Impact of plant domestication on spermosphere and rhizosphere microbiome composition

Perez Jaramillo, J.E.

Chapter 4

Deciphering the microbiome assembly of wild and modern common bean (*Phaseolus vulgaris*) grown in native and agricultural soils from Colombia

Juan E. Pérez-Jaramillo, Mattias de Hollander, Camilo A. Ramírez,

Rodrigo Mendes, Jos M. Raaijmakers and Víctor J. Carrión

Submitted for publication

Abstract

Modern crops are typically cultivated in agriculturally well-managed soils far from the centres of origin where their wild relatives thrive. How this physical decoupling impacted plant microbiome assembly is not well understood. Here, we investigated if the transition from a native to an agricultural soil affected rhizobacterial community assembly of wild and domesticated common bean (*Phaseolus vulgaris*) and if this led to a loss of rhizobacterial diversity. The impact of the bean genotype on rhizobacterial assembly was more prominent in the agricultural soil than in the native soil. Only 113 operational taxonomic units (OTUs) out of a total of 15,925 were shared by all eight bean accessions in both soils, representing 25.9% of all sequence reads. More OTUs were exclusively found in the rhizosphere of common bean in the agricultural soil as compared to the native soil, and in the rhizosphere of modern bean accessions as compared to wild accessions. Co-occurrence analyses further showed a reduction in complexity of the interactions going from native to agricultural soil. Collectively, these results suggest that habitat expansion of common bean from its native soil environment to an agricultural context had an overall positive effect on rhizobacterial diversity and led to a stronger bean genotype-dependent effect on rhizosphere microbiome assembly.

Keywords: native soil, agricultural soil, core microbiome, networks

Introduction

Plant domestication and the agricultural revolution provided a more continuous food supply to early human hunter-gatherer groups and were key drivers of the conformation of stable human settlements (Purugganan and Fuller 2009). Domestication led to major changes both in phenotypic and genotypic traits of crop varieties including larger seed size, loss of dispersal mechanisms and determinate growth (Gepts 2004; Doebley *et al*., 2006). However, domestication also led to a reduction in genetic diversity, referred to as the domestication syndrome (Doebley *et al*., 2006). Recent studies further showed that domestication affected rhizosphere microbiome assembly of several plant species, including sugar beet (Zachow *et al*., 2014), barley (Bulgarelli *et al*., 2015), sunflower (Leff *et al*., 2017), and common bean (Pérez-Jaramillo *et al*., 2017). For common bean, we previously revealed that wild accessions were enriched in bacterial taxa from the phylum Bacteroidetes, whilst modern accessions were enriched in Actinobacteria and that this compositional shift was associated with plant genotypic as well as root phenotypic traits (Pérez-Jaramillo *et al*., 2017).

Plant domestication not only comes with changes in plant traits, but is also accompanied by progressive changes in the habitat and crop management practices to promote high yields and to protect the domesticated plants from biotic and abiotic stress factors (Pérez-Jaramillo *et al*., 2016). Hence, the transition from native habitats to agricultural soils may have led to a loss of plant-associated microbes thereby affecting specific, co-evolved beneficial functions of the plant microbiome. For example, long-term nitrogen fertilization resulted in the evolution of less-mutualistic rhizobia, providing fewer benefits to the host (Weese *et al*., 2015). Similarly, it was shown that nitrogen amendments suppressed soil respiration and microbial biomass, promoting copiotrophs such as Actinobacteria and Firmicutes while reducing the abundance of oligotrophs such as Acidobacteria and Verrucomicrobia (Ramirez *et al*., 2012). It has been also shown that

the occurrence of members of the phylum Bacteroidetes, whose members are known for their abilities to degrade complex polymeric organic matter, is negatively affected by agricultural soil management practices (Wolińska *et al*., 2017). Similarly, conversion of the Amazon rainforest to agriculture resulted in biotic homogenization of soil bacterial communities and a net loss of microbial diversity (Rodrigues *et al*., 2013). For most crop plants, however, there is little knowledge on the co-evolutionary trajectory between plants and their microbiomes during habitat expansion and if domestication indeed led to a reduced microbial diversity and a loss of specific microbial genera in these new habitats. In this study, we used *Phaseolus vulgaris* (common bean) as 'model' plant species. Common bean originated in central Mexico and as a wild species spread throughout Central and South America (Gepts 1998; Bitocchi *et al*., 2012; Desiderio *et al*., 2013). Geographical isolation of wild common bean resulted in the establishment of the Mesoamerican and Andean genetic pools (Gepts and Bliss 1985) which were the basis of two independent domestication processes. As a consequence, domesticated common bean underwent several morphological and physiological changes as well as a significant reduction in genetic diversity (Gepts and Debouck 1991; Sonnante *et al*., 1994; Chacón *et al*., 2005). We selected wild, landraces and modern accessions of Mesoamerican common bean originating from Colombia based on a number of traits (Pérez-Jaramillo *et al*., 2017). Here, we hypothesized that the transition from a native soil environment to an agriculturally managed soil led to a loss of bacterial diversity in the plant rhizosphere microbiome. Therefore, the eight bean accessions were grown in a native soil and in an agricultural soil, both collected in the Colombian highlands, followed by rhizobacterial community profiling, species abundance modelling and co-occurrence network analyses.

Material and Methods

Selection of soils and plant accessions

Two types of soil were selected for this study in the north-west region of Colombia. The native soil was collected in a rural area near to the municipality of Angostura (Antioquia, 6° 53′ 7″ N, 75° 20′ 7″ W). This region has the same climatic conditions, altitude and local plant diversity that have been reported for wild common bean populations collected in Colombia (Toro *et al*., 1990). A successional forest was identified in the region and soil samples were taken from the top layer (0-20cm) after cleaning the litter, wood and unwanted material. Several landraces typically associated with Mesoamerican traits were collected near this region; therefore we refer to this soil as "native". The agricultural soil was collected in a common bean producing farm in rural area of the municipality of El Carmen de Viboral (6°4′55″ N, 75º20′3″W). This soil was under cultivation for the last 30 years in a crop rotation system composed of maize, common bean and potato. Tillage, liming, chemical fertilization (N-P-K) and the application of poultry waste are typical agricultural practices in the region. The climatic conditions, the altitude and the local plant diversity in this region are not suitable for the growth of Colombian Mesoamerican wild common bean populations, but are optimal for Andean domesticated common bean varieties. Physicochemical analyses were performed in the Soil Science Laboratory from the National University of Colombia in Medellín, using standard procedures (Table S1). Two wild, three landraces and three improved varieties (cultivars) of common bean (*Phaseolus vulgaris*) were selected according to the following characteristics: belong to the Colombian Mesoamerican genetic pool, same race, similar phaseolin type, same altitudinal range, adapted to the same climatic conditions and same growth type. The seeds were kindly provided by the Genetic Resources Program at the International Centre for Tropical Agriculture - CIAT – in Palmira, Colombia. A genotypic analysis was performed on the common bean accessions to validate the domestication status provided

by the original passport (Pérez-Jaramillo *et al*., 2017). As a result, we reclassified the accessions as two wild (A1, A2), one landrace (L1), and five modern accessions (M1 to M5).

Experimental Design

Seeds of the eight bean accessions were surface-sterilized twice with sodium hypochlorite (0.5%) during three minutes and rinsed in sterile water four times. 100 µL of the last rinsing step was cultured in Tryptic Soy Agar (TSA, Oxoid) and in Potato Dextrose Agar (PDA, Difco) media by triplicate in order to check the growth of bacteria and fungi, respectively. Disinfected seeds were germinated on filter paper with sterile tap water; after two to five days all the seeds had germinated. The native and agricultural soils were air dried, passed through a 2-mm-mesh sieve and distributed into 1 L PVC pots, with 700 g of dried soil per pot. Seedlings were transferred to the pots, with one plant per pot and four replicates per bean accession and per soil. The plants were cultivated in a growth chamber for one week and then arranged randomly in a screenhouse with an average temperature of 25ºC, 12h of daylight, and daily watered with tap water. Four pots with native soil and four pots with agricultural soil, both without plants were used as bulk soils.

Sampling of rhizospheric soil

At flowering stage, the plants were carefully removed from the pots keeping the root system intact. Soil loosely adhered to the roots was removed by vigorous shaking, and when no more soil could be removed, the root system was submerged in tubes with 5mL of LifeGuard Soil Preservation Solution (Mo Bio Laboratories, Carlsbad, CA, USA) and vigorously shaken in order to wash the roots and recover around 1g of rhizospheric soil per sample for total DNA isolation. For the bulk soils, approximately 1 g of soil was collected from each control pot and also submerged in 5 mL of the LifeGuard solution.

The LifeGuard Soil Preservation Solution can prevent microbial growth while maintaining nucleic acid integrity. All samples were kept at -20ºC until further use.

DNA isolation

For each plant accession in each soil, four replicates of rhizospheric soil were used for total DNA extraction as well as four replicates of control soil. To obtain the total DNA, a two-step approach was followed in order to recover RNA and DNA from the same sample. First, RNA was isolated using the RNA PowerSoil Isolation Kit (Mo Bio Laboratories, Carlsbad, CA, USA) according to the manufacturer's instructions, with slight modifications as follows. After adding the phenol:chloroform:isoamyl alcohol solution to the bead tube containing the bead solution, solution SR1 and SR2, as well as the soil sample, the agitation step was applied for 40 min. This modification allowed us to increase the RNA yields. The RNA was stored at -80ºC for further use. For DNA isolation, the RNA PowerSoil® DNA Elution Accessory Kit (Mo Bio Laboratories, Carlsbad, CA, USA) was used. Briefly, after elution of the RNA from the RNA capture column, this column was transferred to another tube and the DNA elution procedure was performed according to manufacturer's instructions. Each obtained DNA sample was then cleaned with the PowerClean® DNA Clean-Up Kit (Mo Bio Laboratories, Carlsbad, CA, USA). Agarose gel electrophoresis and a ND1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) were used to control RNA and DNA yield and quality. DNA samples were stored at -80° C until further use.

16S amplicon sequencing and preparation of the OTU table

Total DNA was used to amplify and sequence the V3-V4 region of the 16S rRNA that creates an amplicon of approximately ~460bp. The library preparation and the sequencing protocol were done according to the general guidelines of Illumina MiSeq Reagent Kit

v3 (Illumina 2013). The RDP extension to PANDASeq (Masella *et al*., 2012) named Assembler (Cole *et al*., 2014) was used to merge paired-end reads with a minimum overlap of 10 bp and at least a PHRED score of 25. Primer sequences were removed from the per sample FASTQ files using Flexbar version 2.5 (Dodt *et al*., 2012). Sequences were converted to FASTA format and concatenated into a single file. All reads were clustered into OTUs using the UPARSE strategy by dereplication, sorting by abundance with at least two sequences and clustering using the UCLUST smallmem algorithm (Edgar, 2010). These steps were performed with VSEARCH version 1.0.10 (Rognes *et al*., 2015), which is an open-source and 64-bit multithreaded compatible alternative to USEARCH. Next, chimeric sequences were detected using the UCHIME algorithm (Edgar *et al*., 2011) implemented in VSEARCH. All reads before the dereplication step were mapped to OTUs using the usearch_global method implemented in VSEARCH to create an otu table and converted to BIOM-Format 1.3.1 (McDonald *et al*., 2012). Finally, taxonomic information for each OTU was added to the BIOM file by using the RDP Classifier version 2.10 (Cole *et al*., 2014). All steps were implemented in a Snakemake workflow (Köster and Rahmann 2012). The sequence data are deposited at the European Nucleotide Archive (ENA) under accession number PRJEB26084.

