were the most abundant genera in rhizosphere. All of these are reported to show PGP activities. For example, *Pseudomonas* and *Acinetobacter* are well studied for their biofilm properties[38], and siderophore production[39], *Enterobacter cloacae* (most of *Enterobacter* in present study were annotated as *Enterobacter cloacae* in rhizosphere) are associated with heavy metal and salinity stress removal[40,41], and *Bacillus* are reported to possess several beneficial activities including solubilization of phosphate, nitrogen fixation and siderophore production[42]. *UBA2421*, a genus observed with > 3% abundance in all samples, is a yet uncharacterised organism belonging to Planctomycetota phylum. To date, this genera has only been described in databases as a Metagenome assembled genome (MAG) with no known cultured isolates[43]. Several other genera described exclusively by MAGs are expected depending on whether the selected database includes MAGs. This highlights the fact that there are several higher abundant organisms whose roles are yet to be evaluated in environmental context.

Earlier studies have hypothesized and reported that plants initially perform a general recruitment near the root vicinity and then allows microbes with species-specific signatures to colonise within[9]. Our study shows similar results wherein Proteobacteria and specifically *Pseudomonas_M indica* was observed in higher abundance in R1 group and then decreased in abundance. Only two ASVs (ASV1 and ASV31726) were annotated as *Pseudomonas_M* and present in sufficient abundance for inclusion in the study (mean rel. abundance > 0.0001). These ASVs were observed in all samples, including PS samples, at lower abundances. This implies that the organism was already part of soil but enriched during the germination of seeds, similar to what have been observed in previous studies[4]. It could be speculated that *Pseudomonas_M* was part of the bulk and PS soil microbiota, recruited on seed germination, colonised the roots, and then abundance reduced in rhizosphere as the plant matured. It is possible that stable higher abundance of *Pseudomonas_M* could be present in rhizoplane and endosphere microbiota. Many other studies have reported N-fixing and other PGP properties of *Pseudomonas* including *P. indica*[44] and reports are available for their use in fields as growth promoter. *Pseudomonas* are also shown to produce biofilm, which is crucial for colonization of bacteria on plant root surface[45]. *P. indica* was first isolated from oilfields in Gujarat, India while the sample source for this study is also from Gujarat, India[46]. Furthermore, earlier study have also reported higher abundance of Proteobacteria (as high as 50–55%) in the rhizosphere up to 13 days after germination and a comparatively higher proportion in the rhizoplane[4]. Similar to our results, the previous study also highlighted that plants recruit the microbiome from soil rapidly, stabilising the communities within two weeks. Since this study did not involve sterilization of seeds, introduction of microbiota from seed is also a possibility.

Many recent studies have reported on the transmission of microbiota through seeds[47,48]. These microbes inhabit the rhizosphere after germination of seeds through cotyledon defoliation and root development. These microbes are also linked to seedling survival during early stages[49]. It has also been postulated that microbiota of rhizosphere and several other endophytic microbes colonise the seed and are transmitted. PGP organisms like *Pseudomonas*, *Bacillus*, *Acinetobacter*, *Pantoea*, etc. are commonly identified as seed endophytes with the potential to attract more beneficial microbes and suppress pathogens[49,50]. In the current study, there were fewer genera found exclusively in R1 (n = 52) compared to PS (n = 143) (data not shown). Since the farm has sown the same variety/genotype of seeds for the previous three years, it can be speculated that the diversity was already introduced in the soil, explaining the lower number of genera observed exclusively in R1 samples. However, these genera could also be introduced to the soil and have their growth elevated by irrigation. The farm also followed the practice of using water from wells during early sowing followed by a dependence on rain for water thereafter. This led to an extensive dry period without water to plants of approximately forty days during the sampling cycle. However, the samples collected during this period (R2) and after the first rainfall (R3) showed similar microbiota profile (pairwise Adonis p-value = 0.057). Moisture content is one of the important abiotic

factors influencing microbiota structure. Therefore, we tried to minimise possible biases arising due to favourable environment provided by moisture content by collecting samples a minimum of three days after the last rainfall[51].

