Poor competitiveness of *Bradyrhizobium* in pigeon pea root colonisation in Indian soils

Danteswari Chalasani[1†], Anirban Basu[1†], Sarma V.S.R.N. Pullabhotla[1†], Beatriz Jorrin[2], Andrew L. Neal[3], Philip S. Poole[2], Appa Rao Podile[1]* and Andrzej Tkacz[2]*

[1]Department of Plant Sciences, School of Life Sciences, University of Hyderabad, Hyderabad, 500 046, Telangana, India
[2]Department of Plant Sciences, University of Oxford, Oxford, OX1 3RB, UK
[3]Department of Sustainable Agriculture Sciences, Rothamsted Research, North Wyke, EX20 2SB, UK

*Corresponding authors
Dr. Andrzej Tkacz: andrzej.tkacz@plants.ox.ac.uk
Prof. Appa Rao Podile: podilerao@gmail.com, arpsl@uohyd.ernet.in

[†]These authors contributed equally to this work

Abstract:

Background

Pigeon pea, a legume crop native to India, is the primary source of protein for more than a billion people in developing countries. The plant can form symbioses with N_2-fixing bacteria, however reports of poor crop nodulation in agricultural soils abound. We report here study of the microbiota associated with pigeon pea, with a special focus on the symbiont population in different soils and vegetative and non-vegetative plant growth.

Results

Location with respect to the plant roots was determined to be the main factor controlling the microbiota followed by developmental stage and soil type. Plant genotype plays only a minor role. Pigeon pea roots have a reduced microbial diversity compared to the surrounding soil and select for Proteobacteria and especially for *Rhizobium* spp. during vegetative growth. While *Bradyrhizobium*, a native symbiont of pigeon pea, can be found associating with roots, its presence is dependent on plant variety and soil conditions. A combination of metagenomic survey, strain isolation and co-inoculation with nodule forming *Bradyrhizobium* spp. and non-N_2 fixing *Rhizobium* spp. demonstrated that the latter is a much more successful coloniser of pigeon pea roots.

Conclusions

Poor nodulation of pigeon pea in Indian soils may be caused by a poor *Bradyrhizobium* competitiveness against non-nodulating root colonisers such as *Rhizobium*. Hence, inoculant strain selection of symbionts for pigeon pea should not only be based on their nitrogen fixation potential but more importantly on their competitiveness in agricultural soils.

Introduction:

Pigeon pea [*Cajanus cajan* (L.) Millspaugh] is one of the most important legume crops with diverse uses as food, feed, fodder and fuel, besides enriching soil through biological nitrogen fixation. Globally, the crop is grown on about 7 million hectares (FAOSTAT 2019), mainly as a rain-fed crop in semi-arid tropical and subtropical regions of South Asia, East Africa, Latin America and the Caribbean. It is the primary source of dietary protein for over a billion people in developing countries. Millions of resource-poor smallholder farmers grow this multipurpose crop with minimal inputs to sustain their livelihoods. Domestication of the wild progenitor species *Cajanus cajanifolius* (endemic to the Indian subcontinent) resulted in the origin of the cultivated pigeon pea in central India more than 3,500 years ago, from where it subsequently spread to other parts of the globe (Fuller et al. 2019, Varshney et al. 2017). India is the largest producer of pigeon pea, accounting for 72% of global production (FAOSTAT 2019). It is the second-largest cultivated legume crop (after chickpea) in India, contributing 15% by area and 17% by production (Tiwari and Shivhare 2018). The major pigeon pea growing zones in India can be divided into the south zone (Andhra Pradesh,

Telangana and Karnataka), central zone (Madhya Pradesh, Maharashtra and Gujarat) and northern plain zone (Uttar Pradesh) (Singh 2013). The states of Andhra Pradesh, Madhya Pradesh and Uttar Pradesh record the highest yields (Tiwari and Shivhare 2016). The soil types in these three states located in the south, central and northern zones are red soil (Alfisol), black soil (Vertisol) and alluvial soil (Inceptisol), respectively, based on the United States Department of Agriculture (USDA) Soil Taxonomy (Table S1). A large number of pigeon pea varieties are cultivated in India exhibiting a vast genetic and phenotypic diversity of agro-morphological traits, including variations in plant type, branching pattern, pod and seed size, seed colour, protein content, grain yield, resistance/tolerance to abiotic and biotic stresses, crop duration, photoperiod sensitivity, days to flowering and maturity (Semwal et al. 2018).

The root-associated microbiomes of many plants and crop species have been extensively studied, including those associating with other N_2-fixing legumes such as soybean, alfalfa, red clover, common bean and *Lotus japonicum* (Hartman et al. 2017, Liu et al. 2019, Mendes et al. 2014, Mendes et al. 2018, Perez-Jaramillo et al. 2017, Xiao et al. 2017, Zgadzaj et al. 2016). However, the microbiome of tropical grain legumes such as pigeon pea have not been described yet. The root microbiomes of legumes differ from those of non-legumes owing to the symbiotic association with diverse rhizobia in the root nodules. The legume hosts exert a strong influence on the rhizobial diversity patterns in the soil and different parts of the root microbiome (Miranda-Sanchez et al. 2016).

Studies based on strain isolation from India suggest that pigeon pea can be nodulated by *Rhizobium* spp. (Rajendran et al. 2008, Singh et al. 2018, Singha et al. 2018), *Bradyrhizobium* spp. (Nautiyal et al. 1988), *Sinorhizobium/Ensifer* (Dubey et al. 2010, Gosai et al. 2020, Jain et al. 2020), *Mesorhizobium* (Jain et al. 2020, Singha et al. 2018) or even *Burkholderia* (Singha et al. 2018). However, in other geographical regions including Cote d'Ivoire (Fossou et al., 2016), Ethiopia (Degefu et al., 2018), Dominican Republic (Araujo et al., 2015) and Brazil (Rufini et al., 2016), pigeon pea is nodulated solely either by *Bradyrhizobium* spp., or *Sinorhizobium*/*Ensifer* (López-Baena et al. 2016), suggesting that any other species found in the nodulation studies may need to be re-evaluated.

Apart from rhizobial symbionts, pigeon pea harbours diverse non-rhizobial root colonisers belonging to *Bacillus*, *Brevibacillus*, *Paenibacillus*, *Lactobacillus*, *Pseudomonas*, *Agrobacterium*, *Enterobacter*, *Klebsiella*, *Chryseobacterium*, *Streptomyces*, *Serratia*, and other genera (Bind and Nema 2019, Dubey et al. 2010, Fossou et al. 2016, Rajendran et al. 2008).

Most studies have concentrated on isolation and characterisation of pigeon pea nodule and root bacteria and their use in inoculation assays to promote plant growth. No comprehensive study on the root-associated microbiome of pigeon pea or any other legumes grown in the Indian soil has been undertaken to date, neither has any throughput screening of common pigeon pea symbionts been performed. Genomic tools and high throughput sequencing technologies allowed characterisation of the genetic and genomic diversity of pigeon pea, including whole-genome sequencing (Varshney et al. 2017), although this did not cover the root microbiome. The present study was designed to a) identify microbial taxa associated

with pigeon pea roots and surrounding soil, b) investigate the factors shaping the pigeon pea root-associated microbiome in various Indian soils and c) identify nodule symbionts.
Pigeon pea is a legume, and is able to obtain nitrogen through symbiosis. However, its growth is often supplemented with inorganic and organic fertilisers as there are reports of weak nodulation in some parts of India (Arora et al. 2018). Some varieties of pigeon pea have a low symbiosis potential, and it is possible to obtain nodulation deficient lines by simply crossing less efficient lines between each other (Rupela and Johansen 1995).
We hypothesise that the reported low nodulation efficiency of pigeon pea is an outcome of either the low number of compatible symbionts in the soils and/or their low competitiveness in colonising the host plant.

To capture the representative microbiome of pigeon pea, we have sampled different parts of plant microbiome (root endosphere, rhizoplane, rhizosphere and soil not attached to the roots – loose soil, representing only a very weak plant influence). We did not separate nodules from the roots and sequence them, but rather collected them in a separate experiment for isolation study. Based on our previous experience (Tkacz et al. 2020) we know that nodule microbiota structure, obtained through next generation sequencing is likely to include data for bacteria attached to the outside of the nodules and/ those able to (re)colonise the nodule especially in its later developmental stages.
To separate the effect of actively growing roots secreting photosynthetic products into the surrounding soil from plant generally weaker plant secretions (Chaparro et al. 2014) and potentially caused only as an attachment point for microbes (Bulgarelli et al. 2012) we have sampled plants at two contrasting developmental stages; vegetative and flowering.
Widely grown pigeon pea cultivars Asha (ICPL 87119), Durga (ICPL 84031) and Mannem Konda Kandi (MKK; ICPH 2740) were grown in three soils collected across the Indian subcontinent, representing three major geological substrates for soil formation: Alfisols – red soils originating from highly weathered rocks, typical tropical climate process of leaching most soil minerals accumulating insoluble aluminium and iron, Inceptisols – alluvial soils originating from recent flood deposit and Vertisols – black soils originating from older fluvial deposits and representing a typical clay-rich tropical agriculture soil (Murthy 1988).
We compared the Indian pigeon pea data with our previously characterised legume and non-legume plant microbiota of British soils. The data is fully comparable as the same methods were applied for all the samples.
We show in the metagenomic screen, bacterial isolation assay and a gnotobiotic experiment that pigeon pea roots are predominantly colonised by non-symbiotic *Rhizobium* spp., rather than symbiotic *Bradyrhizobium* spp.

Methodology

Experimental design

We collected three different soil types from farmers' fields in the principal pigeon pea producing regions of India during the pre-sowing season between June and August 2017. Alfisols were collected from Andhra Pradesh (Rompicharla of Guntur district, 16.213900N,

79.921386 E), Vertisols from Madhya Pradesh (Athner of Betul district, 21.6406552 N, 77.91300 E), and Inceptisols from Uttar Pradesh (Sitamarhi of Allahabad district, 25.2782289 N, 82.28691 E) (Fig. S1). Rompicharla has a tropical climate with an average annual temperature of 28.5 °C (24.1-33.6 °C) and average annual precipitation (rainfall) of 906 mm. Athner also has a tropical climate with an average annual temperature of 24.6 °C (19.1-32.4 °C) and average annual precipitation (rainfall) of 943 mm. Sitamarhi has a subtropical climate with an average annual temperature of 25.7 °C (16.1-34.2 °C) and average annual precipitation (rainfall) of 981 mm (Fig. S2).

Physicochemical characterisation of collected soils (Table S1) was performed using HiMedia Soil Testing Kits (HiMedia Laboratories, Mumbai, India) according to the manufacturer's instructions. Three popular pigeon pea cultivars (genotypes) were selected for this study, *viz.* Asha (ICPL-87119), Durga (ICPL-84031) and Mannem Konda Kandi (MKK; ICPH-2740) (Table S1). The seeds were procured from International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India. Seeds were surface sterilised using $HgCl_2$ (0.1%) and ethanol (70%) and germinated on MS agar. Three seedlings of each genotype were transplanted into pots (pot size=7.5 kg) individually filled with three collected soil samples. The plants were grown using six biological replicates in a glasshouse at University of Hyderabad, Hyderabad, India under identical conditions of light, temperature and humidity until flowering stage. Six pots of soil for each soil type (without growing any plant) were used as bulk soil control. Plants, as well as control pots, were watered as needed with sterilised distilled water every alternate day without adding any other fertilisers.

Sampling of soil and root fractions:

Plants were harvested at two developmental stages; vegetative (1 month after seedling emergence) and flowering stages (3 months after seedling emergence for Durga; 4 months after seedling emergence for Asha and MKK). Uprooted plant roots were briefly shaken to remove loosely attached soil, collected as 'loose soil' fraction. The soil bound tightly to the roots was collected without damaging the root and root nodules by vortexing and centrifugation at 4,000 rpm for 10 min to yield the 'rhizosphere' fraction. After removing the rhizosphere soil, roots were washed and transferred to 15 mL falcon tube (with 10 mL sterile water) and sonicated for 5 min at full intensity in an ultrasonic bath. Roots were removed, and the falcon tube was centrifuged at 4,000 rpm for 10 min to collect 'rhizoplane' fraction as a pellet. Washed and sonicated roots were ground to a powder using liquid nitrogen and defined as the 'endosphere' fraction.

DNA extraction, PCR and sequencing:

Metagenomic DNA was extracted from the bulk and loose soil, rhizosphere, rhizoplane and root endosphere samples (0.3 g each) using NucleoSpin® Soil Kit (Machery Nagel, Germany) according to the manufacturer's instructions. The V4 hypervariable region of the bacterial 16S rRNA gene was amplified using double-barcoded 515F/806R primer pairs (Caporaso et al. 2011). The PCR reaction mixture consisted of Phusion high fidelity (0.2 μL),

high fidelity buffer (4 µL) (Thermo Scientific, Waltham, USA), dinucleotide triphosphates (0.4 µL), primers (1 µL of each 10 pmol stock), template DNA (1.5 µL of 5 ng μL^{-1}) and H_2O up to a 20 µL final volume. For rhizoplane and endosphere fractions, peptide nucleic acid (PNA) for targeting plastid (pPNA, 5′-GGCTCAACCCTGGACAG-3′) and mitochondrial (mPNA, 5′-GGCAAGTGTTCTTCGGA-3′) DNA (PNA Bio, Newbury Park, USA) of 1 µM as PCR clamps (Lundberg et al. 2013) were added. PCR cycles were as follows: 98 °C for 1 min, 35 cycles of 98 °C for 30 s, 57 °C for 30 s and 72 °C for 45 s with a final elongation step of 72 °C for 7 min. Each DNA sample was amplified in triplicate, followed by purification using a PCR clean-up kit (D4014, Zymo Research). For each amplification run in 96-well plates, water was used as a negative control (no-DNA control). Samples were pooled and sequenced using Illumina MiSeq platform using v3 chemistry of 300PE run at Molecular Research DNA laboratory in Texas, USA.