Diversity and abundance analysis

For downstream analysis we took the obtained OTU table and prepared a "filtered table" using QIIME (1.9.1) custom scripts (Kuczynski *et al*., 2012). First, we extracted from the OTU table the Bacteria domain using the command *split_otu_table_by_taxonomy.py*. Next, we discarded singletons, doubletons, Chloroplast and Mitochondria sequences using the command *filter_otus_from_otu_table.py*. With the "filtered_OTUtable", we first calculated the alpha diversity. Using the command *alpha_rarefaction.py*, the OTU table was rarefied to counts up to 35,000 reads. The reason to use this value was because

this was the lowest sequencing depth obtained from a sample. To calculate the diversity indexes, we used the *alpha_diversity.py* and *alpha_rarefaction* commands, obtaining Shannon, Observed OTUs, Chao1 and Faith's Phylogenetic Diversity metrics. One-way ANOVA and Tukey HSD, as well as statistical tests to validate ANOVA assumptions were performed in R (3.4.1) (R Core Team, 2015). For the Beta diversity, the unrarefied "filtered_OTUtable" was first normalized using the R package metagenomeSeq (v.1.12) (Paulson *et al*., 2016). We used a cumulative-sum scaling (CSS) method to avoid the biases generated with current sequencing technologies due to uneven sequencing depth (Paulson *et al*., 2013). With the normalized OTU table we calculated Bray-Curtis and weighted and unweighted Unifrac dissimilarity matrices and use it to build Principal Coordinate with Phyloseq package (v.1.10) (McMurdie and Holmes, 2013). The nonparametric *adonis* test was used to assess the percentage of variation explained by the soil type along with its statistical significance using Vegan (v.2.4-0) package (Oksanen *et al*. 2016), all performed in R. For the differential abundance analysis and the construction of the heat maps the STAMP software (v.2.1.3) was used (Parks *et al*., 2014). Rarefied OTU tables from Agricultural and Native rhizosphere and bulk soil data (35,000 reads per sample) were used for pairwise comparisons. Welch's t-tests followed by Bonferroni corrections were performed at Phylum and at Family level between soils. Dendograms were built in STAMP with an average neighbor method (UPGMA) and the rows included all the bacterial phyla observed in rhizosphere and bulk soil samples along with their relative abundance.

In order to compare the number of shared and exclusive rhizobacterial genera between common bean accessions in agricultural and native soil, we selected wild accessions A1 and A2 and modern bean accessions M3 and M4. These 4 accessions were selected in order to normalize the number of samples as well as the number of reads to compare. Regarding the modern accessions, M3 and M4 showed to be the most modern accessions

available both for native and agricultural soils (Pérez-Jaramillo *et al*., 2017). To depict the taxa exclusively found in a particular soil or accession we used the online tool Venny (2.1) (Oliveros, 2015) and to graphically represent the exclusive genera we built euler diagrams using the shiny app eulerr of the homonymous R package (Larsson, 2017). Euler diagrams are area-proportional generalized venn diagrams for which the requirement that all intersections be present is relaxed. Euler diagrams were built for the exclusive and shared genera per soil type and plant domestication status.

In order to have a better understanding of the composition of the bacterial diversity in the rhizosphere of the common bean accessions, we calculated several species abundance distribution models and determined whether neutral or niche-based mechanisms were governing the bacterial assembly. We hypothesized that the agricultural soil would be driven by neutral-based processes and that the native soil would respond to niche-based process. We used the command Radfit from the R package vegan to evaluate several abundance models and a zero-sum multinomial (ZSM) model. Species abundance distributions models and comparison of the models fit based on the Akaike Information Criterion (AIC) were calculated as previously reported (Pérez-Jaramillo *et al*., 2017).

Core microbiome and co-occurrence network analyses

For the core microbiome analyses, rarefied OTU tables (35,000 reads each sample) were used for both soils. The QIIME command *compute_core_microbiome.py* was applied in order to obtain a list of OTUs observed in all the common bean rhizosphere samples regardless of soil type. Core microbiome analyses were also performed for common bean on each soil type. Only core OTUs with a relative abundance > 0.5% were included for graphical purposes. Pie and donut charts were built in R. Network analysis was performed to assess the complexity of the interactions among microbial taxa in the common bean rhizosphere grown in the agricultural soil (n=32) and in the native soil (n=26). Best

practices for co-occurrence network construction were strictly followed. Rarefied OTU tables were filtered to a minimum threshold of 25 sequences per OTU. Non-random co-occurrence analyses were performed using SparCC (Friedman and Alm, 2012). *P*-values were obtained by 99 permutations of random selections of the data table. SparCC correlations with a magnitude > 0.8 or < -0.8 and statistically significant (P<0.01) were further included into network construction. The nodes in the reconstructed networks represent the OTUs at 97% identity, whereas the edges correspond to a strong and significant correlation between nodes. The topology of the network was inferred on a set of measures (number of nodes and edges, modularity, number of communities, average path length, network diameter, averaged degree and clustering coefficient) which were calculated using Gephi (v.0.9.2) (Bastian *et al*., 2009). Network visualizations were constructed using Cytoscape (v. 3.4.0) (Shannon *et al*., 2003). Clusters were calculated using a hierarchical clustering algorithm (HC-PIN) with the Cytoscape plugin Cytocluster (Li *et al*., 2017).

Results

Soil and genotype influence plant growth and development

Eight accessions of common bean, encompassing wild relatives, landraces and modern cultivars were grown in agricultural and native soils collected from the Colombian highlands, at the same time under the same screenhouse conditions (see Material and Methods). The soils differed in several physicochemical characteristics (Table 1). Briefly, the agricultural soil had more organic matter, available phosphorus and calcium as well as higher pH and cation exchange capacity (CEC) as compared to the native soil. The native soil showed higher concentrations of iron and aluminum. At flowering stage, each plant was harvested to collect rhizospheric soil and to assess several plant phenotypic traits. In the agricultural soil, significant differences were observed between the different

bean accessions in root dry weight and the number of days to reach flowering stage (Table 2). Genotype-dependent differences were also observed in the native soil (Table 2). Several replicates of the accession M5 did not grow in the native soil or showed a very poor development, and were therefore not included in further analyses. In general, the mean root dry weight was higher for bean accessions grown in the agricultural soil than in the native soil (Fig. S1a). Finally, the number of days to reach the flowering stage (R6) was higher in the native soil than in the agricultural soil, indicating a delayed development of the bean accessions in the native soil (Fig. S1b).

Table 1. Soil physical and chemical analyses. Physical and chemical characterization of the agricultural and the native soil used in this study to grow the common bean plants for the rhizobacterial characterization.

Unit	Item		Agricultural	Native
%	Texture	Clay	8	24
		Silt	30	12
		Sand	62	64
		Classification	Clay loam	Sandy clay loam
dSm-1	pH		5.8	4.7
%	Organic matter		17.9	11.6
cmolc kg-1	Al		ND	3.00
	Ca		15.0	0.3
	Mg		2.9	1.0
	K		0.6	0.4
	Na		0.03	0.03
	CEC		18.50	4.80
mg kg-1	P		56	3
	S		8	8
	Fe		50	620
	Mn		2	2
	Cu		3	3
	Zn		7	4
	B		0.24	0.21
Methods:				

Texture: Bouyoucos; pH: Water (1:1); Organic matter: Walkley Black; Al: KCl 1M; Ca, Mg, K, Na: Ammonium acetate 1M; CEC: Cation exchange sum; S: monocalcium phosphate 0.008M; Fe, Mn, Cu, Zn: Olsen-EDTA; B: Hot water; P: Bray II.
Units consideration:
ND: Not detectable; dSm-1 = mmho cm-1; cmolc kg-1 = meq/100 g soil; mg kg-1 = ppm

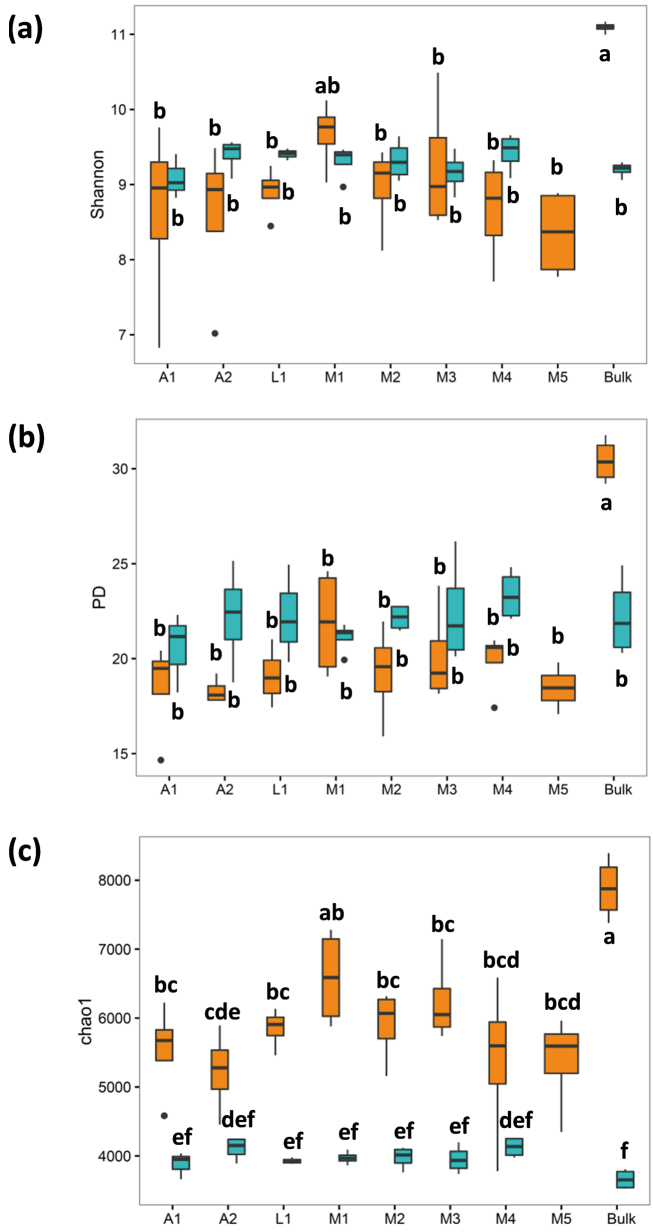

Fig. 1. Comparative analysis of the Alpha diversity of 16S rRNA rhizobacterial sequences from common bean accessions in agricultural and native soils. (a) Shannon, (b) Phylogenetic diversity, and (c) Chao 1 were calculated by soil type and for all bean accessions and the bulk soils. The data was rarefied up to 35,000 counts per sample. Statistically significant differences were determined by one-way ANOVA (P<0.05) followed by post hoc Tukey test. Cyan color was assigned to native soil and dark orange to agricultural samples.

Diversity of rhizobacterial communities is driven by soil type and rhizosphere

For bulk soil and rhizosphere samples, 4.24 million reads were recovered after quality filtering (Table S1), representing 16.727 bacterial operational taxonomic units (OTUs) at 97% sequence similarity. For the α-diversity, rarefaction curves were obtained for Chao1, observed OTUs and Phylogenetic Diversity (PD) indices (Fig. S2). It was observed that the evenness, represented by the Shannon index and the phylogenetic diversity (PD), were in general similar between rhizosphere samples in both soil types (Figs. 1a and b), while the bacterial species richness was significantly higher in the agricultural soil than in the native soil (Fig. 1c). In the agricultural soil, all diversity indices were significantly higher for the bulk soil than for the rhizosphere samples. Additionally, it was observed that agricultural bulk soil samples showed significantly higher values for all the diversity indices as compared to the native bulk soil samples. Regarding the β-diversity, Bray-Curtis metrics and Principal Coordinate Analysis (PCoA) revealed a significant effect of the soil type (Fig. 2a). Soil type alone explained 71.3% of the total variability in the bacterial community composition (PERMANOVA, $P<0.001$). Subsequently, the samples were divided by soil type and analyzed separately. For the total variability of the rhizobacterial community structure, the bean genotype explained 31.2% in the agricultural soil (PERMANOVA, $P<0.001$) and 28.3% in the native soil (PERMANOVA, $P<0.05$) (Figs. 2b and c). Unifrac metrics confirmed the results observed with Bray-Curtis metrics (Fig. S3). The mean Bray-Curtis distances showed that the variability of the rhizobacterial communities within samples of the same accession was significantly lower as compared to the distance between the bean accessions (Agricultural soil, $P<0.001$; Native soil, $P<0.05$; t-test, Bonferroni-corrected). For the agricultural soil, the rhizobacterial community composition of wild bean accession A1 was similar to that of wild accession A2, but significantly different from that of the landrace and the five modern bean accessions ($P<0.001$; t-test, Bonferroni-corrected). In

the native soil, however, rhizobacterial community composition did not differ significantly between the wild and modern bean accessions.

