# The functional potential of the rhizospheric microbiome of an invasive tree species, *Acacia dealbata*

Casper N. Kamutando[1], Surendra Vikram[1], Gilbert Kamgan-Nkuekam[1], Thulani P. Makhalanyane[1], Michelle Greve[2], Johannes J. Le Roux[3], David M. Richardson[3], Don A. Cowan[1], Angel Valverde[1,4*]

[1]Centre for Microbial Ecology and Genomics, Department of Genetics, University of Pretoria, Pretoria, South Africa.

[2]Department of Plant and Soil Sciences, University of Pretoria, Pretoria, South Africa.

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

[4]Department of Microbial, Biochemical and Food Biotechnology, University of the Free State, Bloemfontein, South Africa.

*Corresponding author at: P.O. Box 339, Bloemfontein 9300, Free State, South Africa.

E-mail address: valverdeportala@ufs.ac.za.

**ABSTRACT**

Plant-microbe interactions mediate both the invasiveness of introduced plant species and the impacts that they have in invaded ecosystems. Although the phylogenetic composition of the rhizospheric microbiome of *Acacia dealbata* (an invasive Australian tree species) has been investigated, little is known about the functional potential of the constituents of these altered microbial communities. We used shotgun DNA sequencing to better understand the link between bacterial community composition and functional capacity in the rhizospheric microbiomes associated with invasive *A. dealbata* populations in South Africa. Our analysis showed that several genes associated with plant growth-promoting (PGP) traits were significantly overrepresented in the rhizospheric metagenomes compared to neighbouring bulk soils collected away from *A. dealbata* stands. The majority of these genes are involved in the metabolism of nitrogen, carbohydrates and vitamins, and in various membrane transport systems. Overrepresented genes were linked to a limited number of bacterial taxa, mostly *Bradyrhizobium* species, the preferred N-fixing rhizobial symbiont of Australian acacias. Overall, these findings suggest that *A. dealbata* enriches rhizosphere soils with potentially beneficial microbial taxa, and that members of the genus *Bradyrhizobium* may play an integral role in mediating PGP processes that may influence the success of this invader when colonizing novel environments.

**Key words**

Biological invasions, *Bradyrhizobium*, Plant growth-promoting traits, Rhizosphere microbiome, Shotgun sequencing, Tree invasions

## INTRODUCTION

Rhizosphere microbial communities are highly diverse, both taxonomically and functionally [1]. Taxonomically, rhizosphere microbial communities consist primarily of taxa in the phyla *Proteobacteria*, *Actinobacteria* and *Bacteroidetes* [2]. Recent studies investigating the metabolic capabilities of the rhizosphere microbiomes of several plant species (e.g. soybean [3], wheat and cucumber [4], grapevine [5], wild and domesticated barley [6], and the widespread weed *Jacobaea vulgaris* [7]) have shown that genes overrepresented in the rhizosphere (compared to bulk soils away from focal plant species) include, among others, those encoding for proteins involved in the movement of bacteria toward root exudates and roots-adhesion, genes required for the metabolism of macro- and micronutrients, and stress tolerance-related genes. However, investigations to understand the metabolic capabilities of the rhizosphere microbiome of leguminous plants have received comparatively little attention (but see Mendes et al. [3] for a crop species). Understanding functional attributes of soil microbiomes and how these may change in response to invasive species, may be particularly valuable for understanding the factors promoting plant invasiveness and subsequent invader-mediated impacts on the ecosystems.

Many biological interactions determine plant fitness, including antogonistic (e.g., fungal diseases) and beneficial (e.g., pollination, seed dispersal, mycorrhizae, etc.) interactions, often with implications for the establishment and subsequent invasiveness of non-native plants [8, 9]. Introduced non-native plants can reassemble such interactions upon arrival in new environments through the formation of novel associations with resident species, or through co-introduction with their interacting partners. These two pathways have been postulated to have distinct consequences for the rate and accrual of invader impacts [10].

