This is a post-peer-review, pre-copyedited version of an article published in Plant and Soil. The final authenticated version is available online at: http://dx.doi.org/10.1007/s11104-021-05049-x

Title: A core of rhizosphere bacterial taxa associate with two of the world's most isolated plant congeners

AUTHORS: Johannes J. Le Roux[1,4*†], Pedro W. Crous[2], Casper N. Kamutando[3], David M. Richardson[4], Dominique Strasberg[5], Michael J. Wingfield[6], Mark G. Wright[7], Angel Valverde[8†]

[1] Department of Biological Sciences, Macquarie University, Sydney, NSW 2113, Australia

[2] Westerdijk Fungal Biodiversity Institute, Uppsalalaan 8, 3584CT Utrecht, The Netherlands

[3] Department of Plant Production Sciences and Technologies, University of Zimbabwe, Mount Pleasant, Harare, Zimbabwe

[4] Centre for Invasion Biology, Department of Botany & Zoology, Stellenbosch University, Stellenbosch, South Africa

[5] UMR PVBMT (Peuplements Végétaux et Bio-agresseurs en Milieu Tropical), Reunion University, 97744 Saint Denis, Reunion Island, France[6]Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa

[7] Department of Plant and Environmental Protection Sciences, University of Hawaii at Manoa, Honolulu, Hawaii, USA

[8] Instituto de Recursos Naturales y Agrobiología de Salamanca (IRNASA-CSIC), Consejo Superior de Investigaciones Científicas, Salamanca, Spain

[*] Corresponding Author: jaco.leroux@mq.edu.au, +64 484945367

[†] These authors contributed equally to this work.

Abstract

Aims: Understanding the contributions of abiotic and biotic conditions to soil microbial diversity, structure, and function, remains a central focus in soil biology and biogeochemistry. Here we aim to determine how geography and host plant identity influence these different osphere bacterial communities and endosymbionts associated with *Acacia heterophylla* on La Réunion island (Mascarene archipelago, Indian Ocean) and *A. koa* in the Hawaiian Islands (Hawaiian archipelago, Pacific Ocean). These two tree species are remarkable: they are each other's closest living relatives despite their habitats being more than 16 000 km apart.

Methods: Using 16S rRNA amplicon next-generation sequencing data we show that the structure of rhizosphere communities of these two acacias is largely driven by dispersal limitation between sites and local soil physicochemical conditions within sites.

Results: Despite high taxonomic turnover in soils collected from different sites, we found their predicted functions to be largely similar, suggestive of functional redundancy. We also identify a core of rhizosphere taxa associated with both *Acacia* species in both archipelagos, which included nitrogen-fixing mutualists. Isolation and characterisation of the latter from root nodules of both species further supported strong selection by these plants for the same *Bradyrhizobium* endosymbionts.

Conclusions: Overall, our data suggest that phylogenetically-closely related plants may show remarkably similar selectivity for microbial mutualists over vast geographic distances.

Keywords: acacia; *Bradyrhizobium*; core microbiome; dispersal limitation; host selectivity; island biogeography; rhizosphere soil

Declarations: Not applicable

Introduction

Soil microbes play important roles in mediating plant diversity and productivity (van der Heijden 2008). For instance, communities of rhizosphere microbial symbionts benefit plants by enhancing water and nutrient uptake or suppressing pathogens (Mendes et al. 2012; Vandenkoornhuyse et al. 2015). Interactions between plants and their soil symbionts such as mycorrhizal fungi, and the assembly of rhizosphere communities in general, have deterministic and stochastic components (Nemergut et al. 2013; Fitzpatrick et al. 2018). These represent trait-based and neutral assembly mechanisms, respectively. During deterministic processes, such as plant-microbe compatibility (e.g., Louca et al. 2017) and adaptation (e.g., Fitzpatrick et al. 2018), trait differences are decisive. On the other hand, stochastic processes that govern rhizosphere microbial community assembly involve chance encounters of partners, infection pressure, infection priority or propagule dispersal (Nemergut et al. 2013; Le Roux et al. 2017; Martín-Robles et al. 2018; Ramoneda et al. 2019). The diversity and composition of plant-associated soil microbiomes are also influenced by abiotic conditions such as temperature, soil water availability and organic carbon content (Fierer 2012, Bulgarelli et al. 2013, Philippot et al. 2013). Plants may also change soil abiotic conditions, for example, through the quantity and quality of litter input (e.g., Vitousek and Walker 1989) and rhizodeposits (e.g., Pascale et al. 2020), which may further impact microbial communities. Unsurprisingly, rhizosphere microbial communities can be highly variable between locations and plants species, even over very small spatial scales (Fierer 2017; Fitzpatrick et al. 2018). In addition, as different microbes might be able to perform the same functions (i.e., there is a high level of functional redundancy among soil microbe taxa; Banerjee et al. 2016), the functions of rhizosphere microbiomes may be more stable than their taxonomic make-up (Louca et al. 2017). On the other hand, co-evolution between plants and essential mutualists such as mycorrhizal fungi and rhizobia, will lead to strong taxonomic and functional overlap between the mutualists of phylogenetically closely-related plant species (e.g. Brundrett and Tedersoo 2018; Stępkowski et al. 2018).

Nitrogen-fixing rhizobia form a polyphyletic group of bacteria residing in the α- (e.g., *Rhizobium*) and β-proteobacteria (i.e., *Burkholderia*) classes (Moulin et al. 2001). Most rhizobia are free-living heterotrophs that associate facultatively with legumes (family Fabaceae) by forming specialised structures (nodules) on host plants. Biological nitrogen fixation (BNF) occurs within the nodules, where rhizobia reduce the inorganic atmospheric di-nitrogen into an organic form (i.e., ammonium), which is transferred to the host plant in exchange for carbon-rich photoassimilates. This association is especially important for the establishment, growth and survival of legumes in nutrient-poor environments (van der Heijden

2008). With some exceptions (e.g., Le Roux et al. 2018), studies of rhizobial diversity have focused on the root-nodule environment and very little is known about rhizobia found in rhizospheric soils and whether legumes recruit these from the abundant or rare bacterial biosphere (Pedrós-Alió 2012).

Here we investigate the taxonomic and functional diversity of rhizosphere bacterial communities of two *Acacia* species: *A. koa* A. Gray, endemic to the Hawaiian Islands (Hawaiian archipelago, Pacific Ocean), and *A. heterophylla* Willd., endemic to La Réunion island (Mascarene archipelago, Indian Ocean). These tree species represent an unusual study system because they are closely related, despite a more than 16 000 km separation in their distribution ranges (Le Roux et al. 2014). *Acacia heterophylla* colonised La Réunion island approximately 1.4 million years ago following an extreme long-distance dispersal event from the Hawaiian Islands (Le Roux et al. 2014). Phylogenetic analyses suggest that these two island endemics need taxonomic revision since *A. heterophylla* renders *A. koa* paraphyletic (Le Roux et al. 2014). Some have argued that these two taxa are the tetraploid descendants of a diploid ancestor related Australian black wattle, *A. melanoxylon* R.Br. (Coulaud et al. 1995; Le Roux 0 et al. 2014).

We generated 16S rRNA gene amplicon high-throughput sequencing data for rhizosphere soil bacterial communities and Sanger-sequencing data for rhizobia isolated from the root nodules of both *Acacia* species in their natural ranges for nodulation (*nodA*) and 16S rRNA genes. This approach allowed us to infer the effect of geographic isolation on the composition of rhizosphere bacterial communities, diversity and predicted function. We expected the overall rhizosphere community diversity, composition, and function to be strongly influenced by local soil abiotic conditions and/or dispersal limitation within and between islands. On the other hand, we expected to find that both tree species share a taxonomic and functional "core" rhizosphere microbiome of potential mutualists. We also hypothesised that similar rhizobia will nodulate *A. heterophylla* and *A. koa* given their close phylogenetic relationship and the known specificity between members of the genus *Acacia* and *Bradyrhizobium* bacteria.

Materials and methods

Field sampling

During a field survey in March 2015, we located two populations of *A. heterophylla* in each of two sites in La Réunion island: Parc National (Bébour forest; S21.09563, E55.55148) and Volcano (road to Piton de la Fournaise; S21.21245, E55.6137). At each site we excavated soil from the root zones of four mature trees at least 20 m apart. Using sterilised equipment, we sampled *c*. 200 g of soil directly around the roots of each individual tree (n = 8). Soils were kept in insulated cooler boxes until being transferred to a freezer upon arrival in the laboratory. We also collected between 10-50 root nodules from each tree. Root nodules were placed on silica gel in the field to dehydrate until further use. In September 2015, we identified two populations of *A. koa* in the Hawaiian Islands, one on Oahu (N21.40102, W157.88721) and the other on Hawaii island (N19.68749, W155.46565). We used the same sampling procedure as for *A. heterophylla*.

Soil chemical analyses

pH, N, C, P, Ca, Mg, K analyses of Hawaiian soils were done at the College of Tropical Agriculture and Human Resources Agricultural Diagnostic Center, Honolulu, and analyses of soils from La Réunion at CIRAD, Centre de cooperation Internationale en Recherche Agronomique et Développement, Saint-Denis. Extractable phosphorus levels were determined using the Olsen method (Olsen et al. 1954, Olsen and Sommers 1982). Briefly, an extracting solution of 0.5 M $NaHCO_3$ (pH 8.5) in a soil-to solution ratio of 1:20 with 2.5 g of soil was shaken for 30 min. The slurry was filtered and phosphorus measured on an ICP-OES. For extractable soil cations (Ca, K, Mg), 2.5 g of soil and 50 ml of extractant (1M Ammonium acetate, pH 7.0) was shaken for 10 min and measured on ICP-OES. Total nitrogen and carbon were determined by dry combustion on a LECO CN2000.