It was postulated way back in 1995[52] that plants release exudates to modulate surrounding rhizobacterial communities. It was also postulated that secretion of exudates changes with change in plant development stage[15]. The plants have co-evolved with the surrounding microbes in order to utilize their potential benefits. The plants regulate the microbes through exudate secretion and selects for the best organisms which can help in plant growth promotion through different ways[53]. This makes the rhizosphere microbiome highly dynamic and dependent on abiotic factors as well[10]. One of the objectives of this study was to observe the rhizosphere community throughout the cropping season by looking at crop developmental stage specific trends. In this study, we observed significant changes across R1 to R8 (Fig. 1) and all stages in rhizosphere samples (data not shown) for both number of observed ASVs and Shannon diversity. However, the majority of these variations were due to the S stage samples and M stage samples. The effect of these two stages is visible by looking at the number of significant phyla across developmental stages (Table S5, Figure S8). Such changes were also observed at genus level. For example, PGP organisms like *Bacillus* had comparatively higher abundance in N stage followed by PN and F stages of crop cycle. Notably, abundance of *Bacillus* was least during S and M stage of crop cycle. *Bacillus* are known for their Phosphate-solubilising efficiency which explains their role and presence in higher amount during PN, N and F stages[54,55]. Similar trends were also observed for *Acinetobacter* which are reported to possess several PGP activities[56]. Modulation of rhizosphere microbes has commonly been attributed to change in soil pH[21,57]. In our study, we observed a strong significant link of pH and EC with rhizosphere microbiota similar to previous studies[18,20]. Low pH (< 8) and high EC (> 1.1) was observed during germination of seeds among all rhizosphere samples. While significant changes were observed in the diversity among rhizosphere samples, this was not reflected in DESeq2 based differential abundances. The observation from DESeq2 enrichment corroborates that the differences in the microbiome during the crop cycle are minimal, but these changes are more prominent when compared with PS and PH soil microbiome. Moreover, changes were observed for M stage sample which was much different than previous stage samples and started exhibiting likeness to the bulk soil samples (decrease of Firmicutes and increase of Acidobacteriota and Planctomycetota). Overall, the observations from DESeq2 enrichment and significant differences among taxa indicate that developmental stage dependent effects are more prominent during early or germinative stage and during late or maturation stages and includes recruitment of beneficial and specific taxa . The differences in the intermediate stages, including PN, N, and F or the vegetative and production stage of crop are less prominent and involves modulating the abundance of favourable microbes by a minor degree. However, analysing the multiomics data from functional metagenomics and metabolomics can paint a broader and more detailed picture of the interactions in the rhizosphere.

Methods and materials

Experimental design and sample collection. The experiment was conducted from the part of farm (~ 13,000 m² area from total farm area of ~ 147,000 m²) located near RanaKandorana village (21.618734 N, 69.865700 E) in Gujarat, India during the cropping season of 2017. The farm had history of continuous sowing of G-20 variety of peanut (*Arachis hypogaea* L.) for the previous three years. DAP (Diammonium Phosphate) was added to the farm before sowing the seeds and seeds were irrigated with well water after sowing. The farms in the region followed the practice of early sowing, irrigating once with borewell/well water and then depending on the rainwater for irrigation throughout the season. The selected farm also followed the same practice.

The first samples were collected 15 days prior to sowing of seeds (and five days after addition of Diammonium phosphate fertiliser) (samples termed as PreSowing/PS/Collection-0) (Table 1, Figure S1). During this collection, the field was levelled and tiled ready for sowing of seeds. No intermediate crops were cultivated during the period from last harvest of peanut, making PS soil free from any rhizospheric effects. Next samples were collected 5 days after sowing of seeds, when the seedlings were visible above the ground (samples termed as Collection-1 further grouped into Bulk-1/B1 and Rhizosphere-1/R1). Next samples (Collection-2; Bulk-2/B2; Rhizosphere-2/R2) were collected 6 weeks after Collection-1 because of no rainfall during the period and no irrigation leading to no development in crop (this is common with farms following the practice of early sowing). Next collection (Collection-3; Bulk-3/B3; Rhizosphere-3/R3) was done after 21 days of Collection-2. All subsequent samples (Collection-4 to Collection-8) were collected 14 days after previous collection. Crop was harvested on the very next day of Collection-8. All these samples covered different crop development stages: Seedling (Collection-1), PreNodulating (Collection-2 and Collection-3), Nodulating (Collection-4 and Collection-5), Flowering (Collection-6 and Collection-7) and Matured (Collection-8). Samples were also collected 10 days after harvest of crop (samples termed as PostHarvest/PH/Collection-9). During this collection, the field was completely barren, and samples were collected from the positions where crops were present to check the exact effect on microbiota postharvest.