Table 2. Phenological traits of the common bean accessions grown in native and agricultural soils. Data for root dry weight and the number of days to reach the flowering stage is shown.

	Root dry weight (g)		Days to flower	
Accession	**Native**	**Agricultural**	**Native**	**Agricultural**
A1	0.082±0.015 bcd	0.438±0.141 ab	111.0±12.3 a	93.5±12.4 a
A2	0.139±0.010 a	0.484±0.149 a	103.7±11.4 ab	83.5±5.7 ab
L1	0.118±0.009 ab	0.282±0.080 bc	81.0±31.0 bc	62.5±10.8 c
M1	0.112±0.056 abc	0.219±0.053 c	90.3±17.0 abc	64.8±16.7 bc
M2	0.134±0.039 a	0.374±0.215 abc	75.0±4.0 c	66.5±21.2 bc
M3	0.055±0.016 d	0.260±0.127 c	87.6±11.4 abc	64.8±16.7 bc
M4	0.067±0.019 cd	0.204±0.046 c	106.6±3.5 ab	56±5.2 c
M5	ND*	0.332±0.048 abc	ND	56±5.2 c

The mean values of 4 replicates (agricultural) and 3 replicates (native) per accession are shown, followed by the standard deviation of the mean. The harsh conditions of the native soil prevented us to have 4 replicates for all the accessions; therefore, it was decided to normalize the number of replicates to 3 in native soil samples. Statistical analysis of root dry weight and days to flowering were done between bean accessions per soil type. ANOVA and LSD (P<0.05) tests were applied after checking for assumptions of normality and homoscedasticity. Accessions with the same letter are not significantly different.
*Accession M5 did not grow on the native soil.

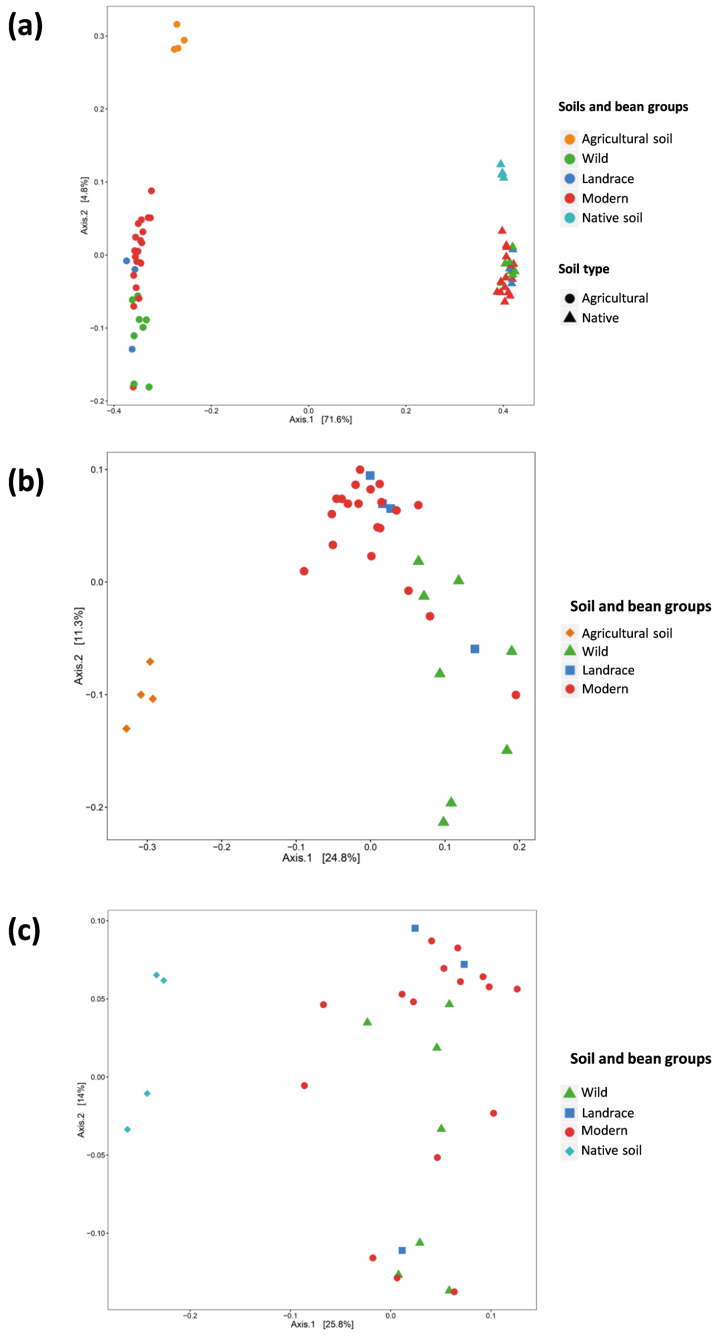

Fig. 2. Rhizosphere bacterial community structure in agricultural and native soils. Principal Coordinate Analysis (PCoA) of 16S rRNA diversity in the rhizosphere of the eight common bean accessions used in this study. (a) Rhizosphere bacterial community of common bean grown in agricultural (circles) and native (triangles) soils. Colors represent the stage of domestication and bacterial communities

from agricultural and native bulk soils. (b) PCoA including only rhizosphere bacterial communities of common bean plants grown in agricultural rhizosphere and bulk soil samples. Colors and shapes represent the stage of domestication and bacterial communities from agricultural and native bulk soils. (c) PCoA including only rhizosphere bacterial communities of bean plants grown in native rhizosphere and bulk soil samples. Colors and shapes represent the stage of domestication and bacterial communities from agricultural and native bulk soils. CSS transformed reads were used to calculate Bray-Curtis distances in (a), (b) and (c).

Profound impact of soil microbiome in the rhizosphere of common bean grown in native and agricultural soils

The observed differences in α and β diversity between the native and agricultural soils and between the eight bean accessions led us to explore more in depth the differences in taxonomic identity and relative abundance of the bacterial taxa for each soil. The most abundant bacterial phyla were Proteobacteria, Acidobacteria and Bacteroidetes in both soils. In the native soil, however, the phylum Acidobacteria showed a higher relative abundance than in the agricultural soil (Figs. S4 and S5). At phylum level, Acidobacteria and Verrucomicrobia were significantly more abundant in the native bulk soil than in the agricultural bulk soil (Welch's t-test, $P<0.05$, Bonferroni-corrected) (Fig. S6a). At class level, Acidobacterial subgroups 1, 2 and 3 were enriched in the native soil, while Acidobacteria subgroup 4 and Betaproteobacteria were more abundant in the agricultural soil (Welch's t-test, $P<0.05$, Bonferroni-corrected) (Fig. S6b). Proteobacteria and Bacteroidetes were consistently more abundant in the rhizosphere of common bean, regardless of the soil type, while Acidobacteria and Verrucomicrobia showed a consistent decrease in the rhizosphere compared to their abundance in bulk soil (Figs. 3a and b). Actinobacteria was significantly more abundant in the rhizosphere of common bean grown in the agricultural soil as compared to the bulk soil (Fig. 3a), whereas this rhizosphere effect was not observed in the native soil (Fig. 3b) (Welch's t-test, $P<0.05$, Bonferroni-corrected).

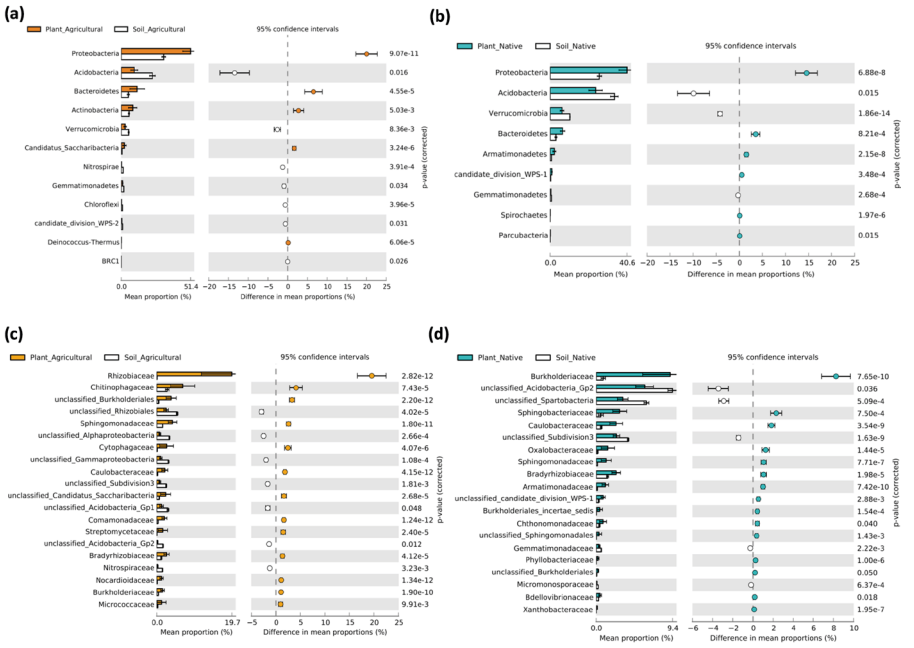

Fig. 3. Differential abundance of bacterial OTUs in agricultural and native soils. Welch's t-tests followed by Bonferroni corrections were performed between merged rhizosphere samples and merged bulk soil samples from agricultural soil and native soil at Phylum (a and c) and Class (b and d) levels. Only differentially abundant Phyla and Classes are shown.

Among the Actinobacteria enriched in the rhizosphere of all bean accessions grown in the agricultural soil, *Streptomycetaceae* and *Nocardioidaceae* were the most abundant families together with *Rhizobiaceae*, *Sphingomonadaceae, Caulobacteraceae* and *Comamonadaceae* for the Proteobacteria, and *Chitinophagaceae* and *Cytophagaceae* for the Bacteroidetes (Fig. 3c). The smaller yet significant rhizosphere effect observed for the bean accessions grown in the native soil was explained by higher relative abundances of *Burkholderiaceae*, *Caulobacteraceae*, *Oxalobacteraceae, Sphingomonadaceae* and *Bradyrhizobiaceae* for the Proteobacteria and *Sphingobacteriaceae* for the Bacteroidetes (Fig. 3d). To further dissect these differences in microbiome composition between rhizosphere and bulk soils, the abundance of the read counts was fitted to several species

abundance distribution (SAD) models. Comparison of Akaike's Information Criterion (AIC) allowed us to find the best-fit value from six distribution models. The results showed that the OTU abundance distributions in the rhizosphere of the bean accessions grown in agricultural and native soils, and the respective bulk soils, are explained by niche-based distributions (Pérez-Jaramillo *et al*. 2017) (Table S2).

Higher bacterial diversity recruited by common bean grown in agricultural soil

We performed a comparison of the bean rhizobacterial community at genus level between soil types and between wild and modern accessions in order describe how habitat expansion and agricultural soil management could have depleted or enriched rhizosphere bacterial genera. The results showed that 143 rhizobacterial genera, representing 28.7% of the total number of genera, were exclusively represented in the agricultural soil (Fig. S7a). Exclusive genera accounted for 2.3% of the total relative abundance in agricultural soil. Some of these 'exclusive' genera such as *Lysobacter* and *Aeromicrobium* accounted for 0.6% and 0.4% of the total relative abundance, respectively. Thirty-one genera including *Cytophaga* and *Acidicapsa* were exclusively found in the native soil, representing 6.2% of the total number of bacterial genera (Fig. S7a). Exclusive genera accounted for 0.2% of the total relative abundance in native soil.