Soil DNA extraction and 16S rRNA amplicon NGS

For rhizosphere bacterial community analysis, total genomic DNA was extracted from 0.25 g of each soil sample (within three days of collection) using the PowerSoil® DNA extraction kit (MO BIO laboratories Inc., Carlsbad, CA, USA) and following the manufacturer's protocol. We used the primers 799F (5'-AAC MGG ATT AGA TAC CCK G-3') and 1391R (5'- GAC GGG CGG TGW GTR CA-3') to amplify the V5-V7 hypervariable regions of the 16S rRNA gene, with sample-specific barcodes in the forward primer. Amplification was done using a 30 cycle PCR and the HotStarTaq Plus Master Mix Kit (Qiagen, Valencia, CA, USA) and the following thermal conditions: 94°C for 3 min, followed by 28 cycles of 94°C for 30 sec, 53°C for 40 sec and 72°C for 1 min, followed by a final elongation at 72°C for 5 min, were used.

After amplification, PCR products were checked on a 2% agarose gel for amplification success and the relative intensity of bands. Multiple PCR samples were pooled in equal proportions based on their molecular weight and DNA concentrations. Pooled samples were purified using calibrated Ampure XP beads (Agencourt Bioscience Corporation, Beverly, MA, USA) and used to prepare DNA libraries by following the Illumina TruSeq DNA library preparation protocol. Sequencing was performed at MR DNA (www.mrdnalab.com, Shallowater, TX, USA) on an Illumina MiSeq instrument (Illumina, San Diego, CA, USA) following the manufacturer's guidelines.

Bioinformatics, functional gene prediction and community compositional analyses of 16S rRNA amplicon sequencing data

Bioinformatic processing was performed using QIIME version 1.9.1 (Caporaso et al. 2010) as described in Kamutando et al. (2018). Sequences were clusterd into OTUs at 97% identity and cluster representative sequences were assigned taxonomy using the SILVA database (release 123). OTUs not classified as bacteria and those classified as chloroplasts or mitochondria were removed. The raw sequencing data were deposited at the National Centre for Biotechnology Information (NCBI) Sequence Read Archive (SRA) (Accession number: to be included upon acceptance of the manuscript).

OTUs were used to predict the functional variation of bacterial communities using phylogenetic investigation of communities by reconstruction of unobserved states (PICRUSt2) (Douglas et al. 2020). PICRUSt2 predicts the abundance of functional genes (expressed as Kegg Orthologs (KOs)) in a prokaryotic community using bacterial and archaeal genomes. To investigate the functional core, we selected 61 KOs (Supplementary file **S6**) that have been reported as relevant for plant fitness (Lemanceaun et al. 2017) and nitrogen cycling (e.g., Dini-Andreote et al. 2016). Nitrogen is the most limiting nutrient to plant growth and therefore rhizosphere nitrogen cycling has received considerable attention.

Taxonomic diversity indices (Species richness, Shannon-Weiner index) were calculated using the vegan (Oksanen et al. 2013) package in R (R Development Core Team 2013). Prior to calculating diversity, all samples were rarefied to the same sequencing depth (20 509 sequences per sample). We used both rarefied and unrarefied data to explore community composition and structure. Soil chemistry data were standardized and pairwise Euclidean distances between samples were computed. Differences in soil chemistry were visualized using principal component analysis (PCA). Bacterial OTU counts and KOs were transformed to relative abundance and pairwise distances computed based on Bray-Curtis

dissimilarity distances. Differences in community composition were visualised using principal coordinate analysis (PCoA). To determine which factors (location, pH, phosphorus and calcium levels) explained most variability in taxonomic community composition we used PERMANOVA in the R vegan package. Carbon, nitrogen, potassium and magnesium levels were excluded from the model because they were highly correlated with other variables (data not shown). The correlation between taxonomic (OTUs) and functional (KOs) community composition was explored using a Mantel test in R. Differences in diversity and relative abundance between phyla and families were tested using nonparametric Kruskal-Wallis tests in R.

To quantify the relative contributions of dispersal limitation and abiotic conditions (habitat filtering) in shaping community phylogenetic assembly, we used the β-Nearest Taxon Index (βNTI) and Bray-Curtis-based Raup-Crick dissimilarities (RC_{bray}) (Stegen et al. 2012) obtained with the iCAMP package in R (Ning et al. 2020). To infer taxa that might be positively selected by the *Acacia* species sampled, we used the approach described by Burns et al. (2016). The contribution of rare and abundant taxa to changes in community composition was assessed using the R otuSummary package (Yang 2020). Taxa (OTUs) were defined as abundant or rare if their average relative abundances were above or below 0.1% across all samples, respectively (Pedrós-Alió 2012). Core taxa were defined as those that were present in all samples (Shade and Stopnisek 2020). Core taxa whose abundances were significantly different between *A. heterophylla* and *A. koa* and rhizosphere soils were identified using the generalised linear models implemented in the R package DESeq2 (Love et al. 2014). Significantly different functional genes (KOs) abundances between *A. heterophylla* and *A. koa* were identified using Welch's t-tests with the statistical analysis of metagenomic profiles (STAMP) software (Parks and Beiko, 2010). Unless stated otherwise, statistical significance was assessed at $\alpha < 0.05$. Where applicable, P-values were adjusted for multiple comparisons using the Benjamini-Hochberg false discovery rate (FDR).

Root nodule rhizobium phylogenies

Three desiccated root nodules/tree were placed in 1 mL distilled water and left overnight to rehydrate. Rhizobia were axenically isolated from single nodules following Somasegaran & Hoben (1994) with minor modification, viz. submersion in 3.5% sodium hypochlorite for 60 seconds instead of acid sterilisation. Rhizobia were grown at 28°C on Yeast Mannitol Agar

(YMA) supplemented with Congo Red dye, and restreaked until purity was achieved. Purity was confirmed through gram-staining.

Pure rhizobial cultures were grown in Yeast Mannitol broth at 28°C for 5 days before genomic DNA extraction using the Sigma Gen-Elute Bacterial Genomic DNA kit (Sigma-Aldrich Co. LLC, USA) according to the manufacturer's specifications. The 16S rRNA gene and nodulation gene, *nodA*, were amplified using the 27F / 1492R (Lane 1991) and nodA-1F / nodA-2R (Haukka et al. 1998) primer combinations, respectively. These regions were separately amplified in 50 µl PCR reactions, each containing: 5 µl template DNA, 5 µl of each forward and reverse primer (5 µM), 5 µl of 10 X buffer, 3µl $MgCl_2$ (25mM), 1µl dNTPs (10mM; Fermentas) and 1µl of Taq polymerase (5U/µl). For the *NodA* region the following PCR conditions were used: initial denaturation at 93°C for 2 min; 35 cycles of denaturation at 93°C for 45 sec, annealing at 53°C for 1 min, extension at 72°C for 2 min; final extension at 72°C for 5 min. For the 16S rRNA gene region the following PCR conditions were used: initial denaturation at 94°C for 2 min; 35 cycles of denaturation at 94°C for 90 sec, annealing at 55°C for 90 sec, extension at 72°C for 2 min; a final extension at 72°C for 10 min. Amplified PCR products were purified using the Qiaquick PCR purification kit (Qiagen GmbH, Germany) and sequenced in both directions using the ABIPRISM BigDye Terminator Cycle Sequencing Ready Reaction kit and an automated ABI PRISM 377XL DNA sequencer (PE Applied Biosystems, Foster City, CA, USA) and the same primers used for PCR amplification. All DNA sequences have been deposited in GenBank (accession numbers: to be included upon acceptance of manuscript).

DNA sequences for both genes were edited in BioEdit v.7.0.5.3 (Hall 1999). Unique sequences were also blasted against data on GenBank (https://blast.ncbi.nlm.nih.gov) and sequences that were highly similar to our data were included in alignments. All sequencing data were aligned using CLUSTAL W (Thompson et al. 1994). Separate phylogenies were reconstructed for the two gene regions using the maximum likelihood search criteria and the MEGAX program (Kumar et al. 2018) and implementing best fit nucleotide substitution models were identified based on the Akaike Information Criterion in JModel Test (Posada, 2008). Topological support for phylogenies was inferred using bootstrapping.

Results

Chemical characterisation of soils

All soils were acidic (pH = 4.9 (mean) ± 0.7 (SD)) and had relatively low levels (mg/kg) of carbon (15.4 ± 6.3), nitrogen (1.1 ± 0.5) and phosphorus (56.7 ± 75.8) (Supplementary Table **S1**) Overall, soil chemistry differed between all three islands and, in the case of La Réunion, sampling locations within an island (Fig. **1**). Location and archipelago (plant species) explained 56% and 15% of the variability in soil chemistry, respectively (PERMANOVA, both factors $P < 0.05$).

Bacterial diversity and distribution

A total of 653 069 sequences (between 20 509-58 217 per sample) were obtained. These were clustered (97% identity cut-off) into 2 508 OTUs (between 516-1 233 OTUs per sample). Of these, 349 OTUs (14%) were unique to Hawaii island samples only, 104 OTUs (4%) from Oahu only, 49 (2%) OTUs from Parc National only and 223 OTUs (9%) from Volcano only; while 679 OTUs (27%) were shared among all four sampling locations. OTUs restricted to Hawaii island accounted for the 0.7% of the sequencing reads, whereas those restricted to Oahu, Parc National and Volcano represented 0.3%, 0.1% and 0.4% of the reads, respectively. Shared OTUs accounted for 91% of the sequencing reads. When grouping OTUs by plant species, 548 were found to be unique to *A. koa* in the Hawaiian Islands, 353 were unique to *A. heterophylla* in La Réunion, while 1 607 OTUs were shared between them. OTUs unique to *A. koa* accounted for 1.5% of the reads and those unique to *A. heterophylla* for 0.8% of the reads. Shared OTUs represented 97.7% of all reads.

There was a statistically significant and positive relationship between occupancy (i.e., the number of soil samples an OTU was found in) and mean relative abundance (Fig. **2**) of all OTUs, 2 271 were identified as being rare and 237 as being abundant i.e. with average relative abundances of < 0.1% and > 0.1%, respectively (Fig. **2**). Rare taxa contributed 90% of the OTU diversity and 23% of the reads. In contrast, abundant taxa contributed only 10% of the OTU diversity but 77% of all reads. This pattern remained the same after rarefaction of sequence reads (Fig. **S1**). In fact, no community metrics changed when using rarefied samples (see online supplementary material) and we therefore only report on analyses based on non-rarefied data here after.