Processing of sequencing data: Initial quality filter and reads alignment was done using Usearch 10 fastq_mergepairs with fastq_maxee using an EE score of 1. After barcode removal only reads of the desired length of 292 bp were used for further analysis. Reads were filtered from plant chloroplast and mitochondria (around 2% of the initial reads were plant origin) using custom-made Bash script (Table S1). Reads were binned into zero-radius Operational Taxonomic Units (zOTUs), including chimera removal according to the Usearch10 pipeline with Unoise3 (Edgar 2016) and annotated using SILVA SSU132 16S rRNA database (Quast et al. 2013).

Statistical analyses:

Permutational multivariate analysis of variance (PERMANOVA), unconstrained Principal Coordinates Analysis (PCoA) and analysis of similarity (ANOSIM) were based on Bray-Curtis dissimilarity matrices calculated from standardised, square-root transformed abundance data and calculated and/or visualised in Primer 6 software (PRIMER-E, Quest Research Ltd., Auckland, New Zealand). Factors influencing the microbial community were statistically assessed using permutation of residuals under a reduced model, sum of squares type III (partial) with 9,999 permutations using unrestricted permutation of raw data model of PERMANOVA. We considered *pseudo-F* values as proxies of a given factor's importance for sample separation and is based on the ratio of the beta-diversity (variation between two or more sample groups) to alpha-diversity (variation between individual samples inside each of these groups). PCoA plots are designed to visualise distance matrices with maximum sample separation along multiple axes (however, for clarity only the first two axes are shown) without prior factorial description. One-way ANOSIM tests based upon 9,999 permutations were used to calculate the difference (the ratio of beta- to alpha-diversity) between each set of data for a given factor. Similarity percentage (SIMPER) was also run in PRIMER 6 on standardised and square-rooted transformed abundance data and aimed to identify microbial taxa with the greatest influence on sample groups separation.

Shannon entropy plot, Volcano plot and taxonomic bar plots were visualised in PRISM 8 (GraphPad, San Diego, USA). Shannon entropy was calculated for each sample as (-1) * sum

of (each zOTU value * ln of each zOTU value), where the sum of zOTU values for each sample equals to 1. For Volcano and taxonomic plots, the taxonomic affiliations of zOTUs were summed into phyla, families and genera. Volcano plots present genera location on XY matrix as a result of their fold change against bulk soil (X-axis) and statistical significance of this change (Y-axis) corrected for multiple testing with false rate discovery (FDR).

Isolation of bacteria from the rhizosphere, roots and nodules

Pigeon pea plants (three cultivars each in three soil types) were harvested at the vegetative stage for isolation of bacteria. Harvested nodules were surface sterilised with 0.1% $HgCl_2$, crushed and serially diluted in saline (0.86% NaCl), streaked onto CRYEM (Congo red yeast extract mannitol) plates, and incubated at 25 °C for up to 7 days. Colonies were selected from CRYEM plates and streaked to new plates to obtain pure cultures. Rhizosphere and root samples were collected as described above, diluted in saline, streaked onto yeast extract mannitol (YEM) plates and incubated at 30 °C for up to 4 days. Single colonies were purified further, their BOX-PCR obtained using the BOX-A1R primer 5'-CTACGGCAAGGCGACGCTGACG-3' (Versalovic et al. 1994) and their 16S rRNA gene sequenced after PCR amplification using 27F (5'-GTTTGATCCTGGCTCAG-3') and 1494R (5'-ACGGCTACCTTGTTACGACTT-3') primers.

WGS of pigeon pea isolates

Eighteen pigeon pea isolates were selected based on their different BOX-PCR pattern. Culture samples were provided to Microbes NG for Illumina sequencing (MiSeq v2, PE 2x250 bp). The closest available reference genome for each sample was identified with Kraken v2 (Wood and Salzberg 2014), and reads were mapped to this using bwa-mem v07.17 (Li and Durbin 2009) to assess the quality of the data. *De novo* assembly was performed with SPAdes v 3.14.1 (Bankevich et al. 2012). An automated annotation was performed using Prokka v1.12 (Seemann 2014). Strains were annotated using their whole genomes with EzBioCloud to the species level.

A local BLAST database of these genomes was generated in Geneious R10 (v10.2.6), and *nodC* and *nifH* sequences from *Bradyrhizobium cajani* AMNPC 1010^T (PRJNA593773) and *Ensifer fredii* NBRC 14780^T (PRJDB6002) were used to assess the presence of these genes in bacteria belonging *Burkholderia*/*Paraburkholderia* spp. (4), *Microbacterium* spp. (3) and *Rhizobium*/*Agrobacterium* spp. (11).

Gnotobiotic inoculation assay

Seeds were surfaced sterilised using ethanol (70%) for 1 min and bleach (4%) for 3 min and placed on water agar until the seedling emergence. Plants were moved to pots (1 L) with vermiculate with N-free rooting solution (400 mL) (Allaway et al. 2000) in a controlled growth chamber and inoculated either with a single *Rhizobium* or *Bradyrhizobium* strain, a community of *Rhizobium* spp. or a community of *Bradyrhizobium* spp. or double inoculated with both the *Rhizobium* spp. and *Bradyrhizobium* spp. communities. *Rhizobium* strains were

isolated from the roots, while *Bradyrhizobium* strains come from pigeon pea nodules. *Rhizobium* strains come from this work, while *Bradyrhizobium* strains were isolated from pigeon pea nodules grown in various Indian soils. Characterisation of *Bradyrhizobium* strains will be covered in a separate publication.

Four-week old pigeon pea plants were harvested, their roots (rhizoplane and endosphere combined) and nodules (if any) crushed using pestle and mortar, the crushed nodule macerate plated in dilution series on AG agar plates and left for 3 days to allow the growth of both *Rhizobium* and *Bradyrhizobium* spp. DNA from individual colonies from highly diluted treatments was isolated and used for ribosomal intergenic spacer (RISA) fingerprinting for species identification using RISA primers; ITSF: 5′-GTCGTAACAAGGTAGCCGTA-3′ and ITSReub: 5′-GCCAAGGCATCCACC-3′ and PCR conditions of 95 °C for 7 min, 30 cycles of 95 °C for 30 sec, 55 °C for 30 sec, 72 °C for 1 min and final elongation of 72 °C for 7 min). All *Bradyrhizobium* strains used have a RISA band of ~900 bp, while all *Rhizobium* strains have the RISA band of approximately 1200 bp allowing for a quick species identification.

Results:

Fraction, developmental stage, soil and genotype shape the pigeon pea microbiome

To understand what is shaping pigeon pea microbiome we used multi-factor PERMANOVA of the following factors: plant fraction (root endosphere, rhizoplane, rhizosphere and loose soil), soil type (Alfisol, Inceptisol and Vertisol), plant genotype (Asha, MKK and Durga) and plant developmental stage (vegetative and flowering) as factors.

We found that the main factor controlling the assembly of pigeon pea microbiome is the plant fraction followed by developmental stage, soil type and the least important, yet still a significant factor is the plant genotype (Fig. 1A and Fig. S3). However, when we look at each fraction separately, the soil is more important than the developmental stage for loose and attached soil, while the developmental stage is the dominant factor for rhizoplane and endosphere (Fig. S3B). Comparing each soil separately, fraction is more important than developmental stage for Alfisol, and they are of approximately the same importance for Inceptisol while developmental stage is more important for Vertisol (Fig. S3C). All factors exert a similar influence for individual plant genotypes (Fig. S3D). Analysing the data with the separation for vegetative and flowering stages, we uncovered that while fraction and plant genotype are of similar importance for both these stages, soil type factor is more important for the flowering plants (Fig. S3E).

We have also compared the major factors of pigeon pea microbiota assembly with our previous findings using legume and non-legume plants grown in the United Kingdom (Tkacz et al. 2020). Even the strongest factor as fraction is dwarfed by the importance of the sample origin (India vs. UK). Some of this difference can be explained by the plant species influence. However, distantly related plants grown in United Kingdom (pea, *Medicago,* wheat and *Arabidopsis*) have a relatively similar microbiota comparing to the Indian-grown pigeon pea suggesting that plant species effect is small.

At the zOTU-level, the origin is approximately ten-fold times more influential than fraction, while this ratio decreases with increasing taxonomic level (6x for genus, 5x for family and 3x for phyla). This change is probably caused by a reduction of alpha diversity with an increase in taxonomic level.

Visualising the microbiota community using PCoA we confirm PERMANOVA results where fraction (Fig. 1B and Fig. S4A) then developmental stage (Fig. S4D), soil (Fig. S4B) and plant genotype (Fig. S4C) can shape the pigeon pea microbiota. PCoA plots demonstrate that plants of all three genotypes shift their microbiota between vegetative and flowering stages, an effect especially observed for Vertisol grown plants (Fig. S4B and D). This observation is a plausible mechanism behind the increase of soil type factor strength for flowering plants already reported using PERMANOVA (Fig. S3E).

ANOSIM and PCoA plots were used to assess the differences between samples based on specific factors. For the fraction factor, the major community shift happens between rhizosphere and rhizoplane (ANOSIM's R=0.301, p <0.01), followed by rhizoplane and endosphere (R=0.269, p <0.01), while there is less difference between loose soil and rhizosphere (R=0.176, p <0.01) or bulk soil and loose soil (R=0.096, not significant). PCoA plot clearly illustrates it with a gentle sample location shift between all fractions, while the main boundary can be drawn between bulk soil-loose soil-rhizosphere cluster and rhizoplane-endosphere cluster (Fig. S4A).
For the soil type, ANOSIM separates plants grown in Vertisol from ones grown in Alfisol (R=0.289, p <0.01) and Inceptisol (R=0.287, p <0.01), while the difference between Alfisol and Inceptisol grown plants is weaker, yet significant (R=0.199, p <0.01) (Fig. S4B). Plant genotype as a factor is a significant factor shaping the microbiota (based on PERMANOVA – Fig. 1A) when the influence of all the three genotypes is considered. However, ANOSIM reveals that no genotype-to-genotype comparison is significant (Asha-Durga, R=0.028, Asha-MKK R=0.018 and Durga-MKK R=0.018, p >0.05 for all). PCoA plot visualises the lack of separation for samples based on their plant genotype factor (Fig. S4C). ANOSIM confirms PERMANOVA findings that developmental stage is one of the strongest factors (R=0.211, p <0.01), while PCoA plot clearly separates samples based on their developmental stage and as expected, loose soil is not affected by the plant developmental stage (a cluster of red-labelled points in the top right corner of Fig. S4D).

To confirm the observed PCoA sample-spread pattern and the main factors driving the community assembly we re-analysed the data originally based on zOTU assignment (sequencing reads similarity) at three higher taxonomical levels of genera, family and phyla (Fig. S5). Irrespective of the taxonomic level, the fraction, followed by developmental stage, soil type and genotype are the main factors controlling the community structure. To no surprise, the higher taxonomic levels due to the reduced number of categories have lower alpha diversity leading to a better separation of different sets of samples. This effect can be observed with an increased PCoA axis (PCO 1 and PCO 2) contribution in explaining the data variation (i.e., PCO 1 axis for genus level explains 18.9, for family 24% and for phyla

43.9%). Moreover, PERMANOVA pseudo-F value also increases for the higher taxonomic level data separation (apart from soil type at the phylum level).

To analyse the Indian pigeon pea microbiome in a wider context, we have supplemented the data with our previous legume and non-legume soil and root assay of plants grown in the United Kingdom (Tkacz et al. 2020). For consistency, we analysed the data at four taxonomic levels: zOTU, genus, family and phyla. However, irrespectively of the taxonomic level used, we see that PCoA plots clearly separate the Indian from United Kingdom samples on their first axis (PCO1), while the fractions within each origin group are separated on the second axis (PCO2) (Fig. 1C and Fig. S6A-D). The fraction separation pattern is similar for both Indian and United Kingdom samples. Root samples (endosphere and rhizoplane) in both cases are separated from the soil fractions (loose soil and rhizosphere) and with a higher taxonomic level root microbiota becomes similar across both geographical locations, irrespective of the plant species or genotype origin.

This community convergence was analysed further with PCA, and both United Kingdom and Indian root communities are highly influenced by Alphaproteobacteria and Bacteroidetes (Fig. S6E). As we analyse legume plants, we present the data with *in-silico* removal of potential symbionts, such as *Rhizobium* and *Bradyrhizobium.* However, such removal does not change the main PCA-based sample separation and overall pattern (Fig. S6F). Pigeon pea roots at the flowering stage (with or without potential symbionts removed) when the rhizodeposition may be reduced have enriched their microbial community to Gammaproteobacteria and Actinobacteria. As we have sampled only vegetative stage plants in United Kingdom samples, we cannot confirm here whether this process is uniform or specific only to the Indian pigeon pea samples.