Rhizosphere (R) and bulk (B) soil samples were collected during all collections from 1 to 8. While PS (Collection-0) and PH (Collection-9) samples were collected from barren farm, B samples were collected from soil between two lines of crop taking care that no roots were visible in dug soil. B soil and samples of PS and PH was collected from depth of 5–10 cm from soil surface. Approximately five gm soil was collected in sterile 50 mL tube for microbiome analysis while around 700gm soil was collected in another sterile container for physicochemical analysis. For collecting the rhizosphere samples, nearby soil was gently brushed aside, and the plants were carefully uprooted. The plants were selected randomly during each collection. The plants were shaken to remove loosely attached soil, which was collected in sterile container for soil property determination. The roots were washed in sterile normal saline (1% NaCl) in flask and the washed soil was then collected in sterile 50 ml falcon tubes for microbiome analysis[14]. During earlier development phases, multiple adjacent plants were uprooted to collect enough soil. The samples for microbiome analysis were transported to lab in − 20 °C portable freezer

and then stored at − 20 °C till further processing. The samples for soil property estimation were transported at room temperature. From each collection, 5 samples of rhizosphere soil (n = 40, 8 collections × 5 replicates) and 5 samples of nearby bulk soil (n = 40, 8 collections × 5 replicates) were collected. Also, five replicates were collected for each PS and PH stages making total of 90 samples.

Sample processing. Soil samples were submitted to a government-approved soil testing facility (Gujarat State Fertiliser Company, GSFC, Vadodara, India) for physicochemical analysis. The samples were tested for physical properties (pH and Electrical Conductivity), macronutrients (%Organic Carbon, concentrations of Phosphate and Potassium) and micronutrient (concentrations of Iron, Sulphur, Manganese, and Zinc and Copper).

The rhizosphere samples for microbiome analysis were thawed and homogenized. The tubes were then centrifuged at 10,000 rpm for 5 min[14]. At this speed, all microbial cells along with soil particles will settle down leaving behind buffer in supernatant which was discarded. The soil was then mixed properly and taken immediately for DNA extraction. Bulk, PreSowing and PostHarvest soil samples were collected without buffer and hence, were processed directly for extraction. DNA was extracted from 1 gm of soil using Qiagen PowerSoil DNA Extraction kit (Qiagen, Germany) following the manufacturer's instructions.

Library preparation and sequencing. 16S rRNA gene amplicon sequencing libraries were prepared from 12.5 ng DNA as starting material following double-pass PCR protocol as given in Illumina 16S library preparation guide (Illumina, USA). The V3-V4 region of 16S rRNA gene was targeted using 341F and 785R primers fused with Illumina adapters[58]. Libraries were verified on Agilent Bioanalyser (Agilent, USA) and quantified using Qubit v3 (ThermoFisher Scientific, USA). The libraries were sequenced on Illumina MiSeq using 250 × 2 v2 chemistry.

Data analysis. The raw fastq data was analysed using the Divisive Amplicon Denoising Algorithm 2 (DADA2) pipeline ("dada2" package version 1.14) in R v3.6.1 following the steps given at https://benjjneb.github.io/dada2/tutorial.html[31,59,60]. Taxonomy of ASVs was assigned using multiple databases namely, SILVA v132, RDP (trainset 16, release 11.5), RefSeq + RDP, GTDB v86 and GTDB v89 using the files hosted at zenodo[61–63]. The taxonomy assigned using GTDB v89 was used for all further analysis.