A total of 100 OTUs were identified as potential mutualists; i.e., OTUs belonging to the families *Bradyrhizobiaceae* and *Burkholderiaceae*. Of these, 87 were identified as rare and 13 as abundant. The rare mutualists contributed 3.5% of the OTU diversity and 1.3% of all reads. In contrast, the abundant mutualist taxa contributed 0.5% of the OTU diversity and 6%

of the reads. Among all 2 508 OTUs, we identified 82 as being core; i.e., those present in all soil samples (Fig. **2**, **S1**). These core OTUs contributed 3.2% of the OTU diversity and 42% of all reads. Most core mutualists (76 OTUs) were abundant and few (6 OTUs) were rare. Moreover, six of the 100 potential mutualist (i.e., *Bradyrhizobiaceae* and *Burkholderiaceae*) OTUs were also core OTUs, contributing 0.2% of the OTU diversity and 3% of the reads.

There was no difference in alpha-diversity, neither between locations nor between archipelagos (i.e., plant species) for any of the metrics we considered (Kruskal-Wallis test, $P > 0.05$). Furthermore, diversity was not influenced by any of the soil chemistry attributes we measured (Spearman rho, $P > 0.05$).

Bacterial community structure and composition

The most abundant phyla (Fig. **3**, **S2**) were Acidobacteria (40% of all reads on average), Proteobacteria (33%), Actinobacteria (10%), Bacteroidetes (8%) and Verrucomicrobia (6%). In general, Acidobacteria were abundant on Oahu in the Hawaiian Islands and Parc National on La Réunion island whereas Actinobacteria were abundant on Hawaii island and Bacteroidetes abundant at Volcano on La Réunion island. However, due to the large variability between samples, these differences were not statistically significant (Kruskal-Wallis, $P > 0.05$). Similar trends were found when samples were grouped according to plant species. The abundance of some of these phyla were correlated with soil chemical conditios (Fig. **S3**)

The most common families (Fig. **3**, **S2**) were *Acidobacteriaceae* (18%), *Koribacteraceae* (17%), *Flavobacteriaceae* (7%) and *Hyphomicrobiaceae* (5%). The families *Bradyrhizobiaceae* (3.6%) and *Burkholderiaceae* (3%) were also among the ten most abundant. Similar to the results for phyla, the abundances of these families were quite variable between the samples. Thus, only the relative abundance of the *Bradyrhizobiaceae* was found to be higher in *A. koa* rhizospheres compared with those of *A. heterophylla* (Kruskal-Wallis, $P < 0.05$).

The 82 core taxa resided in 12 different genera (Supplementary file **S8**; Fig. **4**, also see Fig. **S4**), mainly *Rhodoplanes* (seven OTUs, family *Hyphomicrobiaceae*), *Candidatus* Solibacter (six OTUs, family *Solibacteraceae*), *Pedosphera* (four OTUs, family *Pedosphaeraceae*) and *Burkholderia* (four OTUs, Family *Burkholderiaceae*). However, *Flavobacterium* (family *Flavobacteriaceae*) and *Rubrivivax* (family C*omamonadaceae*), one OTU each, were the most abundant, especially in Volcano soils. We note that 57 of the 82 core taxa could not be assigned to a genus.

The relative abundance of potential mutualists (100 OTUs), i.e. those in the *Bradyrhizobiaceae* and *Burkholderiaceae* families possibly capable of nodulation, was much higher in the rhizospheres of *A. koa* than in those of *A. heterophylla* (Kruskal-Wallis test, $P < 0.05$) (Fig. **4b**). Of these OTUs, 33 belonged to the family *Bradyrhizobiaceae*, of which nine were classified as *Bradyrhizobium*, four as *Bosea*, and one as *Balneimonas*. Nineteen *Bradyrhizobiaceae* OTUs could not be classified below the family level. In contrast, 67 potential mutualist OTUs were in the *Burkholderiaceae*, of which 45 belonged to *Burkholderia*, six to *Salinispora*, one to *Pandoraea,* while 15 could not be assigned below the family level. The six core mutualist OTUs (found in all samples) were classified as *Bukholderia* (four OTUs) and *Bradyrhizobiaceae* (two OTUs). Of these core mutualists, five were abundant and one rare.

Rhizosphere bacterial communities (OTU level) were structured into four groups (Fig. **5**) corresponding to sampling location and archipelago (i.e. plant species). These two factors plus phosphorus level explained 37%, 19% and 7% of the variability, respectively (PERMANOVA, all factors $P < 0.05$). Abundant taxa contributed 74.5% to the total differentiation (Bray-Curtis dissimilarities) between samples; whereas rare taxa contributed 24.5% (Fig. **5**, **S5**). However, in some of the comparisons, rare taxa were important components, representing up to 52.1% of total community variation.

Differentially abundant core taxa

Among the 82 core taxa, generalised linear models identified 24 taxa that were differentially abundant (Fig. **6**, **S6**, DESEq, $P < 0.01$) in the rhizospheres of the two tree species. Of these, nine were overrepresented in *A. heterophylla* and 15 in *A. koa*. The taxa overrepresented in *A. heterophylla* were mainly from the families *Koribacteraceae* and *Acidobacteriaceae*. Taxa overrepresented in *A. koa* were also mostly from the *Acidobacteriaceae.* It is noteworthy that two potential mutualist OTUs from the genus *Burkholderia* and one of *Bradyrhizobiaceae* were also more abundant in the rhizospheres of *A. koa* compared with those of *A. heterophylla.*

Quantifying community assembly processes and host selection

Using both and βNTI and RC_{bray} values, we found that the rhizosphere microbial communities were shaped by dispersal limitation (86.7%) (percentage refers to the percentage of pairs of communities that appear to be driven by dispersal limitation), heterogeneous selection (8.3%), undominated (4.1%) and homogenizing dispersal (0.9%).

According to the predictions of the neutral theory (Hubbell, 2006), 15% (369 of 2 508) of the taxa were selected for by the acacias, 8% (215 of 2 508) were selected against by the

acacias (or are dispersal limited), and 77% (1 924 of 2 508) of the taxa showed neutral patterns. Regarding the core taxa, 94% (77 of 82) appear to be positively selected by the acacias and 6% (5 of 82) showed neutral patterns.

Predicted functional profiles

A total of 6 311 KOs were predicted using PICRUSt2. The functional (KOs level) community composition reflected the patterns of the taxonomic (OTU level) community composition (Mantel r = 0.74, P < 0.01). Functional core KOs were, among others, involved in plant defence, draught adaptation, phosphate solubilization and nitrogen cycling (Supplementary file **S7**). Several of these KOs were differentially abundant in one of the two sister acacia species (Fig. **7**).

Root nodule rhizobium phylogenies

The aligned 16S rRNA gene matrix contained 1391 characters and the *nodA* gene matrix 602 characters. Based on 16S rRNA gene BLAST results, the majority of our strains were of the genus *Bradyrhizobium*, while some isolates from *A. koa* were closely related to *Agrobacterium*/*Rhizobium* strains. Three rhizobial strains isolated from *A. heterophylla* belonged to the ß-proteobacterial genus *Paraburkholderia* (formerly *Burkholderia*, Sawana et al. 2014). Some reference strains included in our phylogeny (e.g., *Bradyrhizobium guangdongense* and *B. ganzhouense;* Fig. **8**) were previously isolated from *Acacia melanoxylon* (Lu et al. 2014; Li et al. 2015), and shared high, or in some instance 100%, sequence identity to some of our isolated strains (Fig. **8**). Alignments including 16S rRNA gene sequences from our NGS data corroborated that most of the *Bradyrhizobium* and *Burkholderia* mutualist OTUs in rhizospheres were not closely related (< 97% identity) to the rhizobia found in root nodules of the two acacias sampled. The exceptions were OTU 872 and OTU 890 (two core OTUs evenly distributed among the samples) that shared high sequence identity (> 98%) with rhizobia we isolated from root nodules of *A. koa* (e.g., isolate AK2_1) and both *A. koa* and *A. heterophylla* (e.g., isolates Ah1_3b and AK7_4, respectively; results not shown). Based on the reference strains included in our phylogeny, OTU 872 shared the highest DNA identity with *Bradyrhizobium elkanii*, while OTU 890 showed high identity to numerous *Bradyrhizobium* species (i.e., *B. japonicum*, *B. guangdongense*, *B. cytisis*, *B. lupini*, *B. canariense*, *B. cytisi*, *B. rifense* and *B. ganzhouense*).

The *nodA* BLAST results revealed that all *A. heterophylla* and *A. koa* strains belonged to the genus *Bradyrhizobium*, including those identified as *Paraburkholderia* based on our 16S

rRNA gene phylogeny. The *nodA* phylogeny identified some *A. heterophylla* strains as the same as those previously isolated from *A. melanoxylon* in Australia (Fig. **8**, Warrington et al. 2019). Some *A. koa* isolates (e.g., strain AK26.1) showed high phylogenetic relatedness to *Bradyrhizobium* strains previously isolated from *Acacia longifolia* in Portugal (e.g., strain U12 EU884543, A113 EU884533; Rodríguez-Echeverría 2010). The majority of rhizobial strains isolated from both acacias clustered in a distinct *NodA* clade, not closely related to any known *Bradyrhizobium* reference strains, and likely represent a new species.

Discussion

We analysed the rhizosphere bacterial communities and rhizobia from root nodules of two sister *Acacia* species found on islands at opposite ends of the world. *Acacia heterophylla* populations on La Réunion island are the descendents of an ancestor that originated directly from the Hawaiian Islands where *A. koa* is endemic (Le Roux et al. 2014). These two species have been isolated for *c*. 1.4 million years (Le Roux et al. 2014). While biogeography largely influenced the structure and composition, but not function, of the rhizosphere communities of these two *Acacia* species, they appear to select for remarkably similar core rhizosphere microbiomes, including their rhizobial mutualists.