Two wild accessions (A1 and A2) and two modern bean accessions (M3 and M4) were selected in order to compare the number of shared and exclusive bacterial genera in the rhizosphere. We found that in the agricultural soil, 85.9% of the rhizobacterial genera were shared between wild and modern accessions, 8.7% was exclusively found in the rhizosphere of the modern bean accessions and 5.4% was exclusively found in the rhizosphere of the wild accessions (Fig. S7b). In the native soil, a similar trend was observed, with 84.8% of the rhizobacterial genera shared between wild and modern bean accessions, 9.0% exclusively found in the rhizosphere of the two modern accessions and

6.3% in the two wild accessions (Fig. S7c). In conclusion, we found more bacterial genera in the rhizosphere of the eight bean accessions grown in the agricultural soil than in the native soil. Additionally, we found more bacterial genera in the rhizosphere of the modern bean accessions than in wild accessions irrespective of the soil type. It should be noted that the abundance of these 'exclusive' bacterial genera in the common bean rhizosphere was relatively low for both soils.

The core microbiome of common bean was represented by a small subset of microorganisms

From the total of 16,727 clustered OTUs, we found 113 OTUs consistently present in the rhizosphere of all eight bean accessions grown in the native and agricultural soils. These 113 OTUs, classified up to genus level, represented only 0.67% of the total number of OTUs but 25.9% of all the sequence reads. This core bean rhizosphere microbiome consisted of 61 Proteobacterial OTUs that made up 68.8% of the mean relative abundance with the genus *Rhizobium* as the most abundant contributor (2 OTUs, 33.4%), followed by *Bradyrhizobium* (2 OTUs, 6.7%), *Burkholderia* (3 OTUs, 4.9%), *Novosphingobium* (3 OTUs, 3.0%) and *Sphingomonas* (1 OTU, 2.2%) (Fig. 4). Other Phyla represented in the core rhizosphere microbiome were Acidobacteria (27 OTUs, 12.2% relative abundance), Actinobacteria (6 OTUs, 4.1%), Verrucomicrobia (8 OTUs, 2.5%) and Planctomycetes (5 OTUs, 1.1%). A core microbiome analysis was done also per soil type in order to dissect the specific contribution of each habitat to the overall core. For the agricultural soil, the core rhizosphere microbiome was composed of 611 OTUs representing 4.97% of the total number of OTUs and 33.07% of the reads. Proteobacteria (219 OTUs), Bacteroidetes (62 OTUs) and Actinobacteria (58 OTUs) were the three most abundant phyla within the core (Fig. S8) with again *Rhizobium* as the most abundant genus (26.7%) followed by *Dyadobacter* (3.3%) and *Streptomyces* (2.1%). In the native

soil, the core rhizosphere microbiome was composed of 812 OTUs representing 12.6 % of the total number of OTUs and 46.4% of the reads. Proteobacteria (237 OTUs), Acidobacteria (190 OTUs), Verrucomicrobia (68), Bacteroidetes (53 OTUs), Actinobacteria (48) and Chloroflexi (17 OTUs) were the most abundant phyla (Fig. S9). Within Proteobacteria, *Ralstonia* was the most abundant genus (4.6%) followed by *Burkholderia* (4.0%), *Herbaspirillum* (1.6%) and *Rhizobium* (1.2%). In the core rhizosphere microbiome of beans grown in the native soil, Acidobacteria was mainly represented by the Acidobacteria subgroups 1, 2 and 3, with 27.4% of the reads. In fact, less than 3% of the OTUs classified in the core as Acidobacteria summed up 12.7% of the total number of reads, evidencing the strong dominance of this phylum in the native soil habitat. Verrucomicrobia represented in total 6.6% of the core microbiome with most of the reads assigned as *incerta_sedis*. Finally, the genus *Mucilaginibacter* and unclassified members of the *Chitinophagaceae* family accounted for most of the abundance of Bacteroidetes representing 5.4% of the core; for Actinobacteria, small contributions mostly by unclassified Acidimicrobiales and Actinomycetales collectively accounted for a relative abundance of 3.1%. In conclusion, these comparative analyses indicated that only a small number of 113 OTUs was consistently present in the rhizosphere of all eight bean accessions grown in the agricultural and native soils, and also revealed that these OTUs represent on average more than a quarter (25.9%) of the total of 4.2 million sequence reads.

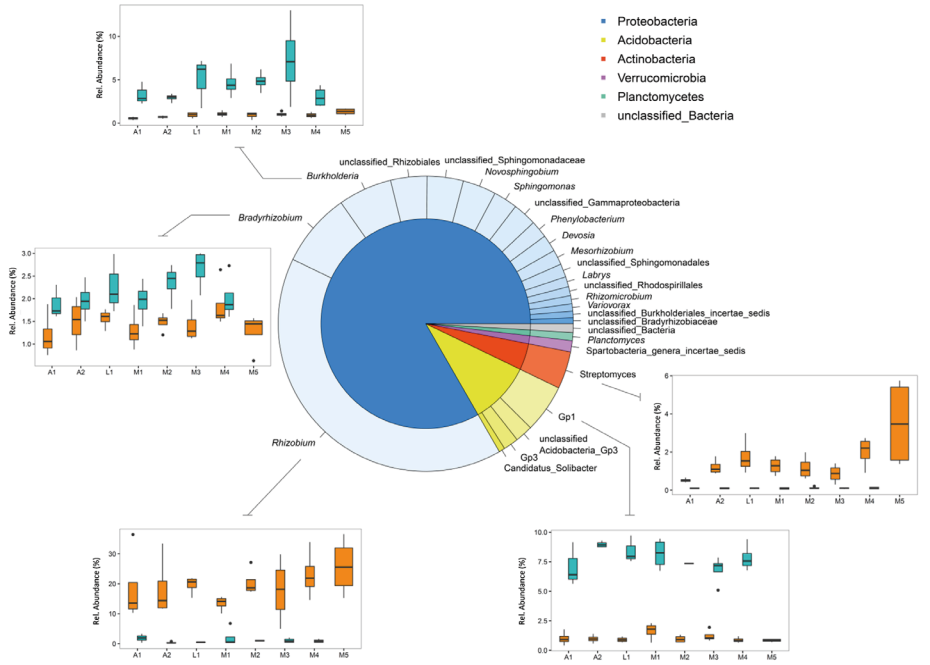

Fig. 4. Core microbiome of the rhizosphere of common bean. The different portions within the inner pie chart represent the bacterial phyla that are part of the common bean core microbiome. The outer donut plot represents the genera that are part of the core, and each genus assigned to the phylum they belong to. The size of the different pie and donut portions represents the contribution of each phylum/genus to the total relative abundance. Satellite box plots depict the relative abundance of selected genera by bean accession (A1 and A2, wild; L1, Landrace; M1 to M5, Modern) and by soil type. Cyan and dark orange colors were assigned to native soil and agricultural samples, respectively.

Higher network complexity in native soil based on co-ocurrence

Co-occurrence network analyses were performed to assess the complexity of the interactions between the rhizobacterial taxa detected in the rhizosphere of common bean grown in native and agricultural soils. The correlations between the occurrence of the rhizobacterial genera were calculated using SparCC (Friedman and Alm, 2012) followed by the graphical inference of the network and the estimation of several topological properties (Table S3). The rhizobacterial network in agricultural soil consisted of 63 nodes and 61 significant correlations, with only one negative connection between OTUs

identified as *Lysobacter* and *Ohtaekwangia* (Fig. 5a). In general, this network presented a simple structure, with four main clusters and few OTUs per cluster. For the native soil, the obtained network contained 89 nodes and 176 significant correlations, with 158 positive and 18 negative (Fig. 5b). Three main clusters were identified, with a high number of nodes per cluster and with a high number of interconnections within each cluster. Remarkably, cluster 2 was connected to the other two clusters only through negative correlations. Global networks statistics were determined for both networks (Table S3). Briefly, modularity and the number of communities were higher in the agricultural soil than in the native soil. Conversely, the average path length and the average degree were higher in the native soil. Using Betweenness Centrality (BC), we aimed to find keystone species within each network. In the agricultural soil, the highest BC values were found for the genera *Lysobacter* (OTU_136), *Rhizobium* (OTU_1), *Niastella* (OTU_10281, OTU_44 and OTU_56), *Ohtaekwangia* (OTU_69), *Terrabacter* (OTU_46), and *Arthrobacter* (OTU_886). For the native soil, the highest BC values were found for *Aquisphaera* (OTU_537, Planctomycetes phylum), two unclassified_Acidobacteria (OTU_62 and OTU_12725), an unclassified_Acetobacteraceae (OTU_175) and *Burkholderia* (OTU_45).

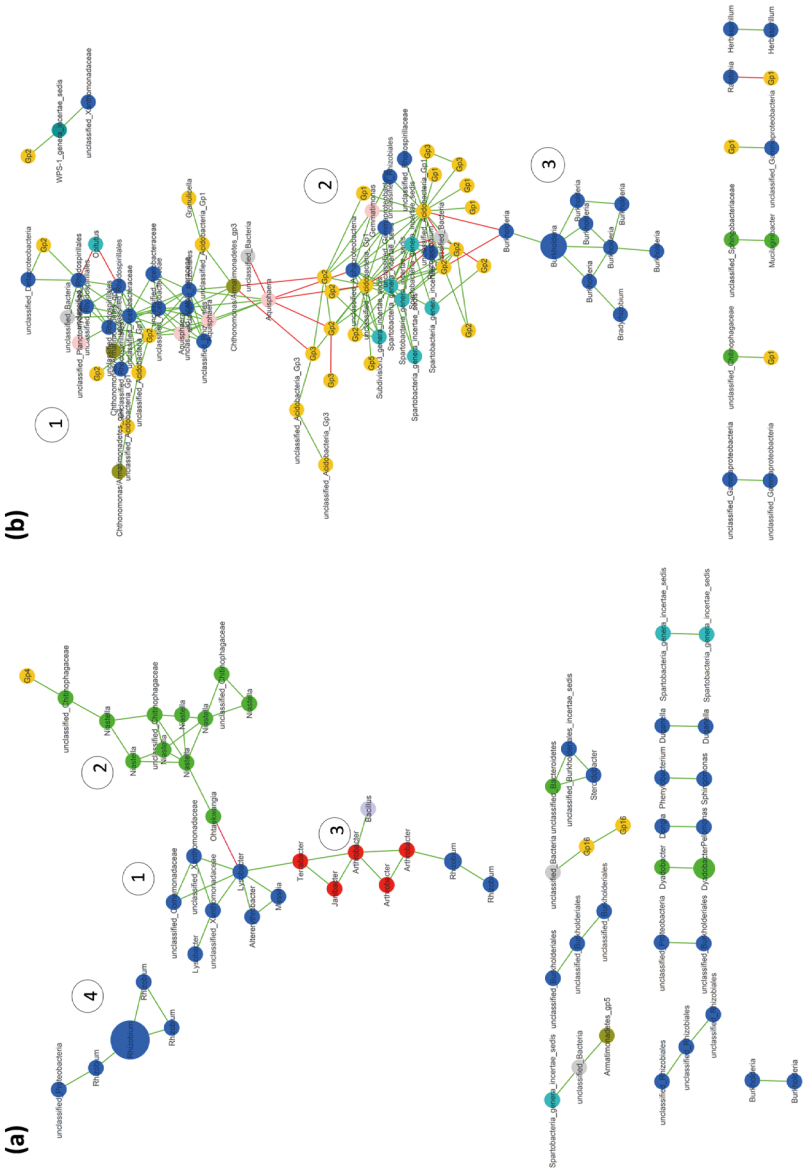

Fig. 5. Common bean rhizobacterial co-occurrence networks in agricultural and native soils. (a) Co-occurrence network of common bean rhizosphere samples in agricultural soil. Cluster 1 was composed of bacterial taxa from several classes of the Proteobacteria phylum. Cluster 2 contained exclusively bacterial taxa from the *Chitinophagaceae* family. Cluster 3 included actinobacterial taxa and one Bacillus, and cluster 4 was composed of the genus *Rhizobium*. (b) Co-occurrence network of common bean rhizosphere samples in native soil. From the three main clusters identified, two were highly abundant in nodes from the Proteobacteria phylum (1 and 3) which held negative connections to cluster 2, mainly composed of phyla Acidobacteria and Verrucomicrobia. Positive interactions are depicted as green edges and the negative interactions are depicted as red edges.