The downstream analysis was done using Phyloseq package v1.30.0 in R v3.6.1 along with other packages like Microbiome v1.9.16, ggpubr v0.2.5, vegan v2.5–6[60,64–67]. In brief, all the data was imported into a Phyloseq object. Alpha diversity was calculated using "alpha" function from Microbiome package. Beta diversity was analysed by Non-Metric Dimensional Scaling (NMDS) on Bray–Curtis distance calculated and plotted using the functions within Phyloseq package. Adonis() function from vegan package and pairwise.adonis() function from pairwiseAdonis package v0.0.1 was used to compare Bray–Curtis distances among groups[68]. Phyloseq package was used to agglomerate taxonomy at phylum and genus levels. ggpubr package was used for comparing different groups. Core-microbiome was determined using the functions in Microbiome package. Differential abundance of taxa was estimated using DESeq2 package v1.26.0 in R along with functions from phyloseq package[69]. All the visualisations were prepared in R using ggplot2 package v3.2.1 along with other packages ggpubr v0.2.5, ggConvexHull v0.1.0, ggnewscale v0.4.1 and ggrepel v0.8.2[70–73]. Other R packages data.table v1.12.8, randomcoloR v1.1.0.1, tidyr v1.0.0, scales v1.1.0 and RColorBrewer v1.1–2 were also used in the analysis[74–78].

The R script used for analysis is available from github.com/ankit4035/peanutRhizosphere (https://doi.org/10.5281/zenodo.4699352) for reproduction of the entire work. The raw data files can be downloaded from the NCBI SRA (Accessions SRR12732102 to SRR12732191) under Bioproject PRJNA665712.

Ethics declaration. The study included the use of soil associated with plants. Since no parts of plants were used for the study, no permission was needed for ethical considerations. Furthermore, the owner/farmer was informed about the study and the type of samples to be collected. Verbal consent and permission were obtained to collect the soil from his farm for the work.

Conclusion

In all, we analysed structure of rhizosphere microbiota from pre-sowing of seeds to post-harvest of crops and compared them with consecutive bulk soil microbiota. Significant variations were observed in alpha-diversity and beta-diversity profiles between bulk and rhizosphere samples, while pre-sowing and post-harvest samples closely resembled bulk soil samples. Differences were also observed among several taxa across all rhizosphere samples especially PGP bacteria. While such patterns demonstrate the "rhizosphere effect", co-analysing this microbiome patterns with patterns of root exudates could reveal a bigger picture. However, due to lack of resources, we cannot analyse such information in the current study. Nonetheless, the results of this study provide a background for further studies on peanut microbiome.

Data availability

The R script used for analysis is available from github.com/ankit4035/peanutRhizosphere (https://doi.org/10.5281/zenodo.4699352) for reproduction of the entire work. The raw data files can be downloaded from the NCBI SRA (Accessions SRR12732102 to SRR12732191) under Bioproject PRJNA665712.