The composition and function of rhizosphere bacterial communities associated with Acacia heterophylla *and* A. koa

The majority of the rhizosphere taxa associated with *A. heterophylla* and *A. koa* were found only in one or few soil samples and were rare. It has been suggested that rare taxa may represent slow growers with small cell sizes, which allows them to avoid predation and viral lysis (Pedrós-Alió, 2006). Despite their low abundance, rare taxa often represent a vast functional gene pool (Jousset et al. 2017). For example, rare plant-associated microbes have been found to produce volatile compounds that protect host plants against fungal pathogens (Hol et al. 2015).

Some cosmopolitan and abundant taxa were also identified in the rhizospheres of the two *Acacia* species, and these contributed the most to community variation. This seems to be a common attribute of soil bacterial communities (Delgado-Baquerizo et al. 2018). Positive relationships between mean relative abundance and occupancy have been previously observed (Burns et al. 2016; Kurm et al. 2017) and conforms to the predictions of the neutral theory for prokaryotes (Sloan et al. 2006). This is because taxa that are abundant are more likely to disperse by chance, while those that are rare have a higher chance of being lost from

communities due to ecological drift. We note that many of the cosmopolitan and abundant taxa identified in this study do not have known close relatives. Taken together, our results suggest that the rhizosphere bacterial communities of *A. heterophylla* and *A. koa* are dominated by a relatively small number of cosmopolitan taxa, some of which need to be described. These findings are in agreement with a recent analyses based on hundreds of soils samples from around the world, that found around 2% of bacterial taxa to account for almost half of all soil bacterial communities worldwide (Delgado-Baquerizo et al. 2018).

The rhizosphere microbial communities of both *Acacia* species were highly variable (average: 12% similarity between soil samples) and separated into four distinct clusters according to sampling location; i.e., the communities showed strong biogeographic structure. Indeed, we found that the major ecological processes shaping these communities to be dispersal limitation (86.7%), followed by variable (heterogeneous) selection (8.3%). Dispersal limitation implies that the movement of microbes between locations is restricted, whereas variable selection results from heterogeneous abiotic and biotic environmental conditions (Zhou and Ning, 2017). Barriers to dispersal are obvious in our case, oceans between all studied islands or high geographic distances between sampling sites within La Réunion island. In both archipelagos, we also sampled soils that differed in key aspects. In the Hawaiian Islands, soils were collected on the older island of Oahu (*c.* 3.7 Myr) and the younger Hawaii island (*c.* 0.43 Myr). In La Réunion island we sampled soils on an extinct volcano (Parc National, Bébour forest) an in close proximity to an active volcano (Volcano, Piton de la Fournaise). These differences within and between archipelagos likely contributed to heterogeneus selection on soil microbial communities.

In contrast to taxonomic composition, the predicted functional composition was less variable (85% similarity between all soil samples), although it differed between locations and between the two archipelagos (i.e., between host species). Several of the predicted functions are involved in processes that promote disease suppression/plant defence, hormone balance and nitrogen cycling, which are common in plant rhizosphere communities and may play major roles in plant growth. Since rhizosphere microbial communites were highly variable, these findings suggest that functional similarity rather than taxonomic similarity shapes microbial community assembly in the rhizosphere of these acacias, which agrees with previous observations (Ofek-Lalzar et al. 2014).

The occupancy of a large proportion of the core taxa (94%) in the rhizospheres of both acacias was higher than predicted by the neutral theory, suggesting that the rhizosphere environment created by these two *Acacia* species selects for certain bacterial taxa. Among these

were several well-known root-associated taxa such as *Bradyrhizobium*, *Burkholderia* and *Flavobacterium*. The genera *Bradyrhizobium* and *Burkholderia* contain many species known to enhance plant performance, primarily through nitrogen fixation (Yeoh et al. 2017). *Flavobacterium* is abundant in rhizospheres and seems to increase plant resistance to pathogens (Kolton et al. 2016). For example, *Flavobacterium johnsoniae* strain GSE09 has biocontrol activity against pathogenic fungi (Sang and Kim, 2012). An interesting finding was that a core *Flavobacterium* OTU was overrepresented in Volcano soils in La Réunion island. This suggests that some core taxa are conditionally rare and only become abundant when optimal growth conditions prevail. Estimates suggest that 1.5-28% of microbes can be conditionally rare and that they are essential to understanding community assembly and function (Shade et al. 2014). Whether or not the core taxa identified here positivily interact with these two acacias remains unknown and warrants further investigation.

A shortcoming of our study is the fact that rarity/abundance thresholds and/or the way taxonomic units are defined (e.g., OTUs, zOTUs, ASVs) will ultimately influence the number of common/rare OTUs. However, dividing taxa as common or rare is useful for defining ecological categories that offer additional information from those defined by phylogeny, taxonomy or functional capacity (Magurran and Henderson, 2003). We also do not know how similar or dissimilar bulk soil bacterial communities are in our sites. For example, if bulk soil communities in different locations are profoundly different, then the notion that these two acacias select for the same (or functionally similar) rhizosphere communities is futher supported. On the other hand, if acacia rhizosphere communities at our different sites simply reflect a subset of local bulk soil communities, then selectivity by these two acacias is probably less important in shaping their rhizosphere communities. Previous studies on acacias and their associated rhizobia, however, support host selection rather than stochasticity as underlying our findings. For example, Kamutando et al. (2017) compared rhizosphere microbial communities of invasive *A. dealbata* and bulk soils. Similar to our findings, these authors found *Bradyrhizobium* and *Burkholderia* OTUs to form part of the species' core rhizosphere microbiome (Kamutando et al. 2017). A follow-up metagenomic analyses confirmed *Bradyrhizobium* functional genes to be over-represented in rhizosphere soils of the species (Kamutando et al. 2019). Le Roux et al. (2016, 2018) also characterised rhizobium communities in soils and nodules of various legumes in acacia-invaded (by six *Acacia* species) and uninvaded soils in South Africa. They found that dense acacia stands homogenised rhizobial community structure and enriched rhizosphere soils for *Bradyrhizobium* strains that

nodulate them (Le Roux et al. 2016, 2018). Therefore, at least from a mutualist perspective, it appears that acacias do exert strong selection on soil bacteria.

Rhizobial mutualists associated with Acacia heterophylla *and* A. koa

Using Sanger-sequencing data we found *Bradyrhizobium* strains to be present in root nodules of both *A. heterophylla* and *A. koa*. These observations support the hypothesis that these two acacias select for the same/similar mutualists (also see Le Roux et al. 2016; Keet et al. 2017; Le Roux et al. 2018; Warrington et al. 2019). In Australia, hundreds of native *Acacia* species are primarily nodulated by an endemic lineage of bradyrhizobia (Richardson et al. 2011; Mishler et al. 2014), mainly from the so-called Clade I *Bradyrhizobium* (*sensu* Stępkowski et al. 2018). Clade I bradyrhizobia have also been recorded from areas outside Australia, but always in association with invasive acacias (e.g., Rodríguez-Echeverría S. 2010; Ndlovu et al. 2013; Warrington et al. 2019), suggesting that these bacteria have been co- introduced with their Australian host plants to many parts of the world. For example, Warrington et al. (2019) recently found *A. melanoxylon* to nodulate with Clade I bradyrhizobia in New Zealand and South Africa. These bacteria have also been isolated from the root nodules of invasive *A. longifolia* in Portugal (Rodríguez-Echeverría 2010). Our phylogenetic analyses suggest that the specificity of acacias to Clade I bradyrhizobia also holds, to some degree, for indigenous species found outside Australia. We also isolated fast-growing strains of *Rhizobium* and *Paraburkholderia* from the root nodules of *A. koa*, some of which carried *Bradyrhizobium* nodulation genes, suggestive of horizontal gene transfer (HGT). As far as we know, prior to our study, only two HGT events between α- and β-proteobacterial rhizobia have been known (Lemaire et al. 2015).

Although the 16S rRNA gene region is useful for identifying rhizobia, it rarely provides information about the specificity/functionality of the symbiosis. Nodulation genes play a role in host specificity and can, to some extent, provide such information (Le Roux et al. 2017).

Our phylogeny revealed that all *A. heterophylla* and *A. koa* root nodulating bacteria carried *Bradyrhizobium nodA* genes, many of which shared high sequence identity with those of Clade I bradyrhizobia strains previously isolated from other *Acacia* species (Rodríguez-Echeverría 2010; Warrington et al. 2019). Some of our isolates (e.g., strain AK 7.2) that belonged to the same clade, but did not share high sequence homology with any known *Bradyrhizobium* strains. Whether La Réunion island and the Hawaiian Islands form part of the natural distribution of Clade I bradyrhizobia, or whether these bacteria have been recently introduced to these islands from Australia, could not be determined from the results of this study.

Nevertheless, invasive Australian acacias are present in both archipelagos (Richardson et al. 2011) and therefore the possibility that Clade I bradyrhizobia have been co-introduced with these invasive acacias cannot be ruled out, and this warrants further work.

Conclusion

While the overall taxonomic and, to a much lesser extent, predicted functions of rhizosphere bacterial communities associated with *A. heterophylla* and *A. koa* are largely influenced by dispersal limitation and local abiotic conditions, we also found evidence for the existence of a core of rhizosphere taxa (including rhizobial mutualists with low abundance) that is selected by these two plant species. This suggests remarkable host plant selectivity, especially for bradyhizobium mutualists, over extremely wide geographic ranges. Other plant congeners, or even conspecifics, with similarly disjunct distributions provide exciting oppurtunities to test the generality of such strong host plant selection. For example, *Sophora chrysophylla* and *S. denudate* are endemic to the Hawaiian Islands and La Réunion island, respectively. Numerous *Sophora* species are also found on oceanic islands throughout the Southern Hemisphere (Shepherd and Heenan 2017). The Hawaiian and Mascarene archipelagos also share conspecifics. For example, both archipelagos are home to *Dodonea viscosa*, a species that is native in numerous parts of the world (Harrington and Gadek 2009). Plants such as *Sophora* species and *D. viscosa* provide ideal study systems to further tease apart the roles of biogeography, stochastic and deterministic processes, and plant identity in shaping the diversity, structure and function of rhizosphere communities, especially symbionts.