Discussion

In this study we showed that rhizobacterial diversity of both wild and domesticated common bean (*Phaseolus vulgaris*) was higher in an agricultural soil as compared to a native soil. Furthermore, species abundance analyses revealed niche-based processes for both soils suggesting selection pressures. For the agricultural soil, management practices (fertilization, addition of organic matter) are the most likely drivers of the observed differences in species abundance distributions, whereas edaphic factors, in particular low pH, are the most probable selection pressures for the native soil. Bacterial diversity is generally lower in acidic soils (Fierer and Jackson, 2006; Lauber *et al*., 2009) and pH largely determines the composition of the soil bacterial communities (Kuramae *et al*., 2012). Our results further showed that the impact of the bean genotype on rhizobacterial assembly was more prominent in the agricultural soil than in the native soil where the rhizosphere effect was much smaller and where genotype-dependent effects on rhizobacterial community composition were more homogeneous. An underlying mechanism of this minor and more homogenous rhizosphere effect is that the harsh abiotic conditions in the native soil may have affected the amount and quality of root exudates released into the soil. In the native soil used in this study, the bean plants faced a low soil pH, high aluminum concentrations and low P availability, characteristics that are common for tropical undisturbed soils (Rodrigues *et al*., 2013; Sánchez and Logan, 2013). Also, the lack of nodulation in these acidic conditions (Ferguson *et al*., 2013) could have undermined symbiotic associations for nitrogen uptake and concomitantly the growth and development of the common bean plants with an adverse effect on root exudation.

Common bean grown in the agricultural soil harbored more exclusive OTUs than bean grown in the native soil and we also found more exclusive OTUs in the rhizosphere of modern bean accessions as compared to wild accessions, irrespective of the soil type. The

genera exclusive for the agricultural soil were *Lysobacter* and *Aeromicrobium*. The genus *Lysobacter* is commonly found in agricultural soils (Puopolo *et al*., 2018) and their abundance is strongly modulated by soil type and negatively affected by low pH (Postma *et al*., 2011; Tardy *et al*., 2015). Liming is a common agricultural practice in tropical croplands to increase soil pH (Raboin *et al*., 2016) and is also typically applied in the region in Colombia where the agricultural soil used in our study was collected. Consequently, their exclusive presence in the agricultural soil might be related with the higher pH as compared to the native acidic soil. Also the exclusive genus *Aeromicrobium* prefers neutral to alkaline pH and has been previously isolated from agricultural fields (Yoon *et al*., 2005; Cui *et al*., 2007). In terms of activity, both *Lysobacter* and *Aeromicrobium* species are known to produce diverse secondary metabolites with antimicrobial properties (Miller *et al*., 1991; Hayward *et al*., 2010; de Bruijn *et al*., 2015; Gómez-Expósito *et al*., 2015) which may aid in the protection of the bean plants against biotic stress caused by soil-borne pathogens. Further experimentation is needed to decipher the impact of these 'enriched' microbes for common bean growth and health in agricultural and native habitats.

Bacteria exclusively found in the native habitat of wild relatives of crop plants, comprise representatives of *Cytophaga* and *Acidicapsa* genera. The genus *Cytophaga* is known for its cellulose-degrading capabilities and species such as *Cytophaga hutchinsonii* can be found as indigenous soil inhabitants (Wilhelm *et al*., 2017; Zhu and McBride, 2017). Their exclusive presence in the native soil may be associated with their ability to decompose complex carbohydrates such as plant litter and decaying wood, thereby contributing to carbon cycling in the undisturbed native soils. The genus *Acidicapsa*, which belongs to the phylum Acidobacteria, encompasses strictly aerobic chemo-organotrophs that are adapted to acidic conditions (Kulichevskaya *et al*., 2012; Falagán *et al*., 2017). Acidobacteria members are in general considered oligotrophs and have been

found positively associated with low soil pH (Fierer *et al*., 2007; Jones *et al*., 2009). The diversity and abundance of acidobacterial species in soil, as well as their diversity in metabolic traits, makes this phylum a potential player in soil nutrient cycling (Ward *et al*., 2009; Kielak *et al*., 2016). If these rhizobacterial genera, when re-introduced into agricultural soils, will be able to establish and survive in the rhizosphere of modern bean cultivars and, if they can, provide additional life-support functions (growth, health) for the bean plants remains to be investigated. It is important to highlight that the enriched or depleted bacterial taxa explored in this study are based on amplicon sequences that were classified up to genus level. It may be possible that bacterial species that were classified up to genus level in our analysis are absent in one of the soil types. Therefore, additional analyses that allow taxonomic resolution at the species or even strain level are needed.

A highly abundant taxonomic core was shared by all eight bean accessions in both soils. The core microbiome genera we observed included *Rhizobium*, *Bradyrhizobium, Mesorhizobium, Sphingomonas,* and *Streptomyces*. These results showed that a significant portion of the core microbiome of common bean is composed of bacterial genera with nitrogen fixing capabilities, an important feature of microbes associated with leguminous plant species. However, also for other non-leguminous plant species these rhizobacterial genera are members of the core rhizosphere microbiome (Peiffer *et al*., 2013; Yeoh *et al*., 2016; 2017). We further observed that *Rhizobium* was by far the most dominant core member in the agricultural soil while in the native soil the genera *Burkholderia, Ralstonia* and unclassified Rhizobiales were the dominant core members. These latter genera are most likely better adapted to acidic conditions in the native soil and probably responded more efficiently to root signals, such as flavonoids released by roots of common bean. *Burkholderia* species are indeed well represented in acidic soils (Stopnisek *et al*., 2014) and have been found as nodule-forming rhizobia in symbiosis

with leguminous plants (Moulin *et al*., 2001; Elliot *et al*., 2009; Lemaire *et al*., 2016) including common bean (Talbi *et al*., 2010). To form nodules, however, compatibility between *Burkholderia* spp. and the legume host is a key factor (Lemaire *et al*., 2016). In fact, common bean nodulation in tropical acid soils in South America has been associated with only a few *Rhizobium* species (Martínez-Romero *et al*., 1991; Aguilar *et al*., 2004) which were found in low abundances in the native soil. Despite the high abundance of *Ralstonia* in the native soil, a genus known to harbour soil-borne bacterial pathogens, no disease symptoms in common bean roots were evidenced in our study. *Ralstonia* species may occupy several ecological niches and have been isolated from different environments, including soil (Chen *et al*., 2001). It has been shown that tropical leguminous plants can be nodulated by *Ralstonia taiwanensis* that display functional nitrogenase activity (Chen *et al*., 2003). Nevertheless, whether the *Burkholderia* and *Ralstonia* OTUs detected here in the rhizosphere can establish symbiotic associations with common bean is not known yet.

The co-occurrence network analyses further indicated that the interactions between rhizobacterial taxa in the rhizosphere of common bean accessions grown in a native soil environment were more complex than those in an agriculturally managed soil, where the establishment of copiotrophs in the rhizosphere compartment was favored. Therefore, we hypothesize that the rhizobacterial community assembly for common bean grown in the agricultural soil is less complex and more modular than for common bean in native soil. This in turn, makes it relatively more easy for a given soil bacterial species to invade and establish in the rhizosphere of bean plants grown in the agricultural soil. Following this hypothesis, the higher rhizobacterial diversity observed for common bean in the agricultural soil may represent a less specialized microbiome. Accordingly, N-fertilization into soil has been shown to induce shifts in bacterial community composition,

promoting copiotrophs that rely on labile carbon sources (Ramirez *et al*., 2012) as well as promoting the evolution of less-mutualistic microbes (Weese *et al*., 2015). In the 'agricultural' and 'native soil' networks, we observed positive interactions between nodes, which suggest niche overlap, and we also observed negative interactions, suggesting competition or amensalism (Faust and Raes, 2012). The occurrence of phylogenetically close OTUs was in general positively correlated, forming well-differentiated clusters (Fig. 5). Accordingly, it has been shown that Acidobacteria and Verrucomicrobia phyla tend to co-occur more than expected by chance only (Barberán *et al*., 2012). In this study, we found a similar pattern in the native soil network, where cluster 2 is composed mainly of the oligotrophic phyla Acidobacteria and Verrucomicrobia. This cluster interacts negatively with clusters 1 and 3, abundant in copiotrophic bacterial genera, which presumably respond better to the common bean root exudates. Furthermore, the clustering suggests a strong niche differentiation (Faust and Raes, 2012). For instance, cluster 4 in the agricultural network is composed exclusively of rhizobial OTUs, with no interactions with other clusters. Similarly, cluster 2 consisted mainly of Bacteroidetes that may represent the rhizobacterial hub involved in degradation of complex polymers (Thomas *et al*., 2011; Berlemont and Martiny, 2015). Also cluster 3 in the native soil, mainly composed of *Burkholderia*, might represent a specific hub of nodule-forming rhizobia (Talbi *et al*., 2010). Whether these hubs represent distinct functional groups remains to be investigated by metagenomics and trait-based bioassays.

Conclusions

Our study showed that the transition of common bean from a native soil to an agricultural soil led to a gain of rhizobacterial diversity. We found a low diverse but highly abundant core microbiome which resembles that of other plant species, suggesting a homogenization of rhizobacterial diversity of plants grown in different agricultural

landscapes. It is important to note that the core microbiome analysis presented here is a descriptive analysis based solely on taxonomy, and that functional traits should be taken into account in future analyses for better insight into the impact of habitat expansion on trait-based microbiome assembly (Lemanceau *et al*., 2017). The network structure was simpler in agricultural soil as compared to native soil, which again may reflect the process of biotic homogenization. In this study, we also aimed for the identification of microbes that were lost as a consequence of domestication and habitat expansion of common bean. Indeed, several bacterial genera were exclusively found in the native soil and also as an exclusive member of the rhizosphere of wild bean accessions. These bacterial genera were low-abundant members of the rhizobacterial community. Conversely, the number of bacterial taxa exclusively found in the agricultural soil was considerably higher. The proportion of depleted bacterial genera appears to be overcompensated in the agricultural soil by the number of "gained microbes", many of which were highly abundant in the rhizosphere of all eight common bean accessions. On the other hand, this increased bacterial diversity in the agricultural soil might also correspond to a less specialized microbiome. To what extent these "enriched" and "depleted" bacterial genera have an impact on plant growth and health is not yet known and subject of future experiments. It is important to emphasize that the number of agricultural and native soils tested should be further expanded to resolve if the significant changes we observed between the two soils tested in our study can be extrapolated as general trends in rhizobacterial shifts during domestication.

Acknowledgements

J.E.P-J was financially supported by the Department of Science, Technology and Innovation of Colombia - COLCIENCIAS through the doctoral grant 568-2012-15517825. J.M.R. and V.J.C. were supported by the Dutch STW-program "Back to the Roots" and R.M. by CNPq 443112/2014-2. We are grateful to H. A. Pérez and J. A. Pérez

for assistance in the selection and collection of the native soil for the greenhouse experiments.