References

1. Mendes, R., Garbeva, P. & Raaijmakers, J. M. The rhizosphere microbiome: Significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *FEMS Microbiol. Rev.* **37**, 634–663. https://doi.org/10.1111/1574-6976.12028 (2013).
2. Bhattarai, A., Bhattarai, B. & Pandey, S. Variation of soil microbial population in different soil horizons. *J. Microbiol. Exp.* **2**, 00044. https://doi.org/10.15406/jmen.2015.02.00044 (2015).
3. Liu, F. *et al.* Soil indigenous microbiome and plant genotypes cooperatively modify soybean rhizosphere microbiome assembly. *BMC Microbiol.* **19**, 201. https://doi.org/10.1186/s12866-019-1572-x (2019).
4. Edwards, J. *et al.* Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E911-920. https://doi.org/10.1073/pnas.1414592112 (2015).
5. Vives-Peris, V., de Ollas, C., Gomez-Cadenas, A. & Perez-Clemente, R. M. Root exudates: From plant to rhizosphere and beyond. *Plant Cell Rep.* **39**, 3–17. https://doi.org/10.1007/s00299-019-02447-5 (2020).
6. Qu, Q. *et al.* Rhizosphere microbiome assembly and its impact on plant growth. *J. Agric. Food Chem.* **68**, 5024–5038. https://doi.org/10.1021/acs.jafc.0c00073 (2020).
7. Schmidt, J. E., Kent, A. D., Brisson, V. L. & Gaudin, A. C. M. Agricultural management and plant selection interactively affect rhizosphere microbial community structure and nitrogen cycling. *Microbiome* **7**, 146. https://doi.org/10.1186/s40168-019-0756-9 (2019).
8. Cordero, J., de Freitas, J. R. & Germida, J. J. Bacterial microbiome associated with the rhizosphere and root interior of crops in Saskatchewan, Canada. *Can. J. Microbiol.* **66**, 71–85. https://doi.org/10.1139/cjm-2019-0330 (2020).
9. Bulgarelli, D. *et al.* Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe* **17**, 392–403. https://doi.org/10.1016/j.chom.2015.01.011 (2015).
10. Lundberg, D. S. *et al.* Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488**, 86–90. https://doi.org/10.1038/nature11237 (2012).
11. Leoni, C. *et al.* Plant Health and Rhizosphere microbiome: Effects of the bionematicide *Aphanocladium album* in tomato plants infested by *Meloidogyne javanica*. *Microorganisms* https://doi.org/10.3390/microorganisms8121922 (2020).
12. Vitulo, N. *et al.* Bark and grape microbiome of vitis vinifera: Influence of geographic patterns and agronomic management on bacterial diversity. *Front. Microbiol.* **9**, 3203. https://doi.org/10.3389/fmicb.2018.03203 (2018).
13. Hu, J. *et al.* Rhizosphere microbiome functional diversity and pathogen invasion resistance build up during plant development. *Environ. Microbiol.* **22**, 5005–5018. https://doi.org/10.1111/1462-2920.15097 (2020).
14. Qiao, Q. *et al.* The variation in the rhizosphere microbiome of cotton with soil type, genotype and developmental stage. *Sci. Rep.* **7**, 3940. https://doi.org/10.1038/s41598-017-04213-7 (2017).
15. Baudoin, E., Benizri, E. & Guckert, A. Impact of growth stage on the bacterial community structure along maize roots, as determined by metabolic and genetic fingerprinting. *Appl. Soil. Ecol.* **19**, 135–145. https://doi.org/10.1016/S0929-1393(01)00185-8 (2002).
16. DeAngelis, K. M. *et al.* Selective progressive response of soil microbial community to wild oat roots. *ISME J.* **3**, 168–178. https://doi.org/10.1038/ismej.2008.103 (2009).
17. Ding, L. J. *et al.* Microbiomes inhabiting rice roots and rhizosphere. *FEMS Microbiol. Ecol.* https://doi.org/10.1093/femsec/fiz040 (2019).
18. Fan, K. *et al.* Rhizosphere-associated bacterial network structure and spatial distribution differ significantly from bulk soil in wheat crop fields. *Soil Biol. Biochem.* **113**, 275–284. https://doi.org/10.1016/j.soilbio.2017.06.020 (2017).