Acknowledgements

We thank Jesse Eiben for lab space on Hawaii island, Marelongue field research station (OSU Réunion) for research accommodation on La Réunion, and Réunion National Park for collecting permits. Megan Koordom at Stellenbosch University provided invaluable lab support. PWC and MJW acknowledge support from the DSI-NRF Centre of Excellence in Plant Health Biotechnology. JJLeR and DMR acknowledge the DSI-NRF Centre of Excellence for Invasion Biology. DMR received additional support from the Oppenheimer Memorial Trust (grant 18576/03).

References

Banerjee, S., Kirkby, C.A., Schmutter, D., Bissett, A., Kirkegaard, J.A., Richardson, A.E., 2016. Network analysis reveals functional redundancy and keystone taxa amongst

bacterial and fungal communities during organic matter decomposition in an arable soil. Soil Biology and Biochemistry 97, 188–198.

Brundrett, M.C., Tedersoo, L., 2018. Evolutionary history of mycorrhizal symbioses and global host plant diversity. New Phytologist 220, 1108–1115.

Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E.V.L., Schulze-Lefert, P., 2013. Structure and Functions of the Bacterial Microbiota of Plants, in: Merchant, S.S. (Ed.), Annual Review of Plant Biology, Vol 64, Annual Review of Plant Biology. pp. 807–838.

Burns, A.R., Stephens, W.Z., Stagaman, K., Wong, S., Rawls, J.F., Guillemin, K., Bohannan, B.J.M., 2016. Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. ISME Journal 10, 655–664.

Campbell B.J. (2014) The Family Acidobacteriaceae. In: Rosenberg E., DeLong E.F., Lory S., Stackebrandt E., Thompson F. (eds) The Prokaryotes. Springer, Berlin, Heidelberg.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Tumbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7, 335–336.

Coulaud, J., Brown, S.C., Siljak-Yakovlev, S., 1995. First cytogenetic investigation in populations of Acacia heterophylla, endemic from La Réunion Island, with reference to *A. melanoxylon*. Annals of Botany 75, 95–100.

Crisóstomo, J.A., Rodríguez-Echeverría, S., Freitas, H., 2013. Co-introduction of exotic rhizobia to the rhizosphere of the invasive legume *Acacia saligna*, an intercontinental study. Applied Soil Ecology 64, 118–126.

Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-González, A., Eldridge, D. J., Bardgett, R. D., Maestre, F. T., Singh, B. K., Fierer, N., 2018. A global atlas of the dominant bacteria found in soil. Science 359, 320–325.

Dini-Andreote, F., Brossi, M.J.L., van Elsas, J.D., Salles, J.F., 2016. Reconstructing the genetic potential of the microbially-mediated nitrogen cycle in a salt marsh ecosystem. Frontiers in Microbiology 7, 902.

Douglas, G.M., Maffei, V.J., Zaneveld, J.R., Yurgel, S.N., Brown, J.R., Taylor, C.M., Huttenhower, C., Langille, M.G.I., 2020. PICRUSt2 for prediction of metagenome functions. Nature Biotechnology 38, 685–688.

Fierer, N., Leff, J.W., Adams, B.J., Nielsen, U.N., Bates, S.T., Lauber, C.L., Owens, S., Gilbert, J.A., Wall, D.H., Caporaso, J.G., 2012. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. Proceedings of the National Academy of Sciences of the United States of America 109, 21390–5.

Fierer, N., 2017. Embracing the unknown: Disentangling the complexities of the soil microbiome. Nature Reviews Microbiology 15, 579–590.

Fitzpatrick, C.R., Copeland, J., Wang, P.W., Guttman, D.S., Kotanen, P.M., Johnson, M.T.J., 2018. Assembly and ecological function of the root microbiome across angiosperm plant species. Proceedings of the National Academy of Sciences, USA 115, E1157–E1165.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95／98／NT. Nucleic Acids Symposium Series 41, 95–98.

Harrington, M.G., Gadek, P.A., 2009. A species well travelled – the *Dodonaea viscosa* (Sapindaceae) complex based on phylogenetic analyses of nuclear ribosomal ITS and ETSf sequences. Journal of Biogeography 36, 2313–2323.

Haukka, K., Lindström, K., Young, J.P.W., 1998. Three phylogenetic groups of *nodA* and *nifH* genes in *Sinorhizobium* and *Mesorhizobium* isolates from leguminous trees growing in Africa and Latin America. Applied and Environmental Microbiology 64, 419–426.

Hol, W.H.G., Garbeva, P., Hordijk, C.A., Hundscheid, M.P.J., Klein Gunnewiek, P.J.A., van Agtmaal, M., Kuramae, E.E., de Boer, W., 2015. Non-random species loss in bacterial communities reduces antifungal volatile production. Ecology 96, 2042–2048.

Hubbell, S.P., 2006. Neutral theory and the evolution of ecological equivalence. Ecology 87, 1387–1398.

Jousset, A., Bienhold, C., Chatzinotas, A., Gallien, L., Gobet, A., Kurm, V., Küsel, K., Rillig, M.C., Rivett, D.W., Salles, J.F., Van Der Heijden, M.G.A., Youssef, N.H., Zhang, X., Wei, Z., Hol, G.W.H., 2017. Where less may be more: How the rare biosphere pulls ecosystems strings. ISME Journal 11, 853–862.

Kamutando, C.N., Vikram, S., Kamgan-Nkuekam, G., Makhalanyane, T.P., Greve, M., Le Roux, J.J., Richardson, D.M., Cowan, D., Valverde, A., 2017. Soil nutritional status and biogeography influence rhizosphere microbial communities associated with the invasive tree *Acacia dealbata*. Scientific Reports 7, 6472.

Kolton, M., Erlacher, A., Berg, G., Cytryn, E., 2016. The *Flavobacterium* genus in the plant holobiont: ecological, physiological, and applicative insights *In* Microbial models: from environmental to industrial sustainability (pp.189-207) Springer Singapore, Editors: Susana Castro Sowinski

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution 35, 1547–1549.

Kurm, V., Van Der Putten, W.H., De Boer, W., Naus-Wiezer, S., Gera Hol, W.H., 2017. Low abundant soil bacteria can be metabolically versatile and fast growing. Ecology 98, 555–564.

Lane, D.J., 1991. 16S/23S rRNA sequencing. In: Stackebrandt E, Goodfellow M (eds) Nucleic acid techniques in bacterial systematics, Wiley, Chichester, pp 115–175.

Lemaire, B., Dlodlo, O., Chimphango, S., Stirton, C., Schrire, B., Boatwright, J.S., Honnay, O., Smets, E., Sprent, J., James, E.K., Muasya, A.M., 2015. Symbiotic diversity, specificity and distribution of rhizobia in native legumes of the Core Cape Subregion (South Africa). FEMS Microbiology Ecology 91, fiv118.

Lemanceau, P., Blouin, M., Muller, D., Moënne-Loccoz, Y., 2017. Let the core microbiota be functional. Trends in Plant Science 22, 583–595.

Le Roux, J.J., Ellis, A.G., van Zyl, L.M., Hosking, N.D., Keet, J.-H., Yannelli, F.A., 2018. Importance of soil legacy effects and successful mutualistic interactions during Australian acacia invasions in nutrient-poor environments. Journal of Ecology 106, 2071–2081.

Le Roux, J.J., Hui, C., Keet, J.-H., Ellis, A.G., 2017, Co-introduction vs ecological fitting as pathways to the establishment of effective mutualisms during biological invasions. New Phytologist 215, 1354–1360.

Le Roux, J.J., Mavengere, N.R., Ellis, A.G., 2016. The structure of legume–rhizobium interaction networks and their response to tree invasions. AoB PLANTS 8, plw038.

Le Roux, J.J., Strasberg, D., Rouget, M., Morden, C.W., Koordom, M., Richardson, D.M., 2014. Relatedness defies biogeography: the tale of two island endemics (*Acacia heterophylla* and *A. koa*). New Phytologist 204, 230–242.

Li, Y.H., Wang, R., Zhang, Young, J.P.W., Wang, E.T., Sui, X.H., Chen, W.X., 2015. *Bradyrhizobium guangdongense* sp. nov. and *Bradyrhizobium guangxiense* sp. nov., isolated from effective nodules of peanut. International Journal of Systematic and Evolutionary Microbiology 65, 4655–4661.

Louca, S., Jacques, S., Pires, A., Leal, J.S., Srivastava, D.S., Wegener Parfrey, L., Farjalla, V.F., Doebeli, M., 2017. High taxonomic variability despite stable functional structure across microbial communities. Nature Ecology and Evolution 1, 0015.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15, 550.

Lu, J.K., Dou, Y.J., Zhu, Y.J., Wang, S.K., Sui, X.H., Kang, L.H., 2014. *Bradyrhizobium ganzhouense* sp. nov., an effective symbiotic bacterium isolated from *Acacia melanoxylon* R. Br. nodules. International Journal of Systematic and Evolutionary Microbiology 64, 1900–1905.

Magurran, A.E., Henderson, P.A., 2003. Explaining the excess of rare species in natural species abundance distributions. Nature 422, 714–716.

Martín-Robles, N., Lehmann, A., Seco, E., Aroca, R., Rillig, M. C., Milla, R., 2018. Impacts of domestication on the arbuscular mycorrhizal symbiosis of 27 crop species. New Phytologist 218, 322–334.

Mendes, R., Kruijt, M., de Bruijn, I., Dekkers, E., van der Voort, M., Schneider, J.H., Piceno, Y.M., DeSantis, T.Z., Andersen, G.L., Bakker, P.A.H.M., Raaijmakers, J.M., 2011. Deciphering the rhizosphere microbiome for disease-suppressive bacteria. Science 332, 6033.

Mishler, B.D., Knerr, N., González-Orozco, C.E., Thornhill, A.H., Laffan, S.W., Miller, J.T., 2014. Phylogenetic measures of biodiversity and neo- and paleo-endemism in Australian *Acacia*. Nature Communications 5, 4473.

Moulin, L., Munive, A., Dreyfus, B., Boivin-Masson, C., 2001. Nodulation of legumes by members of the β-subclass of Proteobacteria. Nature 411, 948–950.