There is a significant reduction of alpha diversity expressed as Shannon entropy associated with the rhizoplane and endosphere of pigeon pea, irrespective of their developmental stage. However, alpha diversity is higher during the flowering rather than vegetative stage. There is no consistent soil or genotype influence on alpha diversity (Fig. 1D and Fig. S7)

<u>Pigeon pea roots are colonised by Proteobacteria and Bacteroidetes</u>

Taxonomic profiles of the rhizoplane and root endosphere are different from that of the loose soil and rhizosphere. Root fractions are colonised by Alpha-, Beta- and Gammaproteobacteria, as well as Bacteroidetes, especially during vegetative growth. The Proteobacteria replace Acido-, Actinobacteria and Archaea found in soil (Fig. 2A and Fig. S8).

Comamonadaceae, Sphingomonadaceae and Xanthomonadaceae abundance increases in the rhizosphere and root fractions while Rhizobiaceae are more prevalent in roots of vegetative plants. The separation of soil and root fractions is clearer during vegetative growth than at the flowering stage (Fig. 2B and Fig. S8).

The main genera in the roots of vegetative stage plants are *Rhizobium*, *Pseudoxanthomonas* and *Sphingopyxis*. Some plant roots also have a high abundance of *Bradyrhizobium*

suggesting an active endosymbiosis. Vegetative plants allow *Sphingopyxis* root colonisation, while being replaced by *Brachymonas* in the flowering stage (Fig. 2C and Fig. S8). SIMPER run on the endosphere samples from vegetative and flowering stages places these two genera as the most influential taxa for the community separation between these developmental stages (Table S1).

As Bacteroidetes (along with Proteobacteria) increases its abundance in the root fraction, we analysed this phylum in more details. Bacteroidetes and especially *Chitinophaga* spp. and *Flavobacterium* spp. were found to reduce pathogen load inside the plant roots of sugar beet with target antibiotic production by overexpressing polyketide synthases and non-ribosomal peptide synthetase genes (Carrión et al. 2019). In our study, we found these two Bacteroidetes genera more abundant in the roots than in the rhizosphere or soil, with their abundance especially high in older plants (Fig. 2D).

We investigated genera abundance in more details using Volcano plots. Here we present the increase and decrease microbial abundance with statistical power. For clarity, we compared plant selection in each soil type at the vegetative and flowering stages, where each genus is represented by a dot of a different size according to its total abundance and located on the X-axis according to its abundance in a given fraction against the bulk soil control. The Y-axis indicates the statistical confidence for suppression (if on the left-hand side of the graph) or selection (if on the right-hand side of the graph) (Fig. S9-10).

Loose soil community becomes different from the bulk soil over time, while there are almost no genera either suppressed or selected in the vegetative stage, they do appear during the flowering time signifying at least some plant roots influence over the microbiota thriving in the more distant soil. For both plant developmental stages, the rhizosphere is a place of suppression of *Bradyrhizobium* and *Rhizobium* in Alfisol and *Bacillus* in Vertisol. For the vegetative stage, rhizoplane and endosphere selection are clear, especially for *Rhizobium*, *Pseudoxanthomonas* and *Sphingopyxis* (genera belonging to Alphaproteobacteria). There is soil type specificity in the suppression/selection pattern; plants grown in Alfisol suppress *Bradyrhizobium*, in Vertisol *Bacillus*, while in Inceptisol plants strongly select for *Rhizobium*. In general, plant roots exert a weaker influence in their flowering comparing with the vegetative stage, while *Hydrogenophaga*, *Sphingomonas*, *Opitutus* and *Brachymonas* replace *Rhizobium*, *Pseudoxanthomonas* and *Sphingopyxis* as efficient root colonisers (Fig. S9-10).

<u>High rate of *Rhizobium* in pigeon pea roots</u>

During sampling, we kept the nodules attached to the roots. We therefore expected to observe a spike of nodule symbionts' abundance in the root samples. Bulk soil, as well as loose soil, contain relatively high proportions of *Bradyrhizobium*, while pigeon pea rhizosphere and roots are colonised predominantly by *Rhizobium* (Fig. 3A). The proportion of *Rhizobium* is reduced in the flowering stage (av. 3.5%) relative to the vegetative one (av. 8.7%), while the *Bradyrhizobium* and other Rhizobiales abundance is relatively stable with only *Bradyrhizobium* reduction in loose soil over plant lifetime. Focusing on plant roots only, we

observe both the soil type and plant genotype specificity in Rhizobiales selection (Fig. 3B and C). In general, plants (all the biological replicates for a given condition) are either highly colonised by *Bradyrhizobium* (as Asha in Inceptisol), or by *Rhizobium* (Durga in Inceptisol and Vertisol). Soil type influence on Rhizobiales community inside plant roots is comparable to the general soil type influence (PERMANOVA *pseudo-F* = 17.5 for all microbial taxa and 16.3 for Rhizobiales only), the importance of genotype increases almost two-fold (3.4 to 6.3 for all microbial community and Rhizobiales community, respectively), while developmental stage still being important has reduced *pseudo-F* value from 21.9 to 10.6 (Fig. 3D and Fig. 1A). This signifies that Rhizobiales are more influential than other microbial taxa inside legume plant roots, but their community stays relatively stable over plant lifetime.

In order to explain the higher abundance of *Rhizobium* over the pigeon pea native symbiont *Bradyrhizobium* (and *Ensifer*) we isolated bacteria from the soil and roots of pigeon pea grown in Indian soils.
We isolated and purified 60 colonies from the rhizosphere and 272 colonies from the root endosphere and nodules, of which 13 and 43, respectively were found to be unique strains. Isolates of root-inhabiting *Rhizobium* contribute to 28% abundance, followed by *Burkholderia*, *Microbacterium*, *Paenibacillus* and *Pseudomonas* with 7% each. *Bradyrhizobium* isolates make 5%, while *Ensifer* represents only 2% of the isolated community (Table S1). These values are broadly consistent with the root metagenomic screen output.

We sequenced the genomes of *Rhizobium*/*Agrobacterium* (11), *Burkholderia* and *Paraburkholderia* (4) and *Microbacterium* (3) isolates: none were associated with any *nod* or *nif* genes suggesting these strains are not symbionts. We also sequenced many *Bradyrhizobium* nodule isolates and confirmed their nodulation ability. However, a detailed discussion of nodule isolated strains will be presented in a separate publication.
We confirmed that *Bradyrhizobium* rather than *Rhizobium* nodulates pigeon pea by growing plants in controlled conditions of growth chambers in Oxford in sterile vermiculate. Plants inoculated with *Rhizobium* isolates (either single or in a mixed inoculation) were not nodulated, while their roots contain a high bacterial presence of ~10^7 colony forming units per root. Plants inoculated with *Bradyrhizobium* (either single or in a mixed inoculation) were nodulated (2-5 nodules per plant) and contained a similar bacterial presence to plants inoculated with *Rhizobium* isolates only. Plants inoculated with both *Rhizobium* and *Bradyrhizobium* strains (strain mixtures) formed nodules, but only *Rhizobium* was recovered from the roots (rhizoplane and endosphere combined) (at least 100-fold dominance of *Rhizobium* over *Bradyrhizobium*), while all the visible nodules harboured *Bradyrhizobium* strains only.

Discussion:

We identified that the principal factor controlling the assembly of pigeon pea microbiome was the plant fraction followed by developmental stage, soil type and the least important, but

a significant factor is the plant genotype (Fig. 1A). In previous work using legumes and non-legume plants grown in soils from the United Kingdom, we have also observed fraction, followed by soil and plant species to be the main factors (Tkacz et al. 2020). Similar importance of fractions, soil type, developmental stage and genotype was also observed for rice grown in the United States (Edwards et al. 2015, Edwards et al. 2018). This indicates that the plant presence itself influences the surrounding microbiota, while its exact profile is influenced by other factors as soil type.

We can also conclude that soil type is more important for loose soil and rhizosphere fractions, while plant fractions as rhizoplane and root endosphere are greatly affected by the plant developmental stage and genotype. (Fig. S3B). All our statistical analyses indicate that the plants exert a gradient of influence, greatest towards root and decreasing towards surrounding soil (Fig. 1B).

Irrespective of sample's geographical origin (India or UK), plant species or soil type, rhizoplane and root endosphere is colonised by Alphaproteobacteria and Bacteroidetes (Fig. 2A, Fig. S6 and Fig. S8). Proteobacteria, and especially their Alpha class are common root colonisers found across multiple soils and plant species as *Arabidopsis* (Bulgarelli et al. 2012), *Lotus* (Zgadzaj et al. 2016), barley (Garrido-Oter et al. 2018) and rice (Edwards et al. 2015). Alphaproteobacteria harbours bacterial taxa that are likely to be quick in metabolising plant-derived nutrients (Reinhold-Hurek et al. 2015), and many of them may have genomic traits similar to plant symbionts (Pini et al. 2011).

Bacteroidetes among Proteobacteria, Actinobacteria and Firmicutes were found to contain many genetic adaptations to interact with plant host (Levy et al. 2018). However, their abundance is also correlated with pathogen presence. *Cytophaga* spp. and *Flavobacterium* spp. reduce pathogen load inside the infected plant roots by antibiotic production (Carrión et al. 2019). While we have not tested any Bacteroidetes isolates for antifungal properties, we found three pigeon pea varieties to have an increased abundance of these genera in various soil types.

The main root-inhabiting genera were *Rhizobium*, *Pseudoxanthomonas* and *Sphingopyxis* for vegetative stage plants, suggesting that these genera are especially attracted by young plants possibly by an active plant secretion. Some plant genotypes, when grown in a specific soil, also had a high abundance of *Bradyrhizobium* suggesting an active endosymbiosis.

Vegetative plants were associated with *Sphingopyxis* root colonisation, which was replaced by *Brachymonas* in the flowering stage (Fig. 2C).

Sphingopyxis and *Brachymonas* genera are rarely found in plant microbiomes. A *Sphingopyxis* isolate was found to be an inconsistent root coloniser in competition with a synthetic community (Voges et al. 2019). However, in our case, this genus was consistently associated with roots, irrespectively of the soil type or plant genotype. Conversely, we consider *Brachymonas* to be an opportunistic root coloniser of older plants that no longer invest resources in interaction with its microbiota (Mougel et al. 2006). In general, young plants strongly associate with only a part of surrounding microbiota, as the bacterial diversity is lower on and inside the root comparing to the surrounding soil. However, over time with the flowering stage, the plant loses its selective pressure allowing various other bacteria to colonise the roots (Fig. 1D).

We confirmed elevated *Bradyrhizobium* presence for selected genotypes grown in selected soil types (i.e., Asha in Inceptisol and Durga in Vertisol - Fig. 3 B and C), however, contrary to our expectations, *Bradyrhizobium* presence was generally low in the root fraction, while they can be a dominant genus in the surrounding soil. This genus contains free-living, non-symbiotic strains, which can dominate soil *Bradyrhizobium* community in forest soils (Vaninsberghe et al. 2015) and were also found in agricultural soils (Jones et al. 2016). Here we have not established what proportion of soil *Bradyrhizobium* contain symbiotic properties, a question worth investigating in the future.

Rhizobium was the most abundant root colonising species in our metagenomic assay (Fig. 3A) and we confirmed its dominance over *Bradyrhizobium* with isolation studies from native Indian soils (Fig. 2F). While these *Rhizobium* isolates lack *nod* and *nif* genes and are unable to nodulate pigeon pea, they can outcompete *Bradyrhizobium* in native soils and in gnotobiotic conditions. A similar effect of *Bradyrhizobium* being outcompeted was observed for soybean seedlings containing natural seed epiphytes (Oehrle et al. 2000).

Moreover, the case of pigeon pea is not alone as roots of soybean, the *Bradyrhizobium* host plant, were found to be colonised with a microbial community where this symbiont contributes to only ~1% of the population (Liu et al. 2019, Rascovan et al. 2016). Such low abundance are in contrast with pea plants with ~10-20%, *Medicago* with ~10-60% or *Lotus* with ~10% root presence of their respective symbiotic genus i.e., *Rhizobium, Ensifer* or *Mesorhizobium*, respectively (Brown et al. 2020, Cordero et al. 2020, Horner et al. 2019, Tkacz et al. 2020, Zgadzaj et al. 2016).

Hence, we conclude that *Bradyrhizobium*, while being abundant in Indian soils, is either a poor root coloniser of its host plant under competition from other soil-dwelling bacteria or contain many non-symbiotic strains. Despite that, plants can still be nodulated. We speculate that symbiotic *Bradyrhizobium* colonises the emerging root hairs directly from the soil, where its number is high rather than actively colonising the root and moving towards the emerging nodule regions. This feature may explain poor pigeon pea nodulation in native soils where either their soil population is low or are outcompeted in root colonisation. Hence, any selection for pigeon pea inoculants should not only be based on their N_2-fixing potential, but also on their competitiveness against other root-colonising bacteria.