Supplementary materials

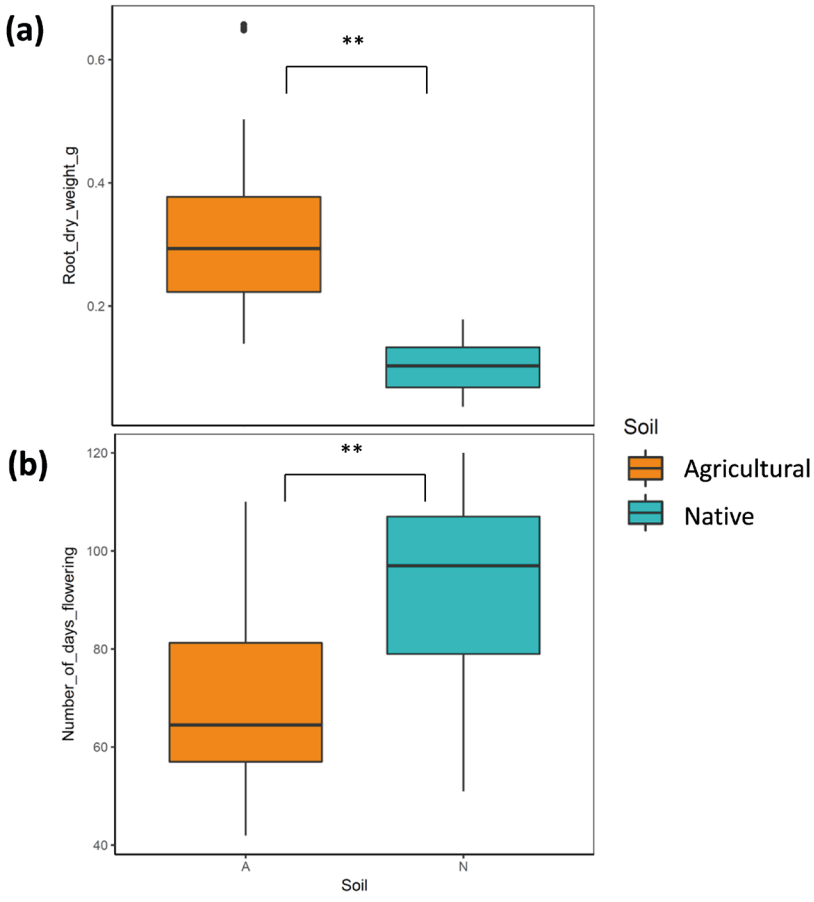

Fig. S1. Phenological traits of the common bean accessions grown in native and agricultural soils. Data for root dry weight (a) and the number of days to reach the flowering stage (b) of common bean accessions grown in agricultural and native soils was merged per soil type and used to built box plots. Subsequently, kruskal and Wallis test was applied followed by a Bonferroni correction. (** significantly different with a $P<0.05$)

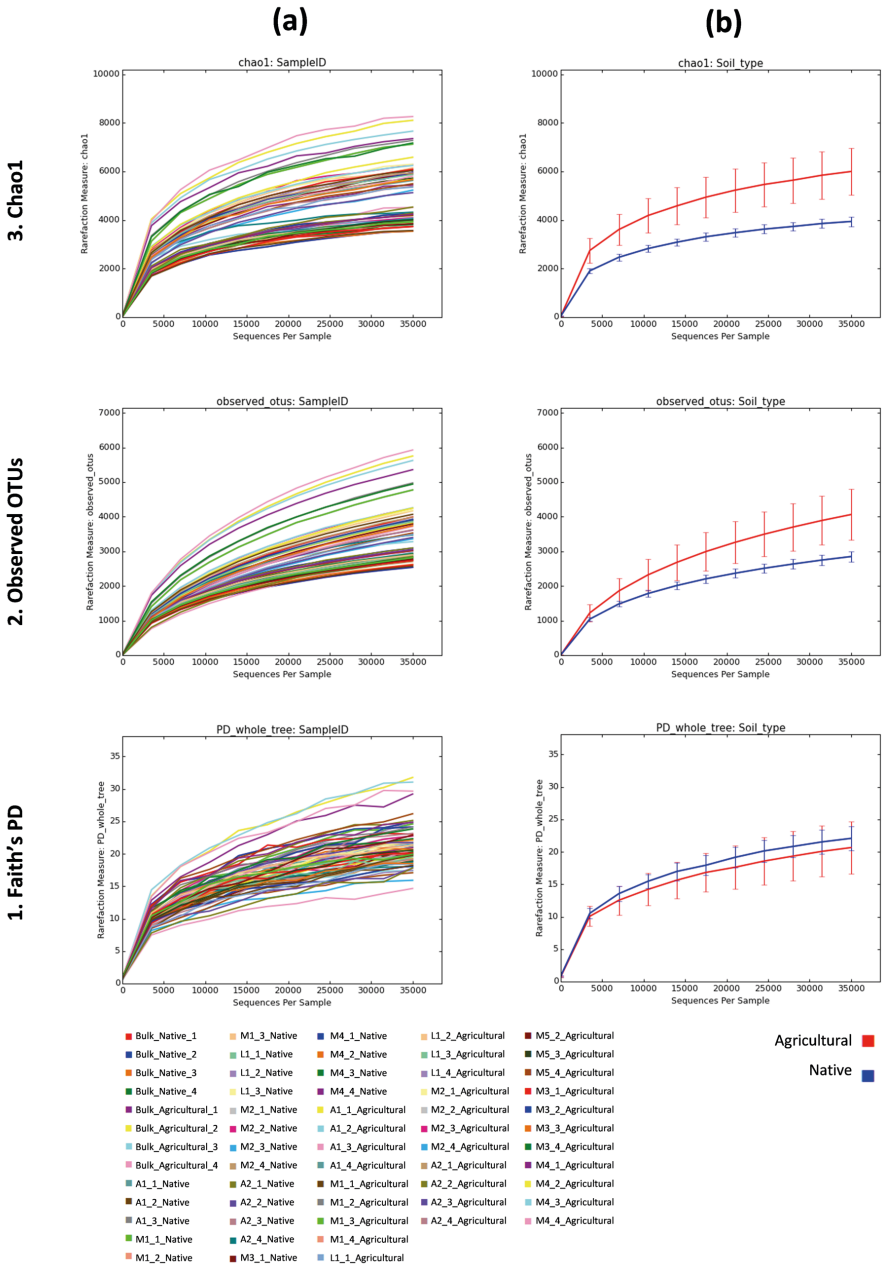

Fig. S2. Rarefactions curves for chao1, observed OTUs and phylogenetic diversity metrics. In the rows, the rarefaction curves are shown by index as follows: 1. Faith's phylogenetic diversity; 2. Number of observed OTUs; 3. chao1 index for estimated OTUs. In the rows, rarefaction curves are shown by (a) Sample; (b) Soil type.

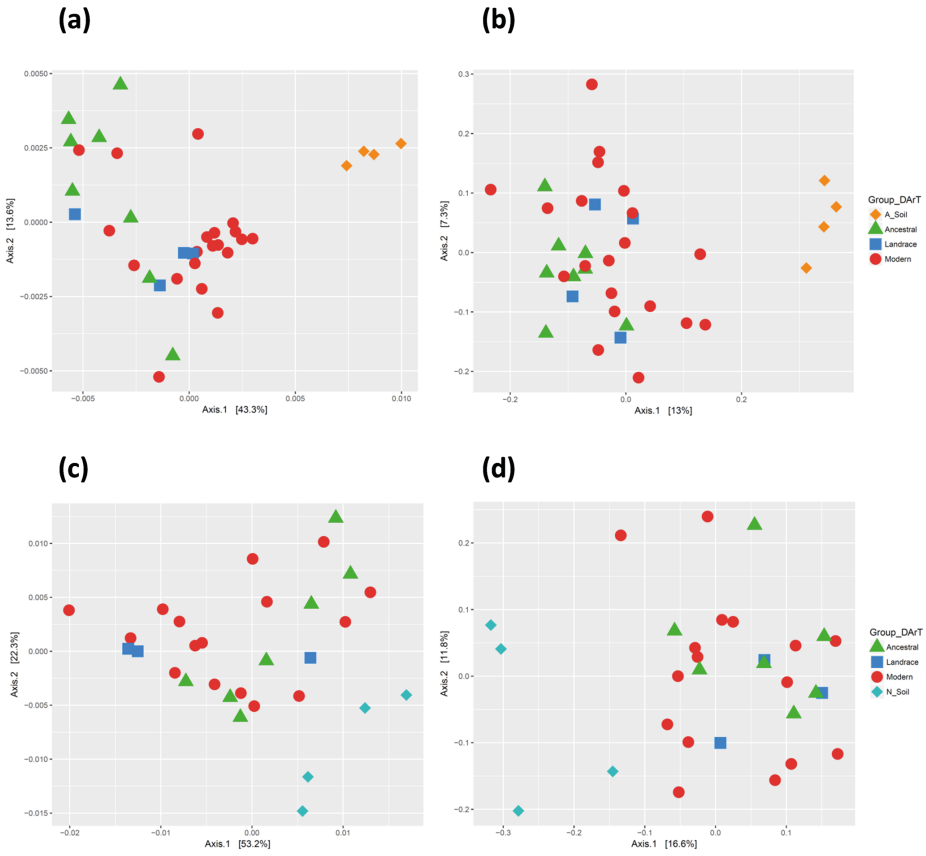

Fig. S3. Rhizosphere bacterial community structure using weighted Unifrac metrics in agricultural and native soils. Principal Coordinate Analysis (PCoA) of 16S rRNA diversity in the rhizosphere of the eight common bean accessions used in this study. (a) weighted and (b) unweighted unifrac metrics of agricultural rhizosphere and soil samples and (c) weighted and (d) unweighted unifrac metrics of native rhizosphere and soil samples.

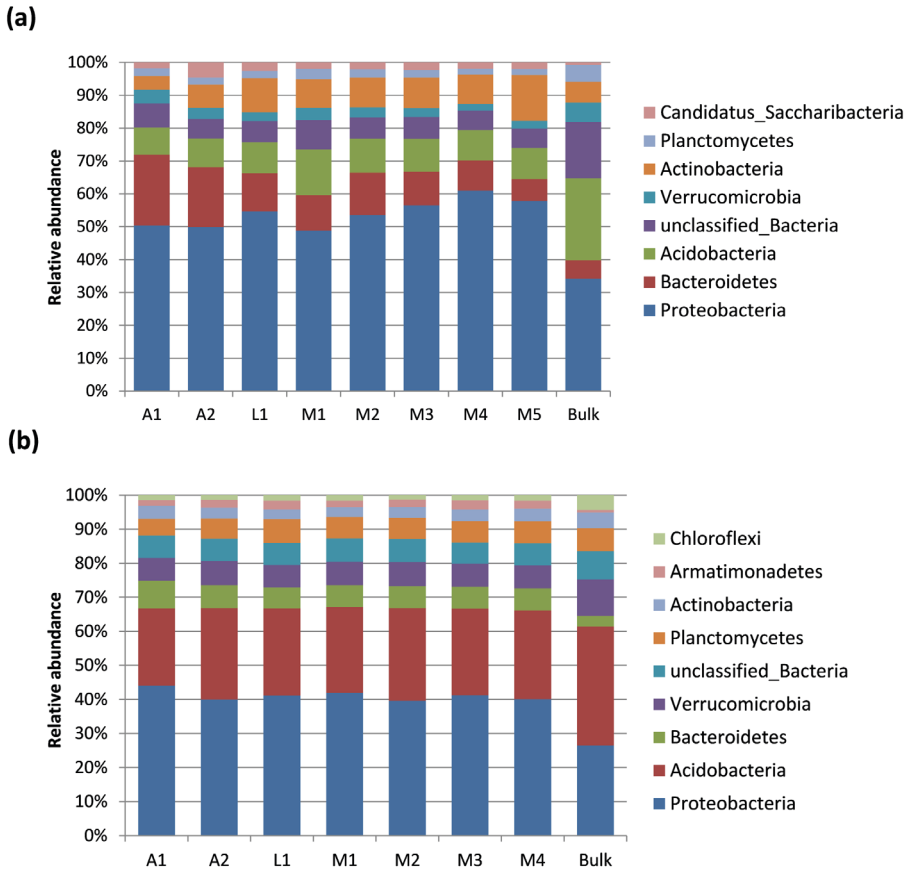

Fig. S4. Relative abundance of the most abundant bacterial phyla in agricultural and native soils. Bar graphs of the relative abundance of the most abundant bacterial phyla in the agricultural (a) and in the native (b) soil are shown. Only phyla with a total relative abundance higher than 1% are depicted in the graphs.