19. Jaiswal, S. K., Mohammed, M. & Dakora, F. D. Microbial community structure in the rhizosphere of the orphan legume Kersting's groundnut [Macrotyloma geocarpum (Harms) Marechal & Baudet]. *Mol. Biol. Rep.* **46**, 4471–4481. https://doi.org/10.1007/s11033-019-04902-8 (2019).
20. Kuramae, E. E. *et al.* Soil characteristics more strongly influence soil bacterial communities than land-use type. *FEMS Microbiol. Ecol.* **79**, 12–24. https://doi.org/10.1111/j.1574-6941.2011.01192.x (2012).
21. Lauber, C. L., Hamady, M., Knight, R. & Fierer, N. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl. Environ. Microbiol.* **75**, 5111–5120. https://doi.org/10.1128/AEM.00335-09 (2009).
22. Mendes, L. W., Kuramae, E. E., Navarrete, A. A., van Veen, J. A. & Tsai, S. M. Taxonomical and functional microbial community selection in soybean rhizosphere. *ISME J.* **8**, 1577–1587. https://doi.org/10.1038/ismej.2014.17 (2014).
23. Peiffer, J. A. *et al.* Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6548–6553. https://doi.org/10.1073/pnas.1302837110 (2013).
24. Perez-Jaramillo, J. E. *et al.* Deciphering rhizosphere microbiome assembly of wild and modern common bean (*Phaseolus vulgaris*) in native and agricultural soils from Colombia. *Microbiome* **7**, 114. https://doi.org/10.1186/s40168-019-0727-1 (2019).
25. Sugiyama, A., Ueda, Y., Zushi, T., Takase, H. & Yazaki, K. Changes in the bacterial community of soybean rhizospheres during growth in the field. *PLoS ONE* **9**, e100709. https://doi.org/10.1371/journal.pone.0100709 (2014).
26. Xu, J. *et al.* The structure and function of the global citrus rhizosphere microbiome. *Nat. Commun.* **9**, 4894. https://doi.org/10.1038/s41467-018-07343-2 (2018).
27. Haldar, S. & Sengupta, S. Impact of plant development on the rhizobacterial population of Arachis hypogaea: A multifactorial analysis. *J. Basic Microbiol.* **55**, 922–928. https://doi.org/10.1002/jobm.201400683 (2015).
28. Dai, L. *et al.* Effect of drought stress and developmental stages on microbial community structure and diversity in peanut rhizosphere soil. *Int. J. Mol. Sci.* https://doi.org/10.3390/ijms20092265 (2019).
29. Desmae, H. *et al.* Genetics, genomics and breeding of groundnut (*Arachis hypogaea* L.). *Plant Breed* **138**, 425–444. https://doi.org/10.1111/pbr.12645 (2019).
30. Pandey, M. K. *et al.* Translational genomics for achieving higher genetic gains in groundnut. *Theor. Appl. Genet.* **133**, 1679–1702. https://doi.org/10.1007/s00122-020-03592-2 (2020).
31. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583. https://doi.org/10.1038/nmeth.3869 (2016).
32. Parks, D. H. *et al.* A complete domain-to-species taxonomy for bacteria and archaea. *Nat. Biotechnol.* **38**, 1079–1086. https://doi.org/10.1038/s41587-020-0501-8 (2020).
33. Parks, D. H. *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996–1004. https://doi.org/10.1038/nbt.4229 (2018).
34. Lalucat, J., Mulet, M., Gomila, M. & Garcia-Valdes, E. Genomics in bacterial taxonomy: impact on the genus pseudomonas. *Genes* https://doi.org/10.3390/genes11020139 (2020).
35. Correa-Galeote, D., Bedmar, E. J., Fernandez-Gonzalez, A. J., Fernandez-Lopez, M. & Arone, G. J. Bacterial communities in the rhizosphere of *Amilaceous Maize* (*Zea mays* L.) as assessed by pyrosequencing. *Front. Plant Sci.* **7**, 1016. https://doi.org/10.3389/fpls.2016.01016 (2016).
36. Xu, Y. *et al.* Influence of salt stress on the rhizosphere soil bacterial community structure and growth performance of groundnut (*Arachis hypogaea* L.). *Int. Microbiol.* **23**, 453–465. https://doi.org/10.1007/s10123-020-00118-0 (2020).
37. Hamonts, K. *et al.* Field study reveals core plant microbiota and relative importance of their drivers. *Environ. Microbiol.* **20**, 124–140. https://doi.org/10.1111/1462-2920.14031 (2018).