Availability of data and materials

16S rRNA gene sequencing data and associated metadata were deposited to EBI SRA repository under accession code PRJEB39218. Genome data is stored in GenBank database as Bioproject PRJNA693523. Detailed documentation of the bioinformatic pipeline and data analysis output used for figure preparation and statistical analysis can be found in Table S1.

Abbreviations

ANOSIM: Analysis of similarities

MKK: Mannem Konda Kandi
PCA: Principal component analysis
PCoA: Principal coordinate analysis
PERMANOVA: Permutational multivariate analysis of variance
SIMPER: Similarity percentage analysis
zOTU: zero-radius operational taxonomic unit

Contributions

AT, PP and ARP conceived and planned the study. AT, DC, AB and SVSRNP designed the study and conducted the experiments, AT, BJ and ALN analysed the data, AT, AB, DC, SVSRNP, BJ and PP drafted the manuscript. All authors have read, critically revised and approved the final version of the manuscript.

Funding

This work was supported by the Biotechnology and Biological Sciences Research Counil [grant number BB/N013387/1]. This work was financially supported by the Department of Biotechnology (DBT), Government of India (GoI) under project No: BT/IN/UK-VNC/41/DLN/2015-16 and Department of Science and Technology (DST), GoI under project No: EMR/2017/002658, granted to ARP. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

The authors gratefully acknowledge Dr. DLN Rao, ICAR-Emeritus Scientist, Indian Institute of Soil Science, Bhopal, India, for his critical inputs in the study. The authors acknowledge DST, GoI for J.C. Bose Fellowship to ARP [JCB/2017/000053] and University Grants Commission (UGC) for research fellowships to AB and DC. ARP, DC, AB and SVSRNP acknowledge DST Funds for Infrastructure in Science and Technology (DST-FIST), Level II and UGC Special Assistance Programme (UGC-SAP) support to the Department of Plant Sciences at University of Hyderabad. DC would like to thank Department of Biotechnology (DBT), GoI, for Newton Bhabha Ph.D. Placement Award [BT/IN/NBPP/CD/28/2018-19]. DC would like to thank Dr. David Hughes, Computational and Analytical Sciences, Rothamsted Research, Harpenden, UK, for technical support.

Ethical declarations

Ethics approval and consent to participate
Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

Allaway, D., E. M. Lodwig, L. A. Crompton, M. Wood, R. Parsons, T. R. Wheeler and P. S. Poole (2000). "Identification of alanine dehydrogenase and its role in mixed secretion of ammonium and alanine by pea bacteroids." Mol Microbiol **36**(2): 508-515.

Arora, N. K., E. Khare, S. Singh and S. Tewari (2018). "Phenetic, genetic diversity and symbiotic compatibility of rhizobial strains nodulating pigeon pea in Northern India." 3 Biotech **8**(1): 52.

Bankevich, A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, S. I. Nikolenko, S. Pham, A. D. Prjibelski, A. V. Pyshkin, A. V. Sirotkin, N. Vyahhi, G. Tesler, M. A. Alekseyev and P. A. Pevzner (2012). "SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing." J Comput Biol **19**(5): 455-477.

Bind, M. and S. Nema (2019). "Isolation and Molecular Characterization of Endophytic Bacteria from Pigeon Pea along with Antimicrobial Evaluation against *Fusarium udum*."

Brown, S. P., M. A. Grillo, J. C. Podowski and K. D. Heath (2020). "Soil origin and plant genotype structure distinct microbiome compartments in the model legume *Medicago truncatula*." Microbiome **8**(1): 139.

Bulgarelli, D., M. Rott, K. Schlaeppi, E. Ver Loren van Themaat, N. Ahmadinejad, F. Assenza, P. Rauf, B. Huettel, R. Reinhardt, E. Schmelzer, J. Peplies, F. O. Gloeckner, R. Amann, T. Eickhorst and P. Schulze-Lefert (2012). "Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota." Nature **488**(7409): 91-95.

Caporaso, J. G., C. L. Lauber, W. A. Walters, D. Berg-Lyons, C. A. Lozupone, P. J. Turnbaugh, N. Fierer and R. Knight (2011). "Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample." Proc Natl Acad Sci U S A **108**(Supplement 1): 4516-4522.

Carrión, V. J., J. Perez-Jaramillo, V. Cordovez, V. Tracanna, M. de Hollander, D. Ruiz-Buck, L. W. Mendes, W. F. J. van Ijcken, R. Gomez-Exposito, S. S. Elsayed, P. Mohanraju, A. Arifah, J. van der Oost, J. N. Paulson, R. Mendes, G. P. van Wezel, M. H. Medema and J. M. Raaijmakers (2019). "Pathogen-induced activation of disease-suppressive functions in the endophytic root microbiome." Science **366**(6465): 606-612.

Chaparro, J. M., D. V. Badri and J. M. Vivanco (2014). "Rhizosphere microbiome assemblage is affected by plant development." ISME J **8**(4): 790-803.

Cordero, J., J. R. d. Freitas and J. J. Germida (2020). "Bacterial microbiome associated with the rhizosphere and root interior of crops in Saskatchewan, Canada." Can J Microbiol **66**(1): 71-85.

Dubey, R. C., D. K. Maheshwari, H. Kumar and K. Choure (2010). "Assessment of diversity and plant growth promoting attributes of rhizobia isolated from *Cajanus cajan* L." Afr J Biotech **9**(50): 8619-8629.

Edgar, R. C. (2016). "UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing." bioRxiv: 081257.

Edwards, J., C. Johnson, C. Santos-Medellín, E. Lurie, N. K. Podishetty, S. Bhatnagar, J. A. Eisen and V. Sundaresan (2015). "Structure, variation, and assembly of the root-associated microbiomes of rice." PNAS **112**(8): E911-E920.

Edwards, J. A., C. M. Santos-Medellín, Z. S. Liechty, B. Nguyen, E. Lurie, S. Eason, G. Phillips and V. Sundaresan (2018). "Compositional shifts in root-associated bacterial and archaeal microbiota track the plant life cycle in field-grown rice." PLOS Biol **16**(2): e2003862.

FAOSTAT. (2019). "Online Agriculture Statistics." 2020, from http://www.faostat.org.

Fossou, R. K., D. Ziegler, A. Zeze, F. Barja and X. Perret (2016). "Two Major Clades of Bradyrhizobia Dominate Symbiotic Interactions with Pigeonpea in Fields of Cote d'Ivoire." Front Microbiol **7**: 1793.
Fuller, D. Q., C. Murphy, E. Kingwell-Banham, C. C. Castillo and S. Naik (2019). "*Cajanus cajan* (L.) Millsp. origins and domestication: the South and Southeast Asian archaeobotanical evidence." Genet Res Crop Evol **66**(6): 1175-1188.
Garrido-Oter, R., R. T. Nakano, N. Dombrowski, K. W. Ma, A. C. McHardy and P. Schulze-Lefert (2018). "Modular Traits of the Rhizobiales Root Microbiota and Their Evolutionary Relationship with Symbiotic Rhizobia." Cell Host Microbe **24**(1): 155-167.e155.
Gosai, J., S. Anandhan, A. Bhattacharjee and G. Archana (2020). "Elucidation of quorum sensing components and their role in regulation of symbiotically important traits in *Ensifer* nodulating pigeon pea." Microbiol Res **231**: 126354.
Hartman, K., M. G. van der Heijden, V. Roussely-Provent, J. C. Walser and K. Schlaeppi (2017). "Deciphering composition and function of the root microbiome of a legume plant." Microbiome **5**(1): 2.
Horner, A., S. S. Browett and R. E. Antwis (2019). "Mixed-Cropping Between Field Pea Varieties Alters Root Bacterial and Fungal Communities." Sci Rep **9**(1): 16953.
Jain, D., A. Kumari, H. Saheewala, S. Sanadhya, D. Maheshwari, R. H. Meena, A. Singh, R. Gera and S. R. Mohanty (2020). "Biochemical, functional and molecular characterization of pigeon pea rhizobia isolated from semi-arid regions of India." Arch Microbiol **202**(7): 1809-1816.
Jones, F.P., I.M. Clark, R. King, L.J. Shaw, J. Woodward and P.R. Hirsh (2016). "Novel European free-living, non-diazotrophic *Bradyrhizobium* isolates from contrasting soils that lack nodulation and nitrogen fixation genes - a genome comparison." Sci Rep **6**: 25858.
Levy, A., I. Salas Gonzalez, M. Mittelviefhaus, S. Clingenpeel, S. Herrera Paredes, J. Miao, K. Wang, G. Devescovi, K. Stillman, F. Monteiro, B. Rangel Alvarez, D. S. Lundberg, T.-Y. Lu, S. Lebeis, Z. Jin, M. McDonald, A. P. Klein, M. E. Feltcher, T. G. Rio, S. R. Grant, S. L. Doty, R., E. Ley, B. Zhao, V. Venturi, D. A. Pelletier, J. A. Vorholt, S. G. Tringe, T. Woyke and J. L. Dangl (2018). "Genomic features of bacterial adaptation to plants." Nat Genet **50**(1): 138-150.
Li, H. and R. Durbin (2009). "Fast and accurate short read alignment with Burrows-Wheeler transform." Bioinformatics **25**(14): 1754-1760.
Liu, F., T. Hewezi, S. L. Lebeis, V. Pantalone, P. S. Grewal and M. E. Staton (2019). "Soil indigenous microbiome and plant genotypes cooperatively modify soybean rhizosphere microbiome assembly." BMC Microbiol **19**(1): 201.
Liu, F., T. Hewezi, S. L. Lebeis, V. Pantalone, P. S. Grewal and M. E. Staton (2019). "Soil indigenous microbiome and plant genotypes cooperatively modify soybean rhizosphere microbiome assembly." BMC Microbiol **19**(1): 201.
López-Baena, F. J., J. E. Ruiz-Sainz, M. A. Rodríguez-Carvajal and J. M. Vinardell (2016). "Bacterial Molecular Signals in the *Sinorhizobium fredii*-Soybean Symbiosis." Int J Mol Sci **17**(5): 755.
Lundberg, D. S., S. Yourstone, P. Mieczkowski, C. D. Jones and J. L. Dangl (2013). "Practical innovations for high-throughput amplicon sequencing." Nat Methods **10**(10): 999-1002.
Mendes, L. W., E. E. Kuramae, A. A. Navarrete, J. A. van Veen and S. M. Tsai (2014). "Taxonomical and functional microbial community selection in soybean rhizosphere." ISME J **8**(8): 1577-1587.
Mendes, L. W., J. M. Raaijmakers, M. de Hollander, R. Mendes and S. M. Tsai (2018). "Influence of resistance breeding in common bean on rhizosphere microbiome composition and function." ISME J **12**(1): 212-224.

Miranda-Sanchez, F., J. Rivera and P. Vinuesa (2016). "Diversity patterns of Rhizobiaceae communities inhabiting soils, root surfaces and nodules reveal a strong selection of rhizobial partners by legumes." Environ Microbiol **18**(8): 2375-2391.
Mougel, C., P. Offre, L. Ranjard, T. Corberand, E. Gamalero, C. Robin and P. Lemanceau (2006). "Dynamic of the genetic structure of bacterial and fungal communities at different developmental stages of *Medicago truncatula* Gaertn. cv. Jemalong line J5." New Phytol **170**(1): 165-175.
Murthy, A. S. P. (1988). Distribution, Properties, and Management of Vertisols of India. In: Stewart B.A. (eds) Advances in Soil Science. vol 8. Springer, New York, NY. https://doi.org/10.1007/978-1-4613-8771-8_4.
Nautiyal, C. S., S. V. Hegde and P. van Berkum (1988). "Nodulation, Nitrogen Fixation, and Hydrogen Oxidation by Pigeon Pea *Bradyrhizobium* spp. in Symbiotic Association with Pigeon Pea, Cowpea, and Soybean." Appl Environ Microbiol **54**(1): 94-97.
Oehrle, N. W., D. B. Karr, R. J. Kremer and D. W. Emerich (2000). "Enhanced attachment of *Bradyrhizobium japonicum* to soybean through reduced root colonization of internally seedborne microorganisms." Can J Microbiol **46**(7): 600-606.
Perez-Jaramillo, J. E., V. J. Carrion, M. Bosse, L. F. V. Ferrao, M. de Hollander, A. A. F. Garcia, C. A. Ramirez, R. Mendes and J. M. Raaijmakers (2017). "Linking rhizosphere microbiome composition of wild and domesticated *Phaseolus vulgaris* to genotypic and root phenotypic traits." ISME J **11**(10): 2244-2257.
Pini, F., M. Galardini, M. Bazzicalupo and A. Mengoni (2011). "Plant-bacteria association and symbiosis: are there common genomic traits in alphaproteobacteria?" Genes (Basel) **2**(4): 1017-1032.
Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies and F. O. Glöckner (2013). "The SILVA ribosomal RNA gene database project: improved data processing and web-based tools." Nucl Acids Res **41**(Database issue): D590-D596.
Rajendran, G., F. Sing, A. J. Desai and G. Archana (2008). "Enhanced growth and nodulation of pigeon pea by co-inoculation of *Bacillus* strains with *Rhizobium* spp." Bioresour Technol **99**(11): 4544-4550.
Rascovan, N., B. Carbonetto, D. Perrig, M. Díaz, W. Canciani, M. Abalo, J. Alloati, G. González-Anta and M. P. Vazquez (2016). "Integrated analysis of root microbiomes of soybean and wheat from agricultural fields." Sci Rep **6**: 28084.
Reinhold-Hurek, B., W. Bünger, C. S. Burbano, M. Sabale and T. Hurek (2015). "Roots Shaping Their Microbiome: Global Hotspots for Microbial Activity." Ann Rev Phytopathol **53**(1): 403-424.
Rupela, O. P. and C. Johansen (1995). "Identification of non-nodulating, and low and high nodulating plants in pigeonpea." Soil Biol Biochem **27**(4): 539-544.
Seemann, T. (2014). "Prokka: rapid prokaryotic genome annotation." Bioinformatics **30**(14): 2068-2069.
Semwal, D. P., S. P. Ahlawat and K. Pradheep (2018). "Pigeonpea [*Cajanus cajan* (L.) Millsp.] and its wild spp. germplasm collection status, diversity distribution and trait-specific germplasm mapping using GIS tools in India." Legum Res **41**(5): 656-662.
Singh, K., S. Dhull and R. Gera (2018). "Genetic Diversity of Diazotrophs Nodulating Pigeon Pea in Arid and Semi-Arid Zones of Haryana, India." Int J Curr Microbiol Appl Sci **7**(3): 3447-3467.
Singh, R. P. (2013). Status paper on pulses. Government of India, Ministry of Agriculture & Farmers Welfare, Directorate of Pulses Development.
Singha, B., P. B. Mazumder and P. Pandey (2018). "Characterization of Plant Growth Promoting Rhizobia from Root Nodule of Two Legume Species Cultivated in Assam, India." Proc Natl Acad Sci India B **88**(3): 1007-1016.