Ndlovu, J., Richardson, D.M., Wilson, J.R.U., Le Roux, J.J., 2013. Co-invasion of South African ecosystems by an Australian legume and its rhizobial symbionts. Journal of Biogeography 40, 1240–1251.

Ning, D., Yuan, M., Wu, L., Zhang, Y., Guo, X., Zhou, X., Yang, Y., Arkin, A.P., Firestone, M.K., Zhou, J., 2020. A quantitative framework reveals ecological drivers of grassland microbial community assembly in response to warming. Nature Communications 11.

Ofek-Lalzar, M., Sela, N., Goldman-Voronov, M., Green, S.J., Hadar, Y., Minz, D., 2014. Niche and host associated functional signatures of the root surface microbiome. Nature Communications 5, 4950. Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M., Wagner, H., 2013. Vegan: Community Ecology Package.

Parks, D.H., Beiko, R.G., 2010. Identifying biologically relevant differences between metagenomic communities. Bioinformatics 26, 715–721.

Pascale, A., Proietti, S., Pantelides, I.S., Stringlis, I.A., 2020. Modulation of the root microbiome by plant molecules: the basis for targeted disease suppression and plant growth promotion. Frontiers in Plant Science 10, 1741.

Pedrós-Alió, C., 2012. The Rare Bacterial Biosphere. Annual Review of Marine Science, Annual Review of Marine Science 4, 449–466.

Pedrós-Alió, C., 2006. Marine microbial diversity: can it be determined? Trends in Microbiology 14, 257–263.

Philippot, L., Raaijmakers, J.M., Lemanceau, P., van der Putten, W.H., 2013. Going back to the roots: the microbial ecology of the rhizosphere. Nature Reviews Microbiology 11, 789–799.

Posada, D., 2008. jModelTest: phylogenetic model averaging. Molecular Biology and Evolution 25, 1253–1256.

R Development Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; http://www.r-project.org

Ramoneda, J., Le Roux, J.J., Frossard, E., Bester, C., Oettlé, N., Frey, B., Gamper, H.A., 2019. Insights from invasion ecology: can consideration of eco-evolutionary experience promote benefits from rootmutualisms in plant production? AoB Plants 11, plz060.

Richardson, D.M., Carruthers, J., Hui, C., Impson, F.A.C., Miller, J.T., Robertson, M.P., Rouget, M., Le Roux, J.J., Wilson, J.R.U., 2011. Human-mediated introductions of Australian acacias – a global experiment in biogeography. Diversity and Distributions 17, 771–787.

Rodríguez-Echeverría, S., 2010. Rhizobial hitchhikers from down under: invasional meltdown in a plant–bacteria mutualism? Journal of Biogeography 37, 1611–1622.

Rodríguez-Echeverría, S., Le Roux, J.J., Crisóstomo, J.A., Ndlovu, J., 2011. Jack-of-all-trades and master of many? How does associated rhizobial diversity influence the colonization success of Australian acacias? Diversity and Distributions 17, 946–957.

Sang, M.K., Kim, K.D., 2012. The volatile-producing *Flavobacterium johnsoniae* strain GSE09 shows biocontrol activity against *Phytophthora capsici* in pepper. Journal of Applied Microbiology 113, 383–398.

Sawana, A., Adeolu, M., Gupta, R. S., 2014. Molecular signatures and phylogenomic analysis of the genus *Burkholderia*: proposal for division of this genus into the emended genus *Burkholderia* containing pathogenic organisms and a new genus *Paraburkholderia* gen. nov. harboring environmental species. Frontiers in Genetics 5, 429.

Shade, A., Jones, S.E., Caporaso, J.G., Handelsman, J., Knight, R., Fierer, N., Gilbert, J.A., 2014. Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. MBio 5, e01371–14.

Shade, A., Stopnisek, N., 2019. Abundance-occupancy distributions to prioritize plant core microbiome membership. Current Opinion in Microbiology 49, 50–58.

Shepherd, L.D., Heenan, P.B., 2017. Evidence for both long-distance dispersal and isolation in the Southern Oceans: molecular phylogeny of *Sophora* sect Edwardsia (Fabaceae). New Zealand Journal of Botany 55, 334–346.

Sloan, W.T., Lunn, M., Woodcock, S., Head, I.M., Nee, S., Curtis, T.P., 2006. Quantifying the roles of immigration and chance in shaping prokaryote community structure. Environmental Microbiology 8, 732–740.

Stegen, J.C., Lin, X.J., Konopka, A.E., Fredrickson, J.K., 2012. Stochastic and deterministic assembly processes in subsurface microbial communities. ISME Journal 6, 1653–1664.

Stępkowski, T., Banasiewicz, J., Granada, C.E., Andrews, M., Passaglia, L.M.P., 2018. Phylogeny and phylogeography of rhizobial symbionts nodulating legumes of the tribe Genisteae. Genes 9, 163.

Somasegaran, P., Hoben, H.J., 1994. Handbook for Rhizobia: Methods in legume–rhizobium technology. Springer-Verlag, New York, USA.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673–4680.

van der Heijden, M.G.A., Bardgett, R.D., van Straalen, N.M., 2008. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. Ecology Letters 11, 296–310.

Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A. and Dufresne, A., 2015. The importance of the microbiome of the plant holobiont. New Phytologist 206, 1196–1206.

Warrington, S., Ellis, A., Novoa, A., Wandrag, E.M., Hulme, P.E., Duncan, R.O., Valentine, A., Le Roux, J.J., 2019. Cointroductions of Australian acacias and their rhizobial mutualists in the Southern Hemisphere. Journal of Biogeography 46, 1519–1531.

Vitousek, P.M., Walker, L.R., 1989. Biological invasion by *Myrica faya* in Hawaii: plant demography, nitrogen fixation, ecosystem effects. Ecological Monographs 59, 247–265.

Yang, S., 2020. otuSummary: Summarizing OTU table regarding the composition, abundance and beta diversity of abundant and rare biospheres. R package version 0.1.1. https://CRAN.R-project.org/package=otuSummary.

Yeoh, Y.K., Dennis, P.G., Paungfoo-Lonhienne, C., Weber, L., Brackin, R., Ragan, M.A., Schmidt, S., Hugenholtz, P., 2017. Evolutionary conservation of a core root microbiome across plant phyla along a tropical soil chronosequence. Nature Communications 8, 215.

Zhou, J., Ning, D., 2017. Stochastic community assembly: Does it matter in microbial ecology? Microbiology and Molecular Biology Reviews 81, e00002-17.

Figures

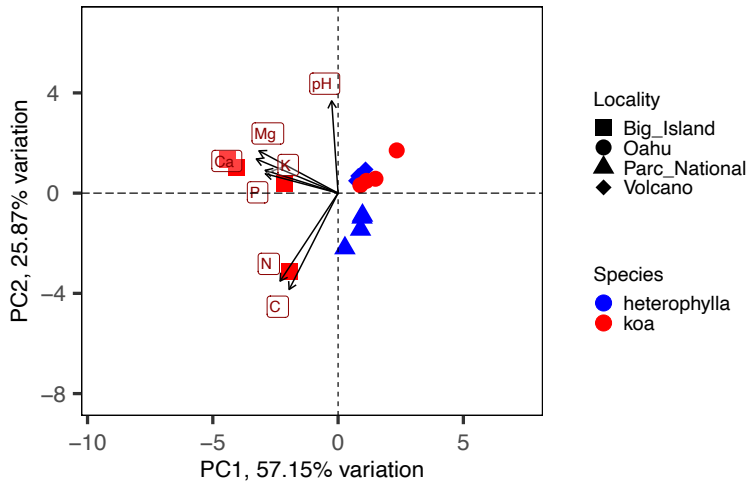

Fig. 1: PCoA of the Euclidean distance matrix (standardised) of soil chemistry for samples form root zones of *Acacia heterophylla* (Hawaiian Islands: Ohau, Hawaii island) and *A. koa* (La Réunion island: Parc National, Volcano). The samples were significantly separated by sampling location and archipelago of origin (i.e., plant species).

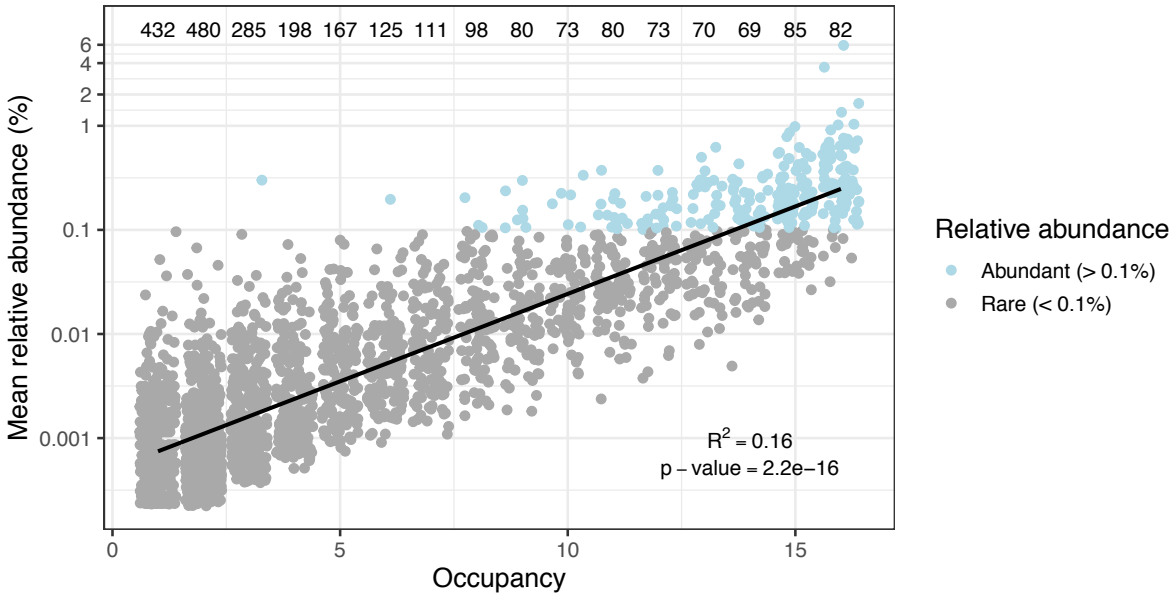

Fig. 2: Linear regression of mean relative abundance and occupancy plot with rare and abundant rhizobia taxa (OTU level) from two *Acacia* species on the Hawaiian Islands in the Pacific Ocean and La Réunion island in the Indian Ocean. Figures at the upper edge of the graph indicate the number of OTUs that occurred in each of the 16 soil samples.