38. Ansari, F. A. & Ahmad, I. Isolation, functional characterization and efficacy of biofilm-forming rhizobacteria under abiotic stress conditions. *Antonie Van Leeuwenhoek* **112**, 1827–1839. https://doi.org/10.1007/s10482-019-01306-3 (2019).
39. Singh, T. B. *et al.* Identification, characterization and evaluation of multifaceted traits of plant growth promoting rhizobacteria from soil for sustainable approach to agriculture. *Curr. Microbiol.* **77**, 3633–3642. https://doi.org/10.1007/s00284-020-02165-2 (2020).
40. Govindasamy, V. *et al.* Multi-trait PGP rhizobacterial endophytes alleviate drought stress in a senescent genotype of sorghum [Sorghum bicolor (L.) Moench]. *3 Biotech* **10**, 13. https://doi.org/10.1007/s13205-019-2001-4 (2020).
41. Abedinzadeh, M., Etesami, H. & Alikhani, H. A. Characterization of rhizosphere and endophytic bacteria from roots of maize (*Zea mays* L.) plant irrigated with wastewater with biotechnological potential in agriculture. *Biotechnol. Rep.* **21**, e00305. https://doi.org/10.1016/j.btre.2019.e00305 (2019).
42. Hashem, A., Tabassum, B. & Allah, F. A. Bacillus subtilis: A plant-growth promoting rhizobacterium that also impacts biotic stress. *Saudi J. Biol. Sci.* **26**, 1291–1297. https://doi.org/10.1016/j.sjbs.2019.05.004 (2019).
43. Parks, D. H. *et al.* Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* **2**, 1533–1542. https://doi.org/10.1038/s41564-017-0012-7 (2017).
44. Gomez-Lama Cabanas, C. *et al.* Indigenous *Pseudomonas* spp. strains from the olive (*Olea europaea* L.) rhizosphere as effective biocontrol agents against Verticillium dahliae: From the host roots to the bacterial genomes. *Front. Microbiol.* **9**, 277. https://doi.org/10.3389/fmicb.2018.00277 (2018).
45. Ansari, F. A. & Ahmad, I. Fluorescent pseudomonas-FAP2 and *Bacillus licheniformis* interact positively in biofilm mode enhancing plant growth and photosynthetic attributes. *Sci. Rep.* **9**, 4547. https://doi.org/10.1038/s41598-019-40864-4 (2019).
46. Pandey, K. K., Mayilraj, S. & Chakrabarti, T. *Pseudomonas indica* sp. nov., a novel butane-utilizing species. *Int. J. Syst. Evol. Microbiol.* **52**, 1559–1567. https://doi.org/10.1099/00207713-52-5-1559 (2002).
47. Shade, A., Jacques, M. A. & Barret, M. Ecological patterns of seed microbiome diversity, transmission, and assembly. *Curr. Opin. Microbiol.* **37**, 15–22. https://doi.org/10.1016/j.mib.2017.03.010 (2017).
48. Adam, E., Bernhart, M., Müller, H., Winkler, J. & Berg, G. The Cucurbita pepo seed microbiome: Genotype-specific composition and implications for breeding. *Plant Soil* **422**, 35–49. https://doi.org/10.1007/s11104-016-3113-9 (2018).
49. Truyens, S., Weyens, N., Cuypers, A. & Vangronsveld, J. Bacterial seed endophytes: Genera, vertical transmission and interaction with plants. *Environ. Microbiol. Rep.* **7**, 40–50. https://doi.org/10.1111/1758-2229.12181 (2015).
50. Kong, H. G., Song, G. C. & Ryu, C.-M. Inheritance of seed and rhizosphere microbial communities through plant–soil feedback and soil memory. *Environ. Microbiol. Rep.* **11**, 479–486. https://doi.org/10.1111/1758-2229.12760 (2019).
51. Frindte, K., Pape, R., Werner, K., Loffler, J. & Knief, C. Temperature and soil moisture control microbial community composition in an arctic-alpine ecosystem along elevational and micro-topographic gradients. *ISME J.* **13**, 2031–2043. https://doi.org/10.1038/s41396-019-0409-9 (2019).
52. Cook, R. J. *et al.* Molecular mechanisms of defense by rhizobacteria against root disease. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 4197–4201. https://doi.org/10.1073/pnas.92.10.4197 (1995).
53. Chaparro, J. M., Badri, D. V. & Vivanco, J. M. Rhizosphere microbiome assemblage is affected by plant development. *ISME J.* **8**, 790–803. https://doi.org/10.1038/ismej.2013.196 (2014).
54. Sharma, S. B., Sayyed, R. Z., Trivedi, M. H. & Gobi, T. A. Phosphate solubilizing microbes: Sustainable approach for managing phosphorus deficiency in agricultural soils. *Springerplus* **2**, 587. https://doi.org/10.1186/2193-1801-2-587 (2013).