Tiwari, A. K. and A. K. Shivhare (2016). "Pulses in India: retrospect and prospects." Government of India, Ministry of Agriculture & Farmers Welfare, Directorate of Pulses Development.
Tiwari, A. K. and A. K. Shivhare (2018). "Pulses in India: retrospect and prospects" - 2018. Government of India, Ministry of Agriculture & Farmers Welfare, Directorate of Pulses Development.
Tkacz, A., E. Bestion, B. Zhiyan, M. Hortala and P. S. Poole (2020). "Influence of Plant Fraction, Soil, and Plant Species on Microbiota: a Multikingdom Comparison." mBio **11**(1): e02785-02719.
Varshney, R. K., R. K. Saxena, H. D. Upadhyaya, A. W. Khan, Y. Yu, C. Kim, A. Rathore, D. Kim, J. Kim, S. An, V. Kumar, G. Anuradha, K. N. Z. Yamini, W., S. Muniswamy, J. Kim, R. V. Penmetsa, E. von Wettberg and S. K. Datta (2017). "Whole-genome resequencing of 292 pigeonpea accessions identifies genomic regions associated with domestication and agronomic traits." Nat Genet **49**(7): 1082-1088.
Vaninsberghe, D., K. R. Maas, E. Cardenas, C. R. Strachan, S. J. Hallam and W. M Mohn (2015). "Non-symbiotic *Bradyrhizobium* ecotypes dominate North American forest soils." ISME J **9**(11): 2435-2441.
Versalovic, J., M. Schneider, F. J. de Bruijn and L. J.R. (1994). "Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction." Methods Mol Cell Biol **5**: 25-40.
Voges, M. J. E. E. E., Y. Bai, P. Schulze-Lefert and E. S. Sattely (2019). "Plant-derived coumarins shape the composition of an *Arabidopsis* synthetic root microbiome." Proc Natl Acad Sci U S A **116**(25): 12558-12565.
Wood, D. E. and S. L. Salzberg (2014). "Kraken: ultrafast metagenomic sequence classification using exact alignments." Genome Biol **15**(3): R46.
Xiao, X., W. Chen, L. Zong, J. Yang, S. Jiao, Y. Lin, E. Wang and G. Wei (2017). "Two cultivated legume plants reveal the enrichment process of the microbiome in the rhizocompartments." Mol Ecol **26**(6): 1641-1651.
Zgadzaj, R., R. Garrido-Oter, D. B. Jensen, A. Koprivova, P. Schulze-Lefert and S. Radutoiu (2016). "Root nodule symbiosis in *Lotus japonicus* drives the establishment of distinctive rhizosphere, root, and nodule bacterial communities." Proc Natl Acad Sci U S A **113**(49): E7996-E8005.

Figures

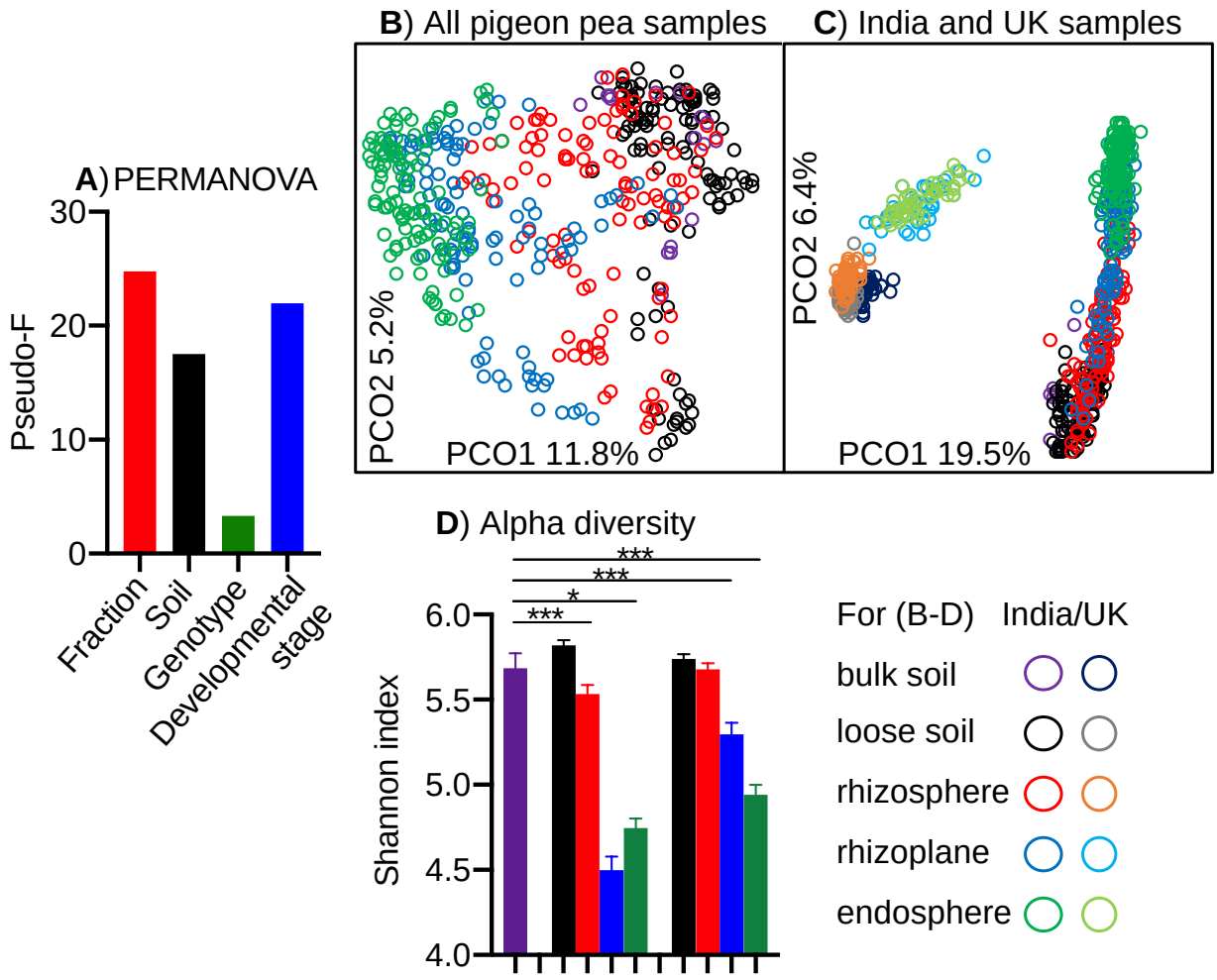

Fig. 1 Community structure and diversity of pigeon pea associated microbiota. A) Influence of different factors on microbiota using PERMANOVA *pseudo-F* value as a proxy, B) PCoA plot representing pigeon pea microbiota and shown with the visual separation fractions, C) PCoA plot representing pigeon pea microbiota and British soils-grown *Arabidopsis thaliana*, wheat, *Medicago truncatula* and pea, D) Shannon entropy shown for each fraction. The outcome of *t*-test for bulk soil against each fraction with Bonferroni correction is indicated above the bar plots; * p <0.05, ***p <0.001. For panels A, B and D, n = 449, while for panel C, n = 713.

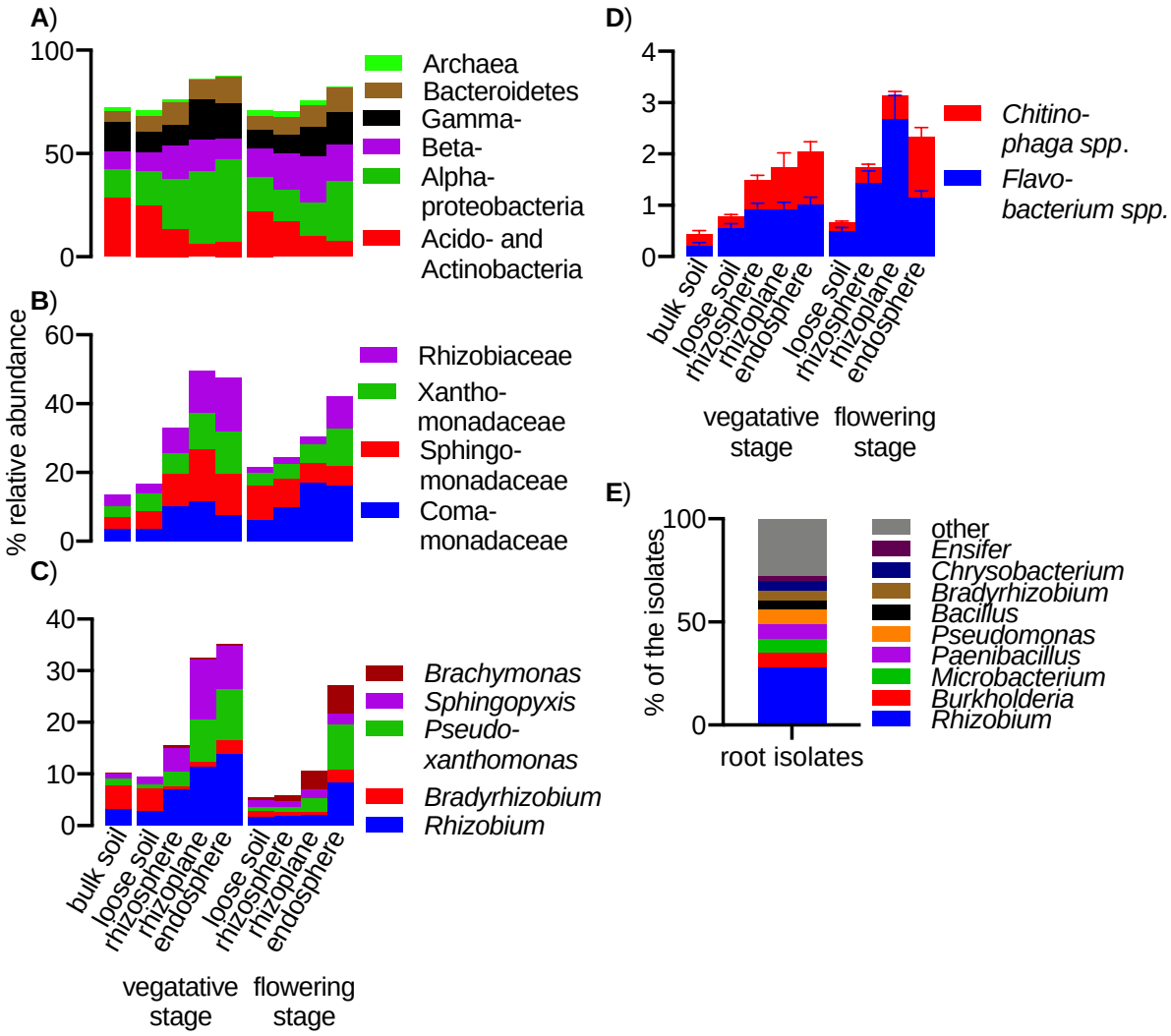

Fig. 2 Taxonomic profile of pigeon pea microbiota at A) phylum, B) family, and C) genus level. Top seven phyla, top four families and top five genera are shown as a percentage of the total community. D) potential beneficial genera belonging to Bacteroidetes phylum and E) a detailed profile of the Rhizobiales community in relation to the whole microbiota. F) Taxonomic profile of the bacterial isolates from roots of pigeon pea grown in native soils