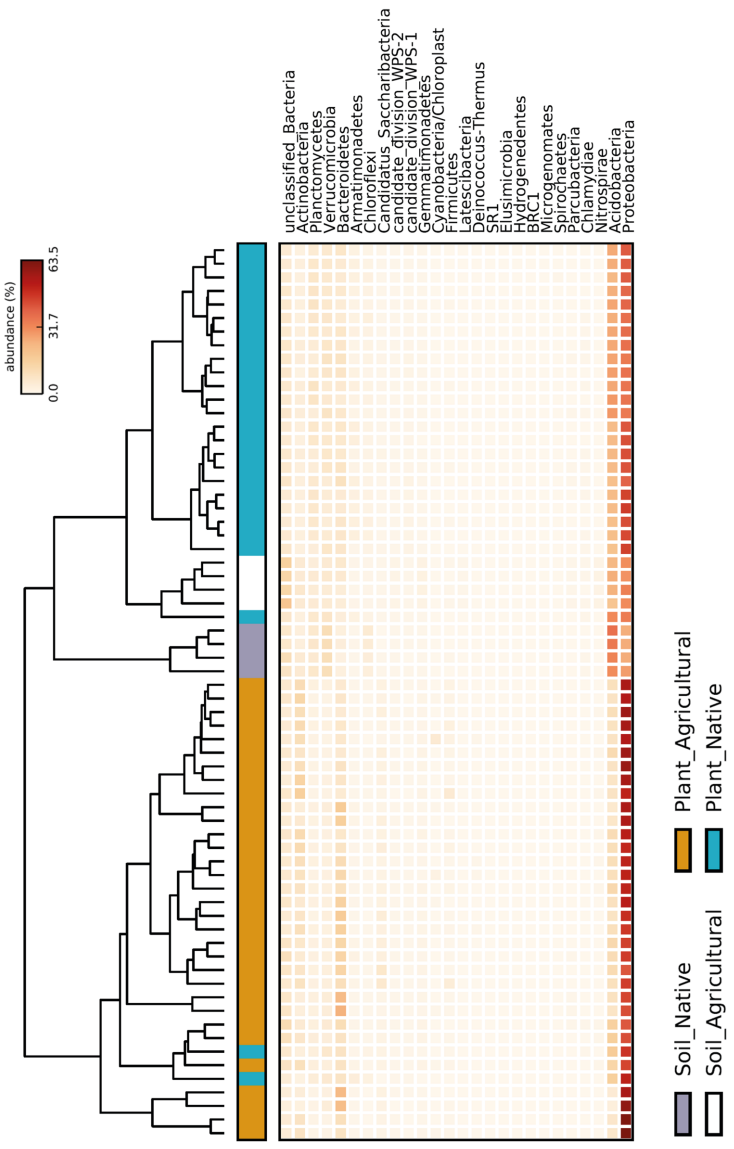

Fig. S5. Heat map of the relative abundance of all the bacterial phyla in agricultural and native soils. The columns indicate the common bean accessions clustered through a dendogram built with an average neighbor method (UPGMA) and the rows shows the bacterial phyla observed in rhizosphere and bulk soil samples. The darker the red squares in the heat map the more abundant the phylum of the specific row. Color codes for the samples are indicated in the lower section of the graph.

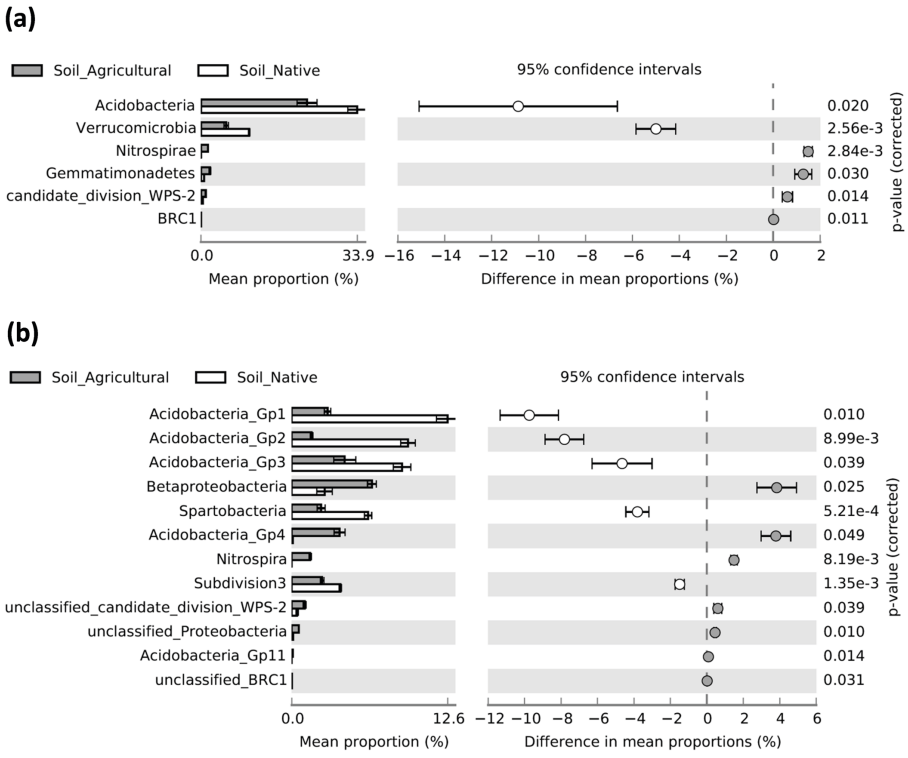

Fig. S6. Differential abundance of bacterial phyla and classes between bulk soil samples of the agricultural soil and the native soil. Welch's t-tests followed by Bonferroni (P<0.05) corrections were performed at Phylum (a) and Class level (b). Only Phyla and Classes differentially abundant are shown.

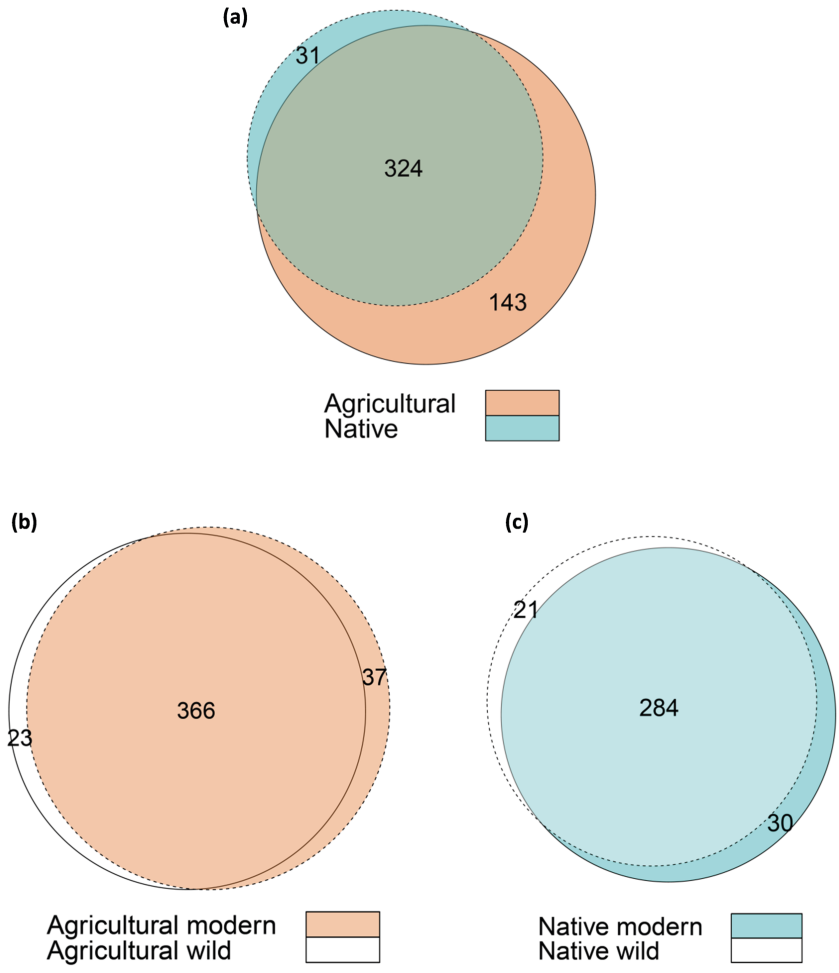

Fig. S7. The enriched and depleted bacteria in the rhizosphere of wild and modern common bean grown in native and agricultural soils. Area-proportional Euler diagrams were built to depict the exclusive and the shared genera. (a) Number of genera shared between agricultural and native soils is depicted within the intersection while the number of genera exclusive to each soil type can be seen out of the intersection zone. The genera exclusive to the native soil are visible in the cyan colored area, while genera exclusive to the agricultural soil are visible in the dark orange colored area. (b) Euler diagram between modern accessions and wild accessions in the agricultural soil. The genera exclusive to the modern accessions is visible in the dark orange colored area, while genera exclusive to the wild accessions is visible in the white area. (c) Euler diagram between modern accessions and wild accessions in the native soil. The genera exclusive to the modern accessions are visible in the cyan colored area, while genera exclusive to the wild accessions are visible in the white area.

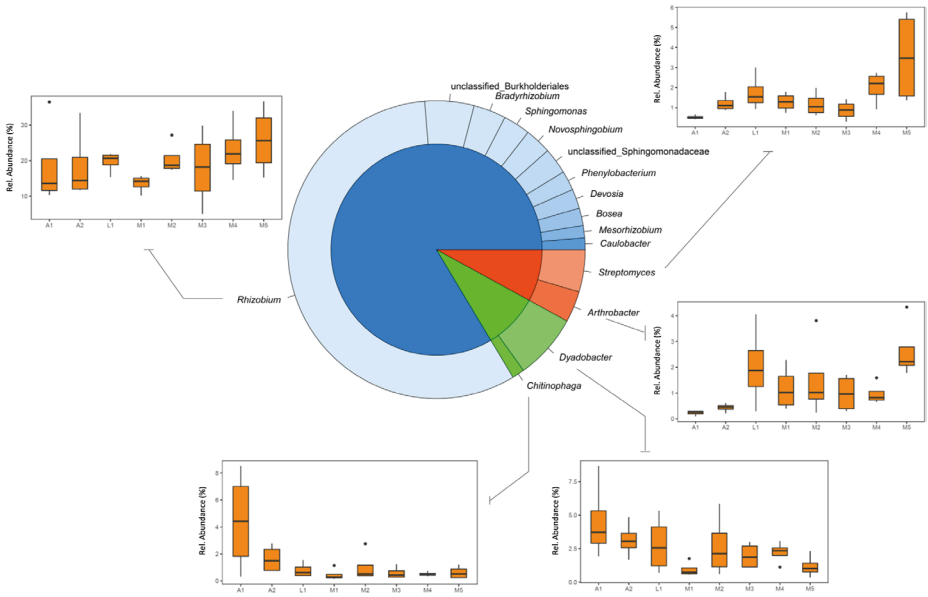

Fig. S8. Core microbiome in the rhizosphere of common bean grown in agricultural soil. The different portions within the inner pie chart represent the bacterial phyla. The outer donut plot represents the genera that are part of the core. The size of the different pie and donut portions represents the contribution of each phylum/genus to the total relative abundance. Satellite box plots depict the relative abundance of selected genera by accession in the agricultural soil.