55. Kumar, A., Prakash, A. & Johri, B. N. In *Bacteria in Agrobiology: Crop Ecosystems* (ed. Maheshwari, D. K.) 37–59 (Springer, 2011).
56. Sachdev, D., Nema, P., Dhakephalkar, P., Zinjarde, S. & Chopade, B. Assessment of 16S rRNA gene-based phylogenetic diversity and promising plant growth-promoting traits of Acinetobacter community from the rhizosphere of wheat. *Microbiol. Res.* **165**, 627–638. https://doi.org/10.1016/j.micres.2009.12.002 (2010).
57. Lareen, A., Burton, F. & Schafer, P. Plant root-microbe communication in shaping root microbiomes. *Plant Mol. Biol.* **90**, 575–587. https://doi.org/10.1007/s11103-015-0417-8 (2016).
58. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, e1. https://doi.org/10.1093/nar/gks808 (2013).
59. Callahan, B. J., Sankaran, K., Fukuyama, J. A., McMurdie, P. J. & Holmes, S. P. Bioconductor workflow for microbiome data analysis: From raw reads to community analyses. *F1000Res* **5**, 1492. https://doi.org/10.12688/f1000research.8986.2 (2016).
60. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/ (2019).
61. Alishum, A. DADA2 formatted 16S rRNA gene sequences for both bacteria & archaea. doi: 10.5281/zenodo.2541239 (2019).
62. Callahan, B. Silva taxonomic training data formatted for DADA2 (Silva version 132). doi: 10.5281/zenodo.1172783 (2018).
63. Callahan, B. RDP taxonomic training data formatted for DADA2 (RDP trainset 16/release 11.5). doi: 10.5281/zenodo.801828 (2017).
64. McMurdie, P. J. & Holmes, S. 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 (2013).
65. Kassambara, A. *ggpubr: 'ggplot2' Based Publication Ready Plots*. R package version 0.2.5. https://CRAN.R-project.org/package=ggpubr (2020).
66. Lahti, L. & Shetty, S. *Tools for microbiome analysis in R Version 2.1.26.* http://microbiome.github.com/microbiome (2017).
67. Oksanen, J. *et al.* vegan: *Community Ecology Package*. R package version 2.5–6. https://CRAN.R-project.org/package=vegan (2019).
68. Martinez Arbizu, P. *pairwiseAdonis: Pairwise Multilevel Comparison using Adonis*. R package version 0.0.1. (2017).
69. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. https://doi.org/10.1186/s13059-014-0550-8 (2014).
70. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
71. Martin, C. *ggConvexHull: Add a convex hull geom to ggplot2*. R package version 0.1.0. http://github.com/cmartin/ggConvexHull (2017).
72. Campitelli, E. *ggnewscale: Multiple Fill and Colour Scales in 'ggplot2'*. R package version 0.4.1. https://CRAN.R-project.org/package=ggnewscale (2020).
73. Slowikowski, K. *ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'*. R package version 0.8.2. https://CRAN.R-project.org/package=ggrepel (2020).
74. Dowle, M. & Srinivasan, A. *data.table: Extension of `data.frame`*. R package version 1.12.8. https://CRAN.R-project.org/package=data.table (2019).
75. Ammar, R. *randomcoloR: Generate Attractive Random Colors*. R package version 1.1.0.1. https://CRAN.R-project.org/package=randomcoloR (2019).
76. Wickham, H. & Henry, L. *tidyr: Tidy Messy Data*. R package version 1.0.0. https://CRAN.R-project.org/package=tidyr (2019).
77. Wichmann, H. & Seidel, D. *scales: Scale Functions for Visualization*. R package version 1.1.0. https://CRAN.R-project.org/package=scales (2019).
78. Neuwirth, E. *RColorBrewer: ColorBrewer Palettes*. R package version 1.1–2. https://CRAN.R-project.org/package=RColorBrewer (2014).

Acknowledgements

Author contributions

A.T.H. was involved in sample collection. A.T.H. and K.J.P. was involved in extraction of nucleic acids and its sequencing. A.T.H. did the data analysis and manuscript writing with inputs from R.J.P. A.T.H., P.G.K. and R.K.K. conceptualised the study and corrected the manuscript. All authors read and approved the final manuscript.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-97071-3.

Correspondence and requests for materials should be addressed to R.K.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.