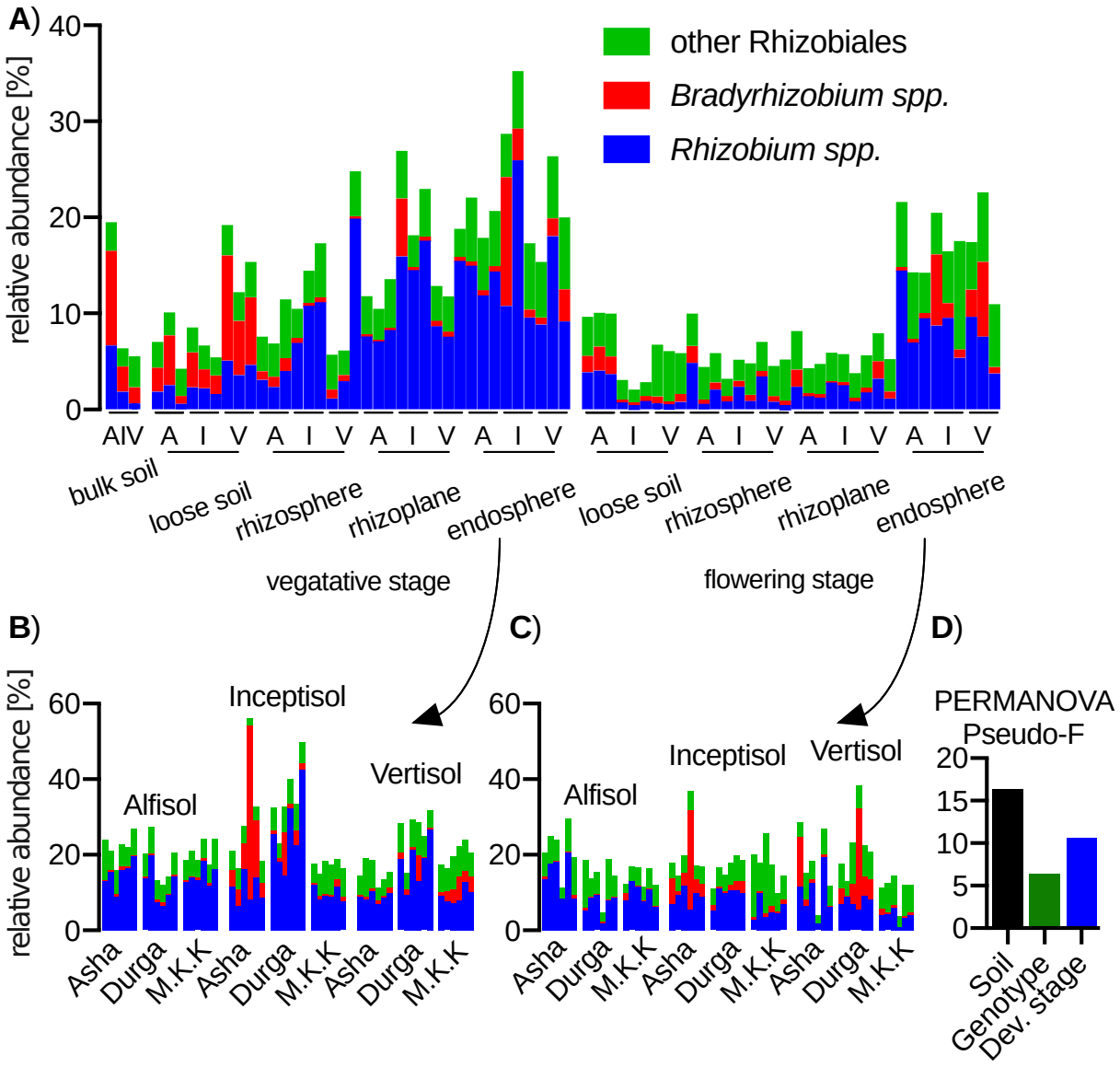

Fig 3. *Rhizobium* spp., *Bradyrhizobium* spp. and other Rhizobiales community taxonomic metagenomic profile associated with pigeon pea plants, A) for all fractions, B) for vegetative stage endosphere showing all biological replicates, C) for flowering stage endosphere showing all biological replicates and D) PERMANOVA output for the relative strength of Rhizobiales on the endosphere community with the separation for the soil type, plant genotype and developmental stage influence (P <0.001 for all comparisons). Soils abbreviated as A – Alfisol, I – Inceptisol, V – Vertisol.

Supplementary figures

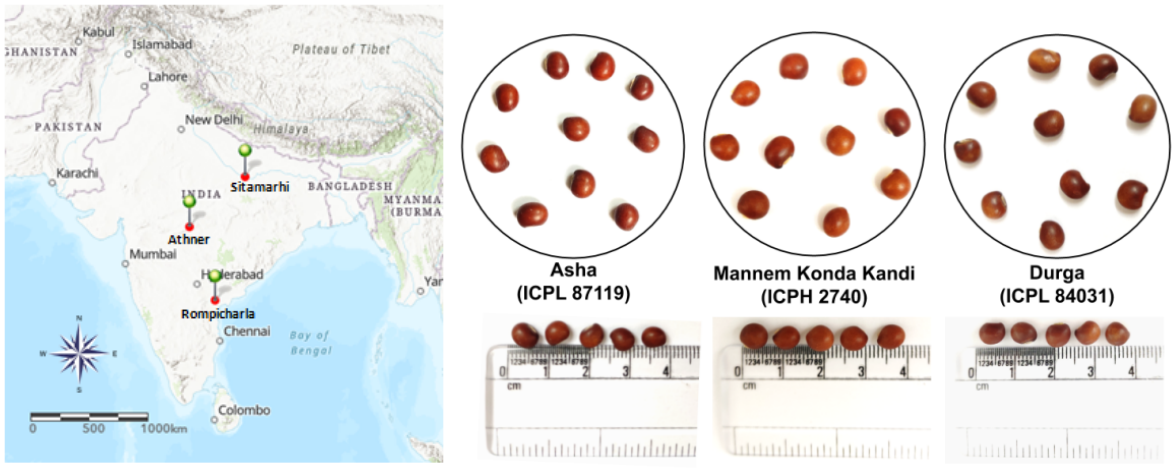

Fig S1: Left: Map of India showing GPS location of pigeon pea sampling sites at Rompicharla (Andhra Pradesh), Athner (Madhya Pradesh) and Sitamarhi (Uttar Pradesh). Map drawn with ArcGIS software.
Right: Photographs depicting the texture and size of seeds of the pigeon pea cultivars Asha (ICPL 87119), Mannem Konda Kandi or MKK (ICPH 2740) and Durga (ICPL 84031).

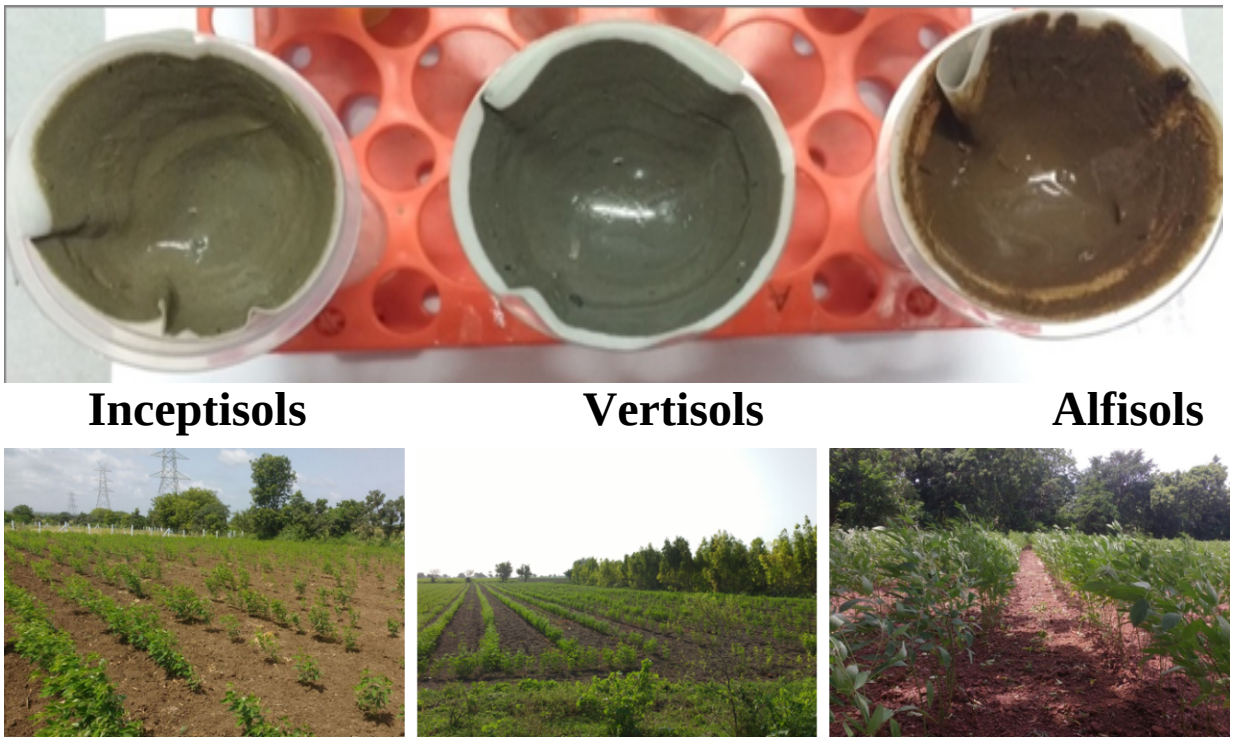

Fig. S2. Soil types used in the study. Top panel (L to R): Inceptisol, Vertisol and Alfisol samples for soil physicochemical characterisation. Bottom panel (L to R): Pigeon pea fields in Mirzapur (Uttar Pradesh), Betul (Madhya Pradesh), and Rompicharla (Andhra Pradesh), having Inceptisol, Vertisol and Alfisol, respectively. Soil samples were collected from the fields pictured above.

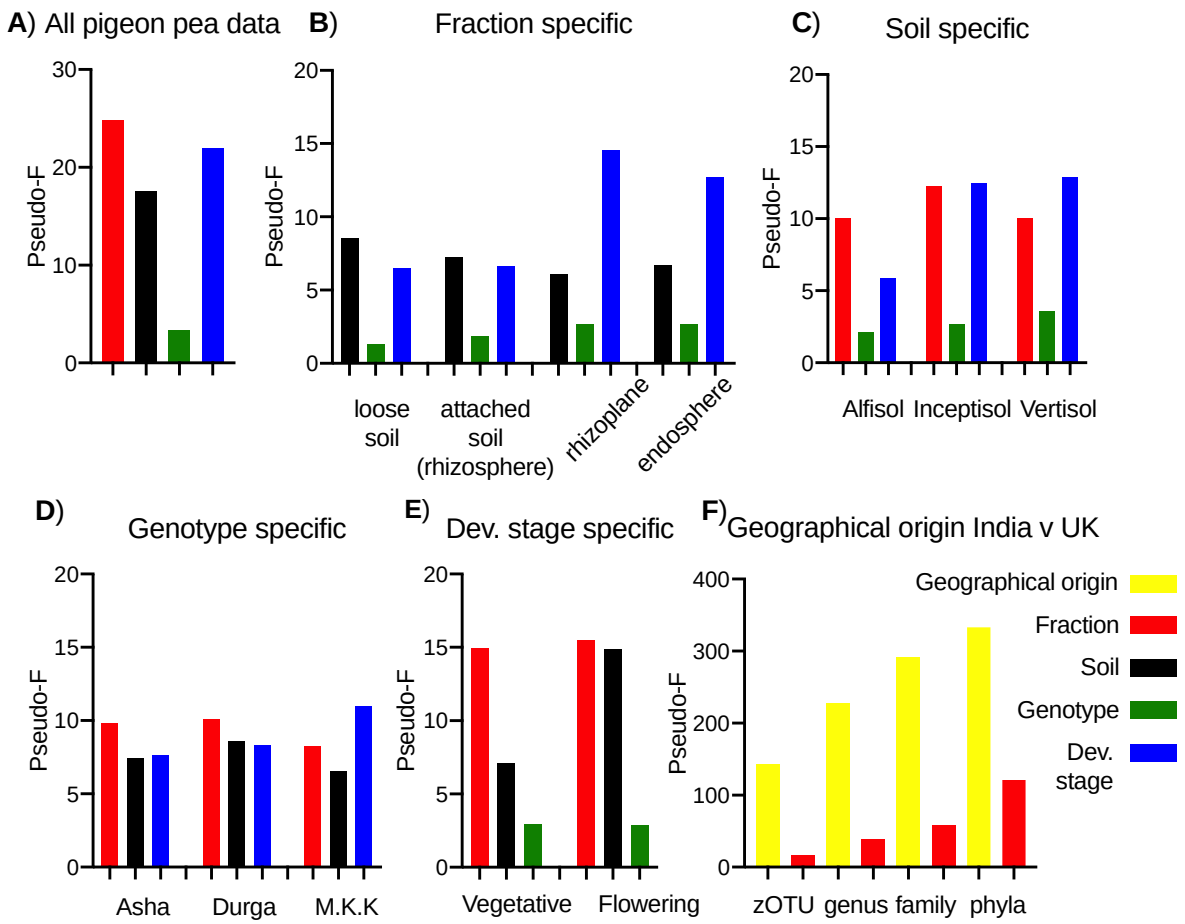

Fig. S3 PERMANOVA output measuring the influence of factors on microbiota using *pseudo-F* value as a proxy. A) All pigeon pea data, B) Pigeon pea data split by each fraction, C) Pigeon pea data split by soil, D) Pigeon pea data split by genotype, E) Pigeon pea data split by developmental stage, F) A comparison of Indian pigeon pea samples and British plants at different prokaryotic taxonomic units.