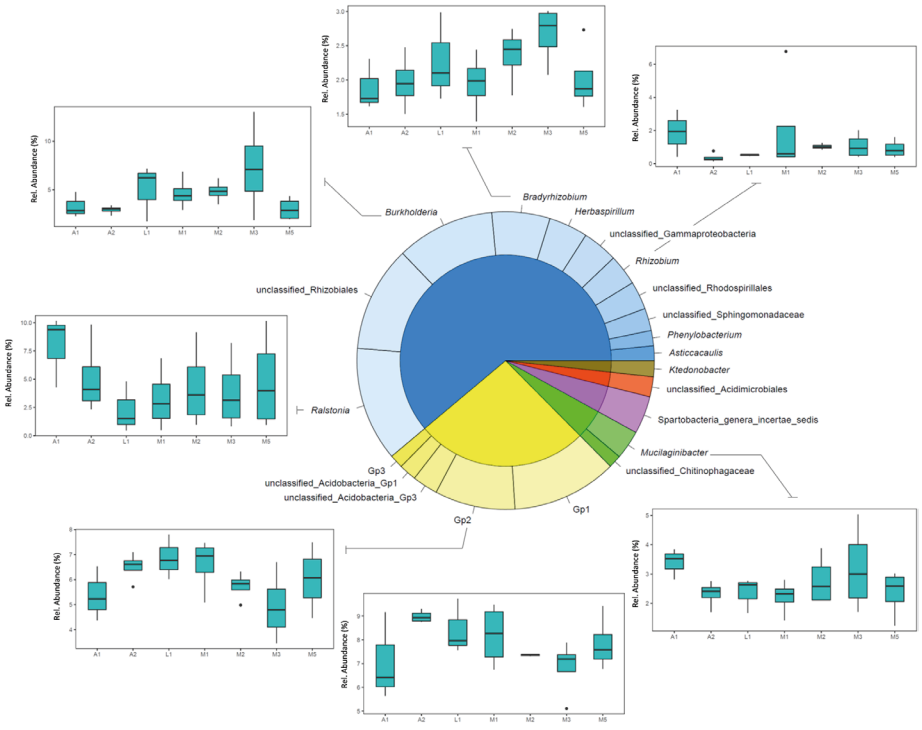

Fig. S9. Core microbiome in the rhizosphere of common bean grown in native soil. The different portions within the inner pie chart represent the bacterial phyla. The outer donut plot represents the genera that are part of the core. The size of the different pie and donut portions represents the contribution of each phylum/genus to the total relative abundance. Satellite box plots depict the relative abundance of selected genera by accession in the native soil.

Table S1. Sequence characteristics obtained in the agricultural soil. In the table the raw read and after processing read counts are presented as well as quality characteristics of the raw reads.

Soil type	Sample	Total Bases	Raw read count	GC (%)	Q20 (%)	Q30 (%)	Processed read count
Agricultural	A1_1	63363676	211312	54.71	86.63	74.13	56996
Agricultural	A1_2	78985945	263296	55.18	87.15	75.19	77864
Agricultural	A1_3	69257541	231142	54.3	87.72	76.1	70601
Agricultural	A1_4	72760654	242636	54.82	87.79	76.02	69084
Agricultural	A2_1	69773872	232626	55.02	86.96	75.01	67646
Agricultural	A2_2	58210761	194134	54.35	88.36	76.93	60724
Agricultural	A2_3	76327328	254550	54.79	87.61	75.93	75603
Agricultural	A2_4	70190462	234008	55.13	86.55	74.44	67151
Agricultural	L1_1	72057959	240260	54.87	87.57	75.82	72136
Agricultural	L1_2	64508620	215002	55.81	87.01	74.95	63095
Agricultural	L1_3	68590650	228646	55.27	87.21	75.28	66964
Agricultural	L1_4	71258796	237562	55.41	87.34	75.6	72020
Agricultural	M1_1	71348421	237868	55.22	88	76.53	71859
Agricultural	M1_2	65375690	217894	55.94	86.24	73.77	58752
Agricultural	M1_3	67841335	226132	55.83	87.02	74.98	63630
Agricultural	M1_4	65532781	218424	55.57	86.86	74.81	63673
Agricultural	M2_1	67221432	224064	55.6	87.34	75.58	67676
Agricultural	M2_2	62678239	209020	55.15	87.26	75.32	61303
Agricultural	M2_3	66567882	221854	55.88	87.19	75.32	67063
Agricultural	M2_4	60348755	201286	54.78	87.97	76.36	61242
Agricultural	M3_1	59335191	197832	55.59	87.46	75.89	60023
Agricultural	M3_2	68861207	229586	55.1	88.35	76.94	70813
Agricultural	M3_3	68907707	229716	55.49	87.93	76.19	71553
Agricultural	M3_4	68367576	227898	55.67	87.18	75.22	68092
Agricultural	M4_1	74886328	249604	55.37	87.99	76.57	79568
Agricultural	M4_2	74054373	246918	55.4	87.51	75.82	73404
Agricultural	M4_3	74769091	249182	55.69	87.19	75.23	79061
Agricultural	M4_4	66904055	223064	55.4	87.95	76.29	69317
Agricultural	M5_1	62075633	206888	55.89	86.95	74.93	60638
Agricultural	M5_2	72482267	241660	55.61	87.82	76.35	74295
Agricultural	M5_3	61109216	203688	55.7	87.48	75.84	62822
Agricultural	M5_4	71365327	237914	55.51	87.73	76.17	72160
Agricultural	Bulk_soil_1	61607085	205346	56.88	86.33	73.98	59869
Agricultural	Bulk_soil_2	65310513	217716	56.66	85.85	73.09	58020
Agricultural	Bulk_soil_3	55162959	183890	56.6	86.69	74.51	51738
Agricultural	Bulk_soil_4	60880405	202966	56.4	86.27	73.96	55876

Table S1 (Continuation). Sequence characteristics of samples obtained in the native soil. In the table the raw read and after processing read counts are presented as well as quality characteristics of the raw reads.

Soil type	Sample	Total Bases	Raw read count	GC (%)	Q20 (%)	Q30 (%)	Processed read count
Native	A1_1	69790674	232642	55.48	87.61	75.83	68643
Native	A1_2	59019732	196740	55.3	87.94	76.32	59343
Native	A1_3	64931567	216392	56.16	86.45	73.82	60661
Native	A2_1	57934044	193128	56.32	86.58	74.47	55570
Native	A2_2	61055744	203522	56.51	87.04	74.95	60234
Native	A2_3	67056980	223502	55.93	86.85	74.57	63884
Native	A2_4	58518207	195056	55.75	87.08	75.01	56197
Native	L1_1	65521365	218470	56	87.59	75.83	65232
Native	L1_2	64785862	215952	55.81	87.13	75.17	65472
Native	L1_3	71386826	238006	56.23	87.42	75.58	70322
Native	M1_1	63411035	211380	55.54	87.88	76.22	62462
Native	M1_2	62074225	206932	56.22	87.78	76.08	63354
Native	M1_3	71515302	238420	56.3	86.88	74.78	68664
Native	M1_4	54123563	180468	55.92	87.98	76.29	55091
Native	M2_1	67140790	223824	56.07	87.71	76.01	67260
Native	M2_2	66608745	222070	56.26	87.66	75.92	65336
Native	M2_3	58726189	195734	56.03	86.89	74.83	56402
Native	M2_4	70057132	233546	55.9	87.05	75.05	69406
Native	M3_1	63700948	212352	55.9	87.25	75.32	63248
Native	M3_2	58356921	194520	56.26	86.7	74.5	55418
Native	M3_3	64816913	216052	56.18	86.53	74.28	60174
Native	M3_4	52757955	175792	55.73	85.03	71.51	45340
Native	M4_1	61945280	206504	55.92	87.14	75.06	60709
Native	M4_2	66279461	220998	55.67	87.04	74.99	63633
Native	M4_3	62281114	207616	55.92	87.18	75.18	59404
Native	M4_4	52537073	175068	56.84	82.13	67.37	35547
Native	Bulk_soil_1	51904334	173074	57.05	86.92	74.74	50134
Native	Bulk_soil_2	57678950	192304	57.01	86.5	74.18	54399
Native	Bulk_soil_3	64011567	213402	56.82	85.94	73.51	59298
Native	Bulk_soil_4	61453027	204890	56.77	86.45	74.16	57379

Table S2. AIC values for six rank abundance distribution models to test niche neutral models in native soil samples. The lowest AIC (Akaike Information Criterion) value for each sample represents the best fit model which is shown in bold.

	AIC					
Sample	**Broken-stick**	**Pre-Emption**	**Log-Normal**	**Zipf**	**Zipf-Mandelbrot**	**ZSM**
A1_1	109193.1	82138.6	**16519.6**	23002.6	22313	27723.2
A1_2	94184.6	69323.1	**12489.5**	17758.8	15560.7	25791.4
A1_3	75917.1	53531.6	**12942.3**	22601.2	nd	31087.8
A2_1	63993.3	46570.4	**13574.7**	22370.6	nd	29473.8
A2_2	71558.9	51409.5	**13565.3**	22980.2	nd	29706
A2_3	97215	73862	**16099**	22244	21654	27916.2
A2_4	67463.6	47477.7	**12704.8**	21889.1	nd	28406.8
L1_1	84656.2	57675.6	**13392.4**	24235.9	nd	29305.8
L1_2	76243.7	53640.4	**13452.2**	24314.4	nd	31883.8
L1_3	83374	53282	**14091**	28556	nd	28726.4
M1_1	102404	78828.6	**15112.9**	18962.9	17328.2	30968.2
M1_2	76043.8	51113.2	**12713**	23954.3	nd	30633.6
M1_3	86084.7	58044.4	**14829**	27170	nd	31285
M1_4	63313.4	45584.3	**12505.8**	21131.8	nd	28744.8
M2_1	81233.1	53271.9	**14479.4**	27657.3	nd	28957.2
M2_2	70881	47165.9	**14569**	27979.2	nd	31003.2
M2_3	82526.2	59967.7	**12385.5**	18537.5	nd	26264.8
M2_4	95319.6	64444	**12840.3**	23598.4	nd	27933.6
M3_1	83007.8	55441.1	**13188.9**	24134.3	nd	29499.2
M3_2	80158.3	60099.8	**13045.7**	18270.1	16701.8	29880.6
M3_3	82194.7	59553.1	**13634.7**	22082.8	nd	33823
M3_4	69381	50489	**10947.5**	15735.1	12638.3	25404.7
M4_1	65597.2	45892.6	**13894.5**	25206.3	nd	33552.8
M4_2	98371.9	75718.3	**16064.2**	21553.6	21041.2	29379
M4_3	75332.4	54770.3	**13519.5**	22214.1	17947.7	28705.4
M4_4	38329.1	29292.2	**11450.2**	17028.1	nd	27612.2
Bulk_soil_1	61291.3	41596.6	**12405.9**	21171.8	nd	24761.1
Bulk_soil_2	70455.2	45471.4	**12031.8**	21858.8	nd	24218.5
Bulk_soil_3	72354.2	47488.2	**12155.4**	22969	nd	25293
Bulk_soil_4	70091.6	45407.7	**12925**	24283.3	nd	25246.2

Table S3. Co-occurrence network properties of 16S rDNA rhizobacterial reads of common bean rhizosphere in agricultural and native soil

Network properties	Common bean Agricultural soil	Common bean Native soil
Number of nodes[a]	63	89
Number of edges[b]	61	176
Positive edges[c]	60	158
Negative edges[d]	1	18
Modularity	0.822	0.584
Number of communities	17	13
Network diameter	11	11
Average path length	3.87	4.06
Average degree	0.968	1.978
Average clustering coefficient	0.5	0.551

[a]Bacterial taxa (genus level) with at least one significant (P<0.01) and strong (SparCC > 0.8 or < -0.8) correlation.

[b]Number of connections/correlations obtained by SparCC analysis;

[c]SparCC positive correlation (> 0.8 with $P < 0.01$);

[d]SparCC negative correlation (< -0.8 with $P < 0.01$)