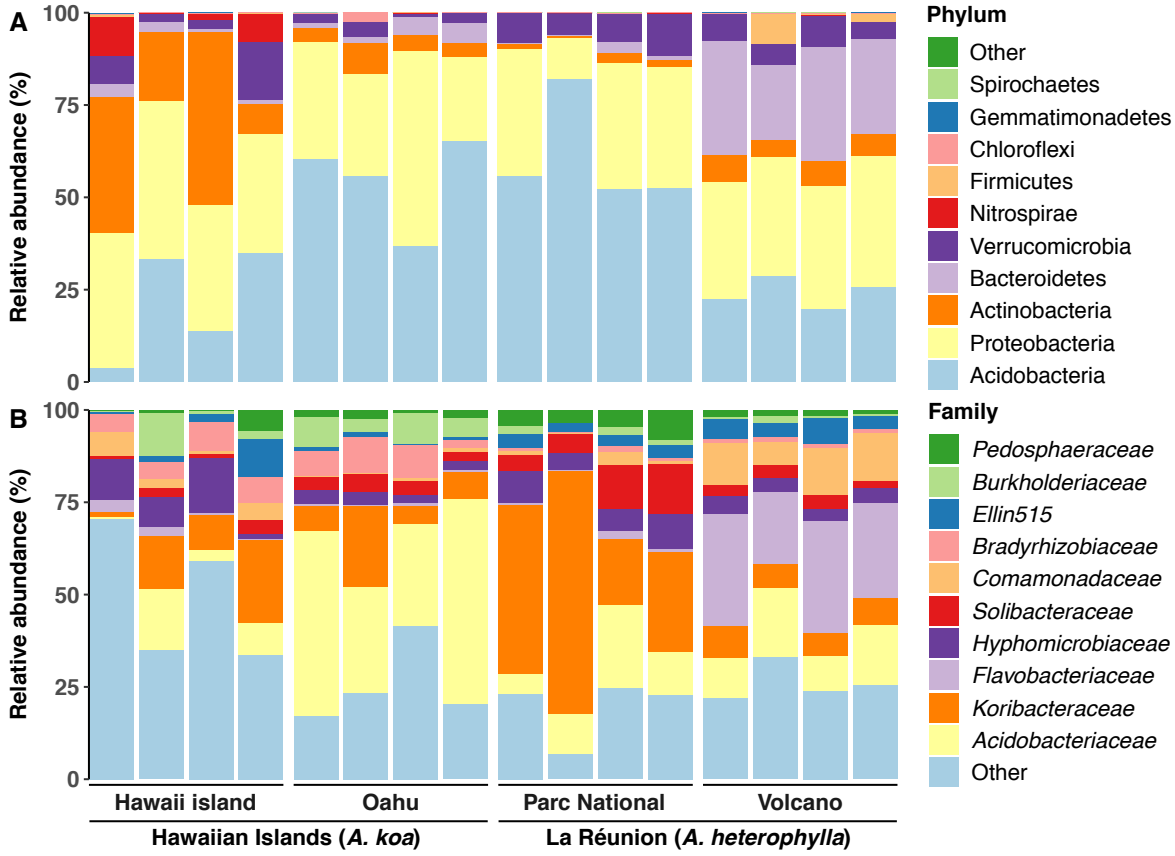

Fig. 3: Relative abundance plots of the most abundant taxa from the root zone of two *Acacia* species and two locations on each of the Hawaiian Islands and La Réunion, across samples at the phylum (a) and family (b) level. The group "other" represents unclassified and minor taxa.

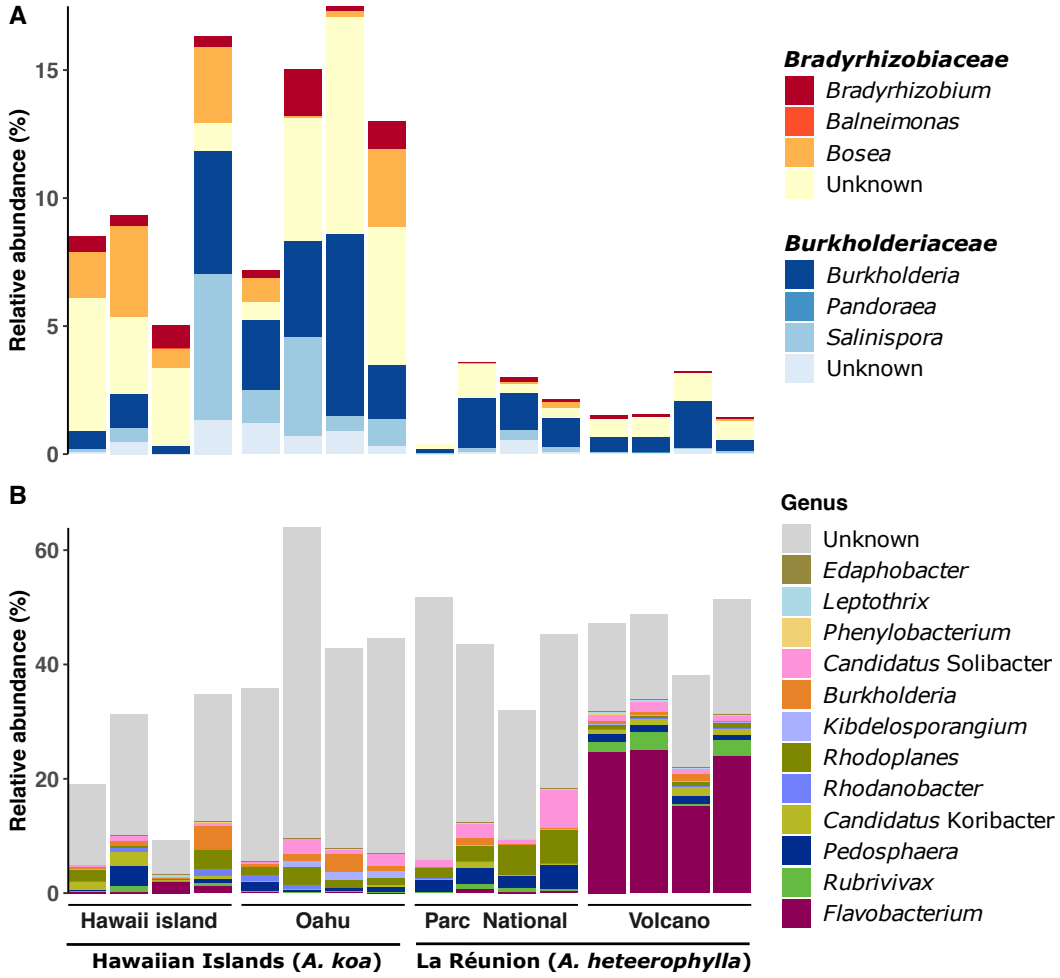

Fig. 4: Relative abundance of (a) core and (b) mutualist (*Bradyrhizobiaceae* and *Burkholderiaceae*) taxa in different rhizosphere soils of *Acacia heterophylla* and *A. koa* collected from different locations on the Hawaiian Islands and La Réunion island. Core taxa contributed 42% of the reads; whereas mutualist taxa contributed 7.3% of the reads.

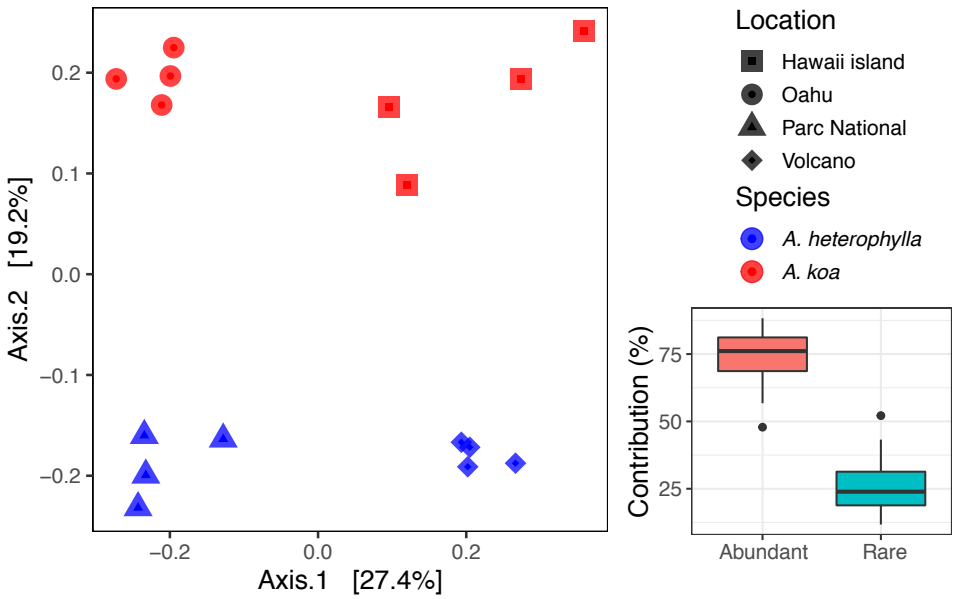

Fig 5: (a) PCoA of the Bray-Curtis distance matrix (after relative abundance transformation). The samples separated by sampling location and archipelago (or plant species) as in figure 1. (b) Boxplot showing quantile summary of the proportional contribution of the abundant and rare taxa to the total Bray-Curtis distance matrices.