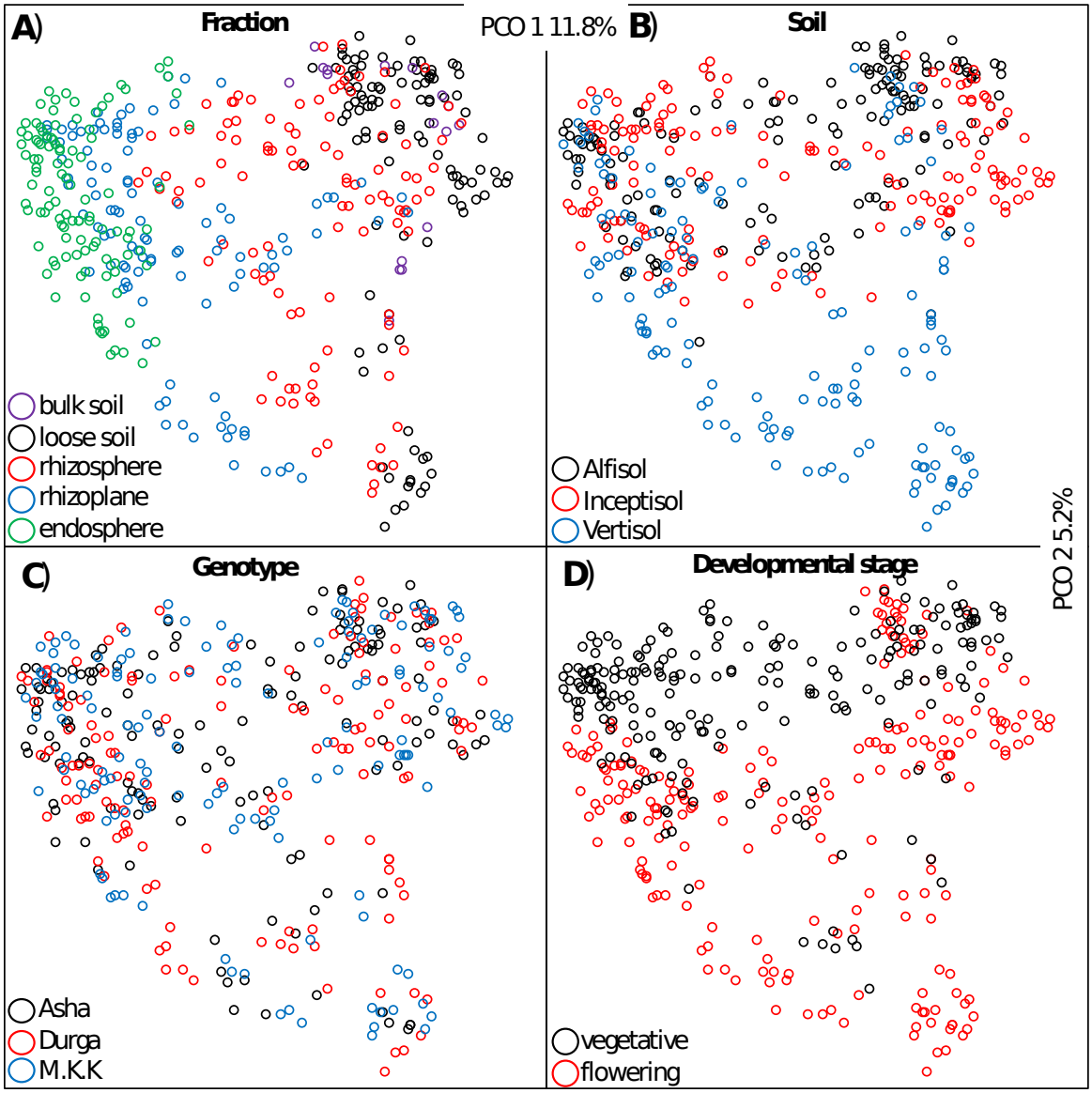

Fig. S4. PCoA plots representing pigeon pea microbiota and shown with the visual separation by A) fractions, B) soil type, C) plant genotype and D) plant developmental stage, *n* = 449.

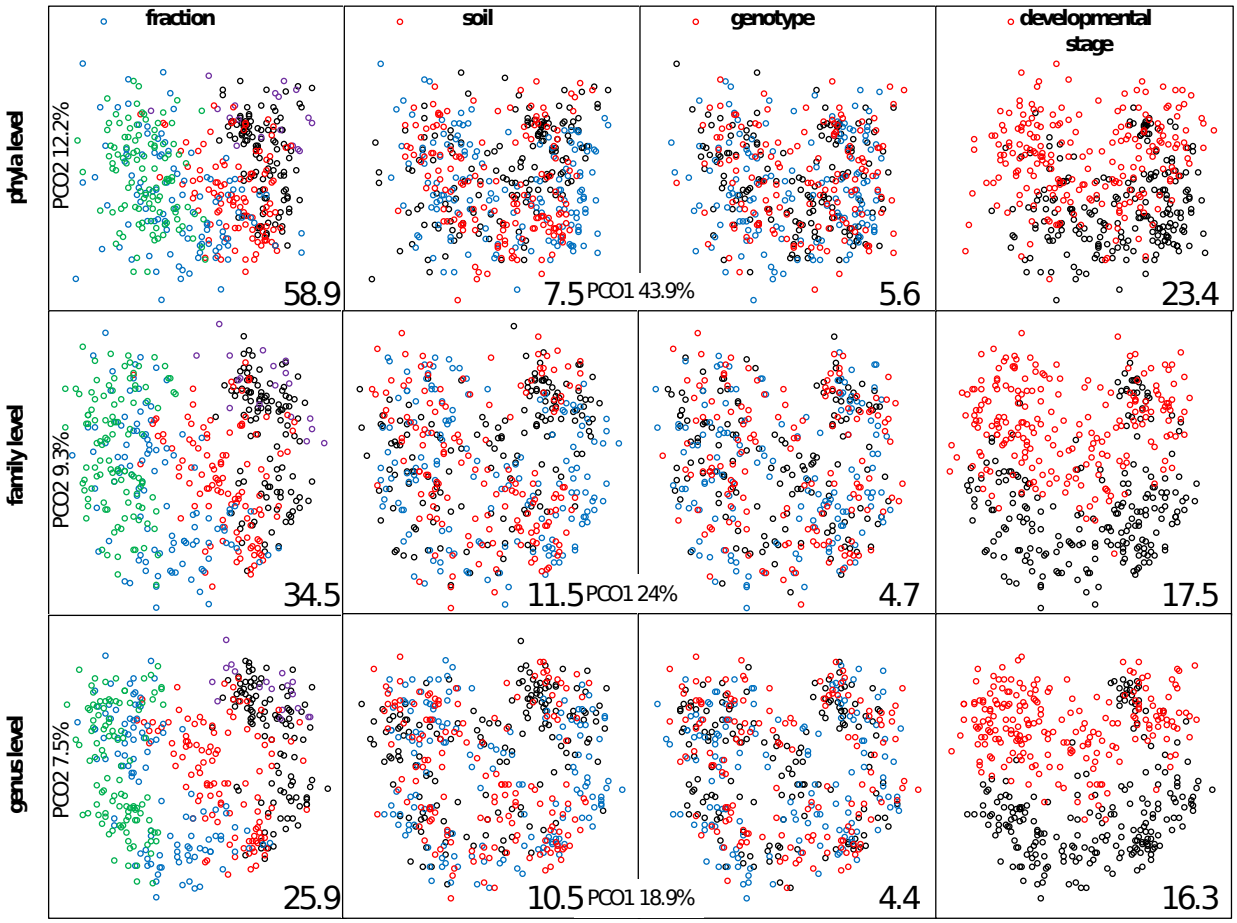

Fig. S5. PCoA plots representing pigeon pea microbial community structure at phyla, family and genus taxonomical level data using fraction, soil type, genotype and developmental stage. The same colours are used as in Fig. S4. For each panel the number in the bottom right indicates PERMANOVA pseudo-F value for a given factor sample separation. All comparisons p <0.01. n = 449 samples.

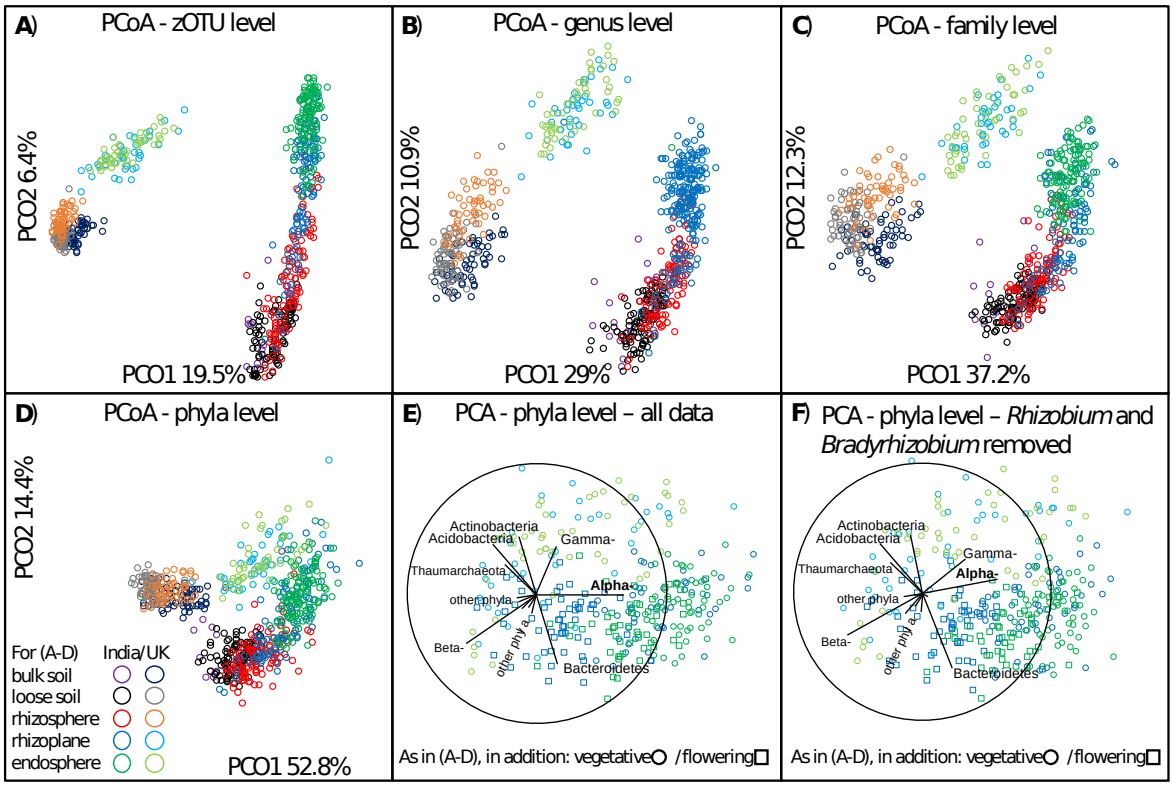

Fig. S6. Microbial community structure visualised with PCoA plots at different taxonomic levels; A) zOTU level, B) genus level, C) family level and D) phyla level, of Indian pigeon pea and British *Arabidopsis thaliana*, bread wheat, *Medicago truncatula* and *Pisum sativum*. E) PCA plot with rhizoplane and root endosphere samples showing six dominant phyla abundance shaping the community (Alpha-, Beta- and Gamma- stand for Proteobacteria classes. F) as E) with *Rhizobium* and *Bradyrhizobium* species *in-silico* removed. For PCoA plots n = 713, for PCA plot n = 303.

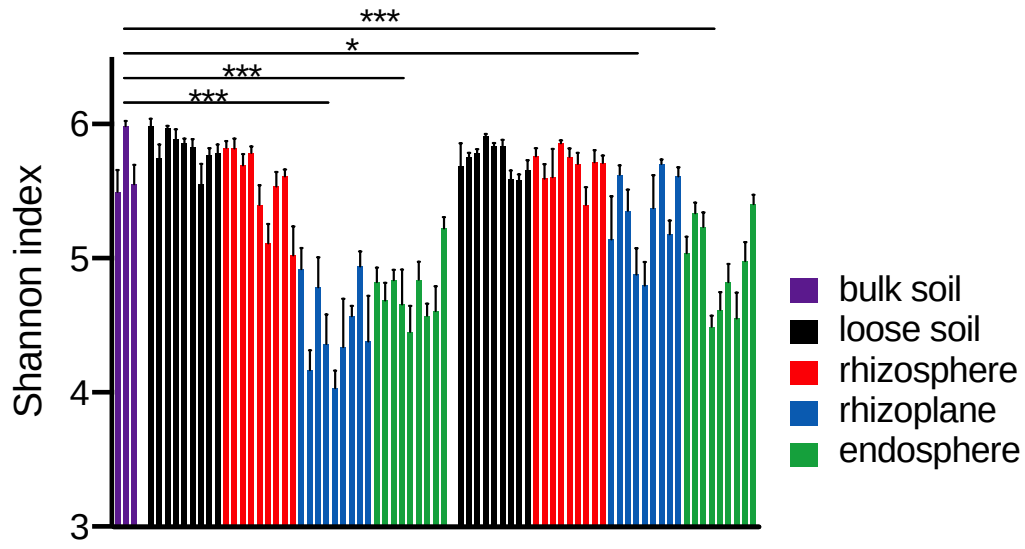

Fig. S7. Shannon entropy associated with each fraction (indicated by colour) and presented for each soil; from left Alfisol, Inceptisol and Vertisol and each genotype inside the soil cluster; from left to right hand side: Asha, Durga and MKK. The outcome of *t*-test for initial soil against each fraction with Bonferroni correction is indicated above the bar plots; * p <0.05, ***p <0.001.

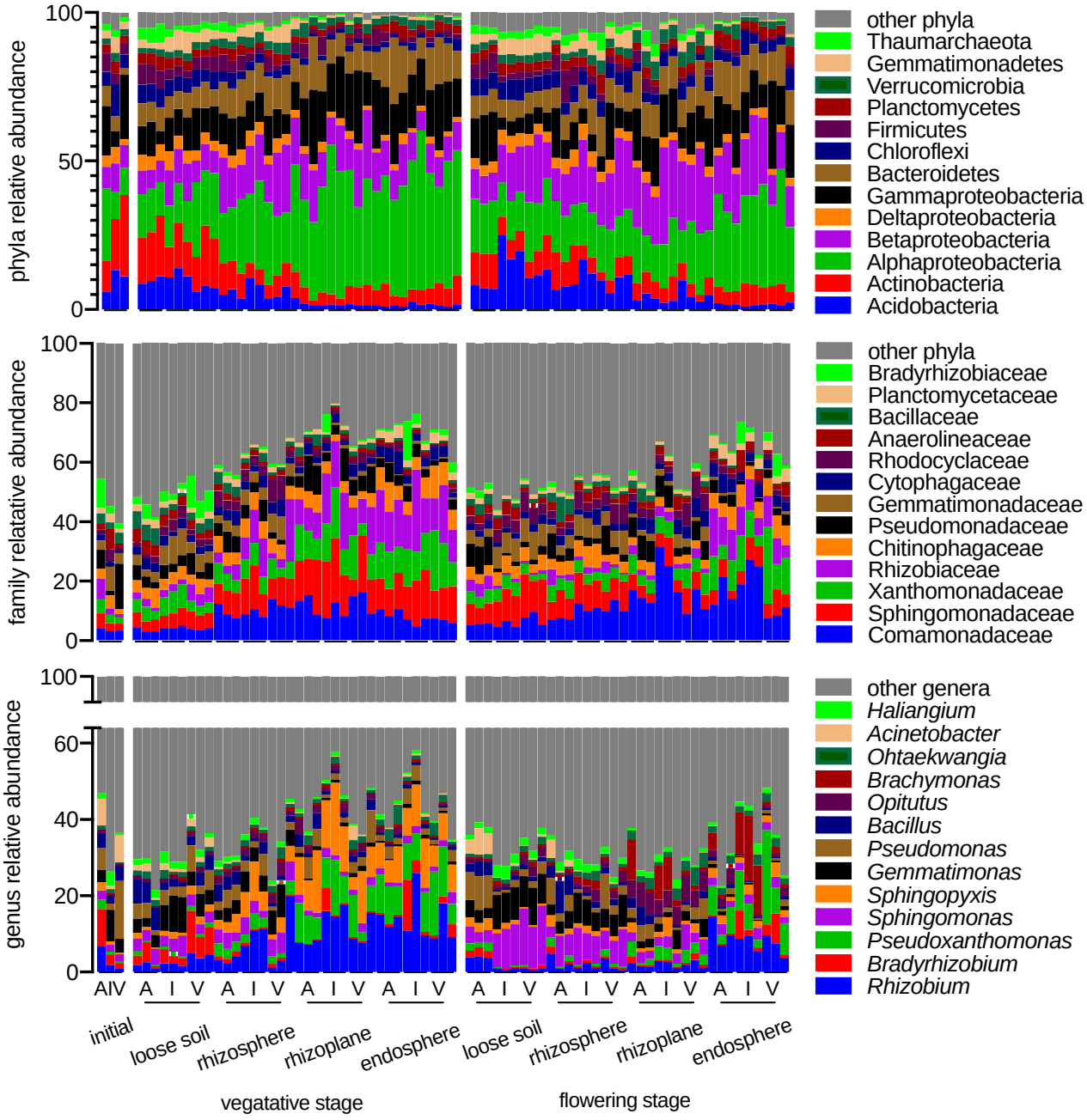

Fig. S8. Pigeon pea microbiota taxonomic profile at phylum, family and genus level for each developmental stage, fraction, soil and genotype. Soils abbreviated as A – Alfisol, I – Inceptisol, V – Vertisol and each genotype inside the soil cluster; from left to right: Asha, Durga and MKK.

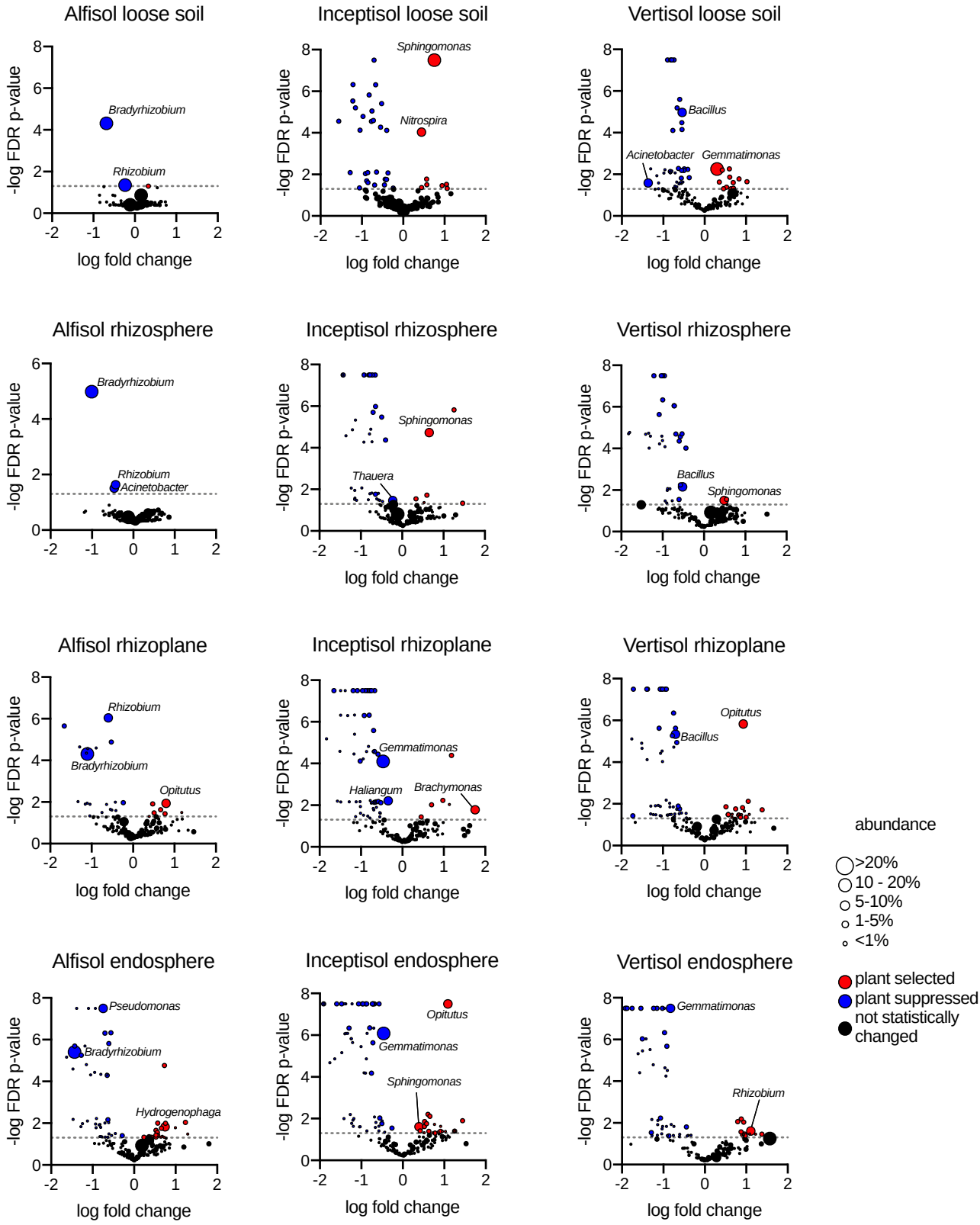

Fig S9. Volcano plots for the microbial community during plant's vegetative stage. Each genus is presented as a dot where X-axis indicates the fold difference between the bulk soil and the respective fraction (in logarithmic scale) and Y-axis indicates the significance according to *t*-test *p*-value corrected with false rate discovery and presented as -log scale (-log(0.05) = 1.301 and presented as a dotted grey line on plots). Selected genera are annotated.

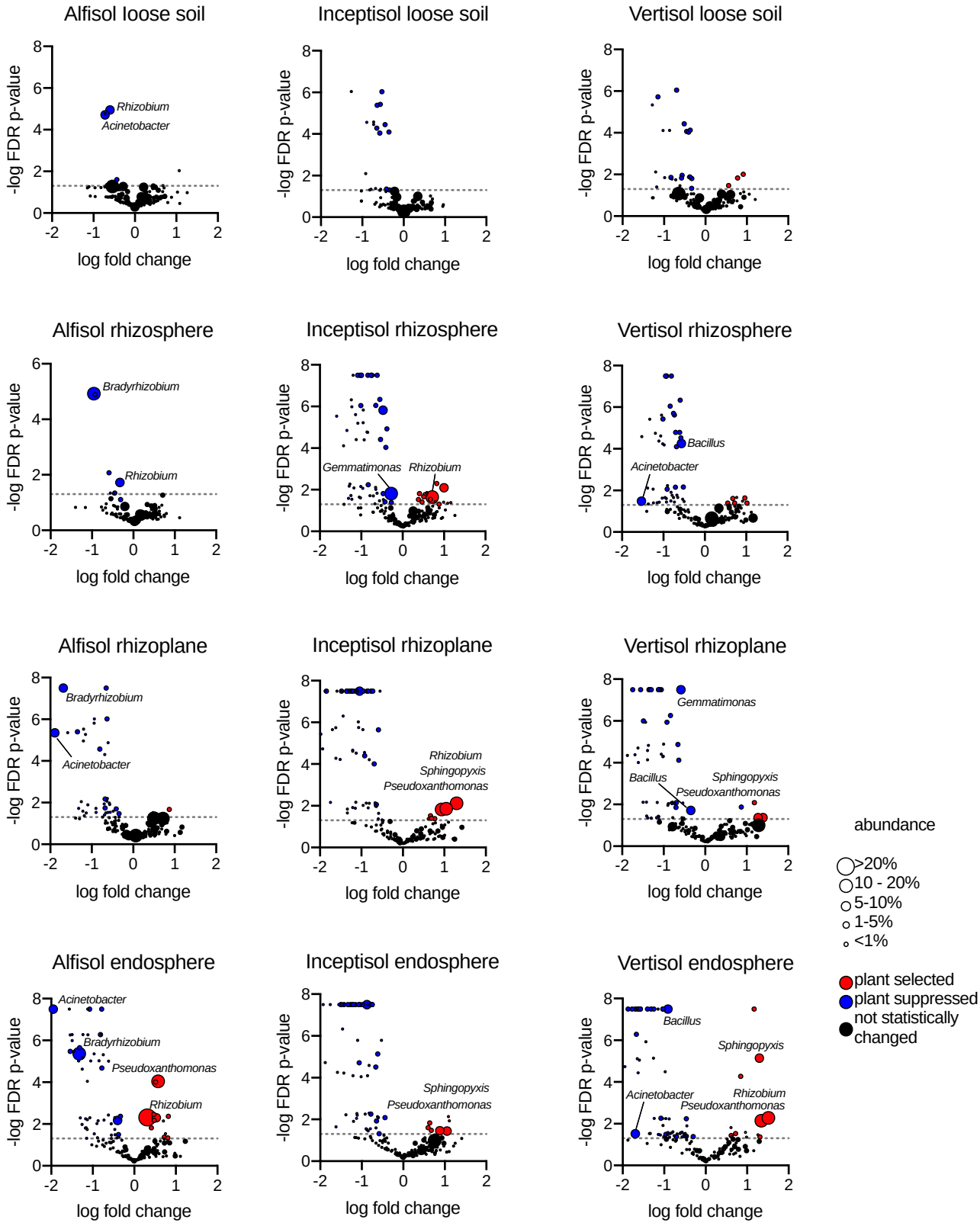

Fig S10. Volcano plots for the microbial community during plant's flowering stage. Each genus is presented as a dot where X-axis indicates the fold difference between the initial and the respective fraction (in logarithmic scale) and Y-axis indicates the significance according to *t*-test *p*-value corrected with false rate discovery and presented as -log scale (-log(0.05) = 1.301 and presented as a dotted grey line on plots). Selected genera are annotated.