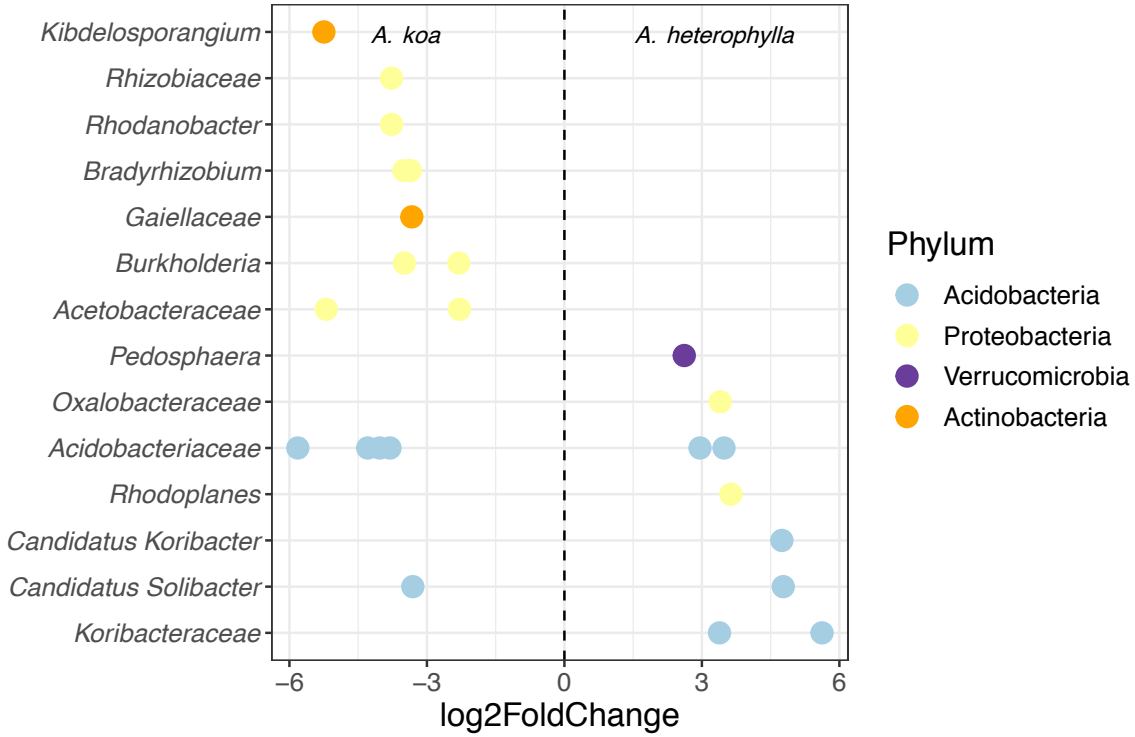

Fig 6: Differential core taxa (OTUs) abundance (DESeq2 analysis) in the rhizospheres of *A. heterophylla* and *A. koa*. Core taxa with positive $\log_2$ fold change values were more abundant in the rhizospheres of *A. heterophylla*, whereas core taxa with negative $\log_2$ fold change values were more abundant in the rhizospheres of *A. koa*.

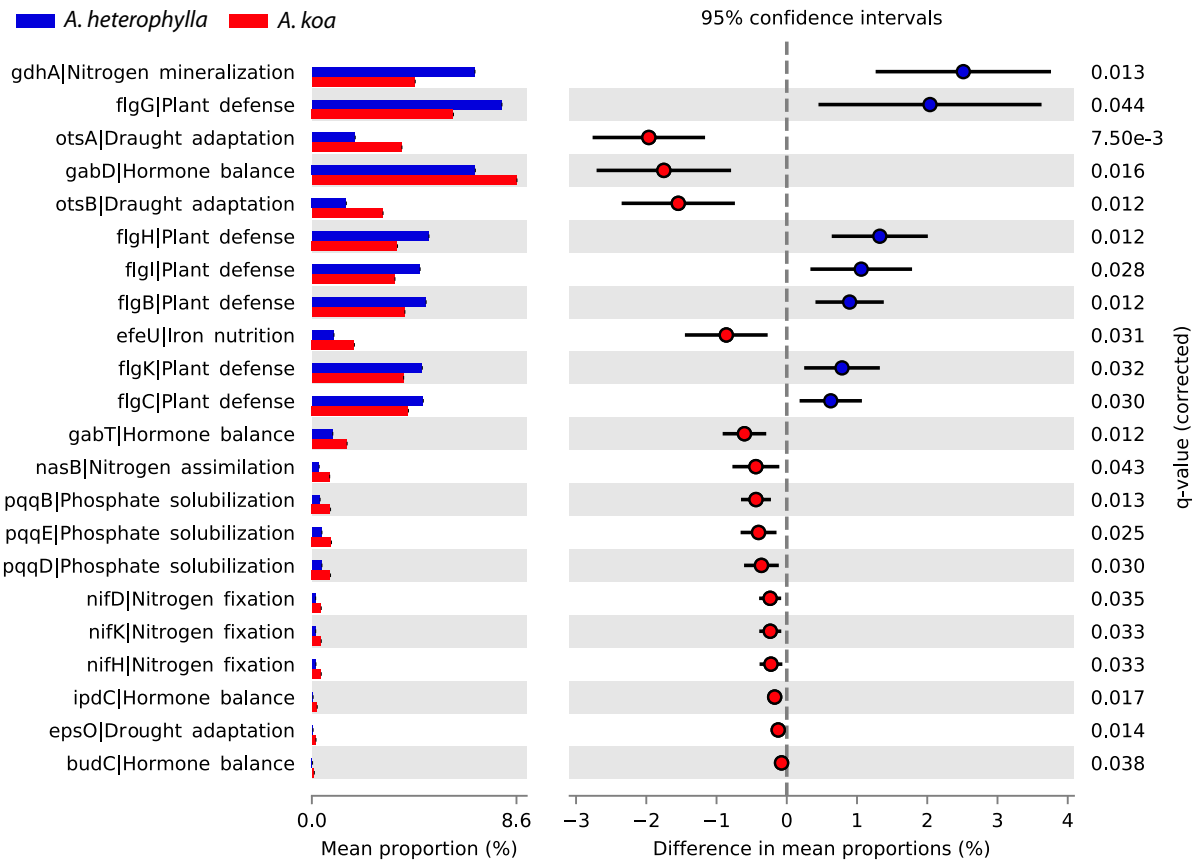

Fig 7: Differences in abundance with functional groups, KO, abundance (mean ± 95% CI) between *Aacacia heterophylla* (La Réunion island) and *A. koa* (Hawaiian Islands).

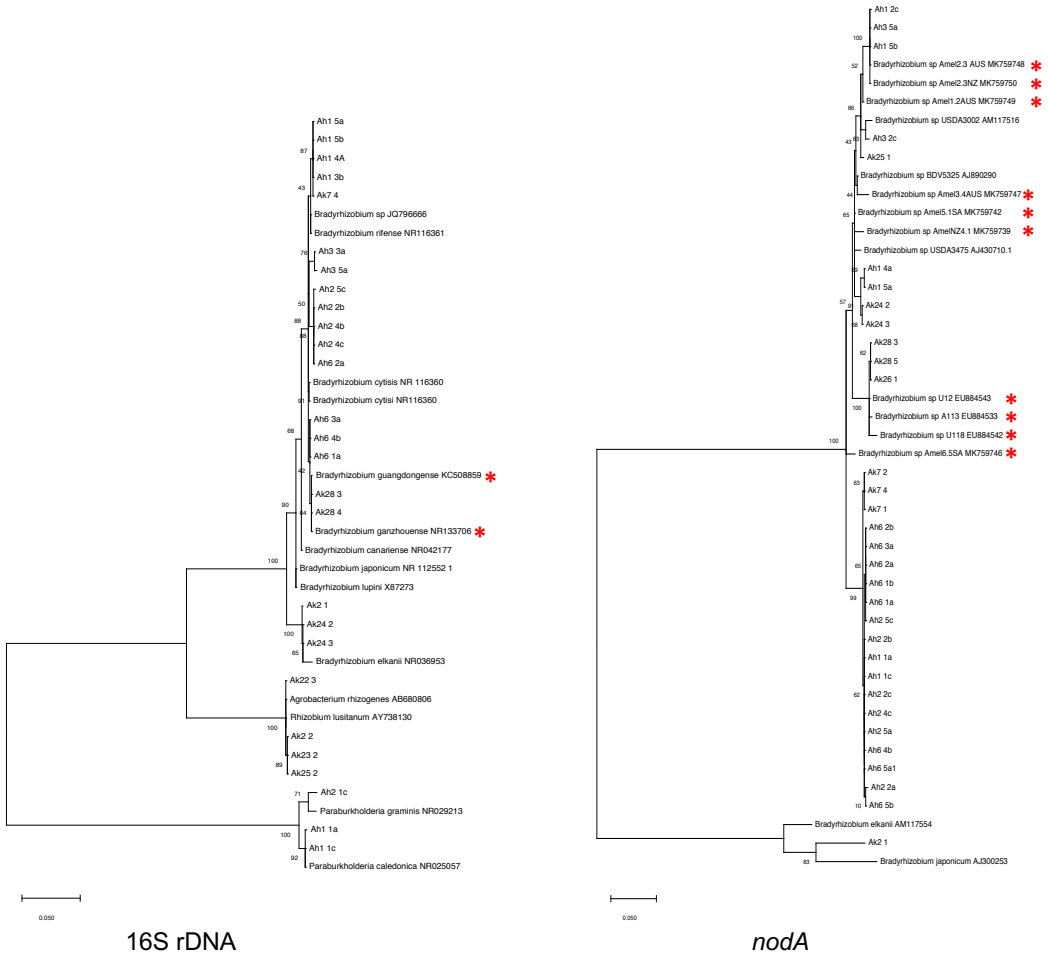

Fig 8: Phylogenetic relationships between root nodule bacteria isolated from *Acacia heterophylla* in La Réunion island (strain numbers starting with 'Ah') and *A. koa* in the Hawaiian Islands (strain numbers starting with 'Ak') and closely related strains based on BLAST results inferred from the 16S rRNA housekeeping (left) and nodulation *nodA* (right) gene sequences. Asteriks indicate strains that have been previously isolated from other *Acacia* species, including *A. melanoxylon*, the sister species of *A. heterophylla* and *A. koa*. These strains originated from both native (Australia) and non-native areas around the world.