*Acacia dealbata* (silver wattle) is a leguminous tree native to South-Eastern Australia and Tasmania [11]. Like many invasive Australian acacias, invasive stands of silver wattle lead to

losses in aboveground plant biodiversity and to diverse alterations to the structure and functioning of rhizoshphere microorganisms in invaded ecosystems [12]. These invasions also affect soil microbial community structure and composition by altering nutrient pools through processes like excessive leaf litter inputs (carbon) and biological nitrogen fixation (nitrogen) [13]. While the microbiome of *A. dealbata* has been characterized [14] and found to be enriched in bacteria of the classes *Alpha-* and *Gammaproteobacteria*, and in fungi of the classes *Pezizomycetes* and *Agaricomycetes*, the functional potential of these microbial communities remains unexplored. However, certain *a priori* hypotheses can be formulated for invasive legumes. For example, the symbiosis between most legumes and bacteria collectively known as rhizobia, may suggest functional enrichment for rhizosphere communities linked to the supply and metabolism of nitrogen [15]. Indeed, various invasive acacias are known to preferentially associate with nitrogen-fixing *Bradyrhizobia*, even when these mutualists are not commonly found in association with native resident legumes (e.g., [16]). Surprisingly, many invasive acacias have been co-introduced with their *Bradyrhizobia* to numerous parts of the world (e.g., [17, 18]). Moreover, molecular evidence suggests that this is also the case for silver wattle invasions in South Africa (S Warrington, personal communication).

Different approaches can be used to elucidate the functional capabilities of rhizosphere microbiomes. For instance, DNA sequencing can be employed to determine the taxonomic makeup of rhizosphere microbial isolates [19], although it has been estimated that 99% of soil bacteria cannot be cultured in the laboratory [20]. Microbial functions can be derived from phylogenetic markers such as the 16S rRNA gene (PICRUSt; [21]); however bacterial strains with identical 16S rRNA gene sequences can exhibit highly divergent genomes and physiologies [22], making this technique ineffective to investigate functionality. Shotgun metagenomics can provide a better insight into community functional potential [23].

In this study, based on shotgun metagenomic data, we substantiated and expanded the preliminary analysis performed by Kamutando et al. [14] to better describe the genetic capacity of the bacterial communities in bulk and rhizosphere soil samples obtained from different *A. dealbata* populations. We also discuss several bacterial plant growth-promoting (PGP) traits. Finally, we suggest that the genus *Bradyrhizobium* may play an integral role in mediating PGP processes and may contribute to the success of this invader when colonizing novel environments.

## MATERIALS AND METHODS

### *Sample collection and DNA isolation*

Rhizosphere and non-rhizosphere (bulk) soil samples were collected as described previously [14]. Briefly, samples were collected from eight sites located in the grassland biome of South Africa. At each rhizosphere soil, firmly attached to the roots, and bulk soil, gathered 10-20 m away from any conspecific tree, were collected using sterile conditions. A total of 80 samples were collected (8 populations ×5 individuals ×2 habitat types (rhizosphere or bulk soil)), of which four rhizosphere and four bulk soil samples from four different *A. dealbata* populations were randomly selected for shotgun metagenomics. Metagenomic DNA was isolated using a PowerSoil DNA isolation kit (MO Bio Laboratories, Inc. Carlsbad, CA) as indicated by the manufacturer.

### *DNA library preparation and sequencing*

Metagenomic DNA was processed using a Nextera dual-indexed DNA sample preparation kit following the manufacturer's instructions (Illumina Inc., San Diego, CA, USA). Library fragments of 300-850 bp were selected from approx. 50 nanograms of metagenomic DNA, using the Experion Automated Electrophoresis System (Bio-Rad, Hercules, California 94547,

USA). Individual libraries were quantified using the Sequencing Library qPCR Quantification kit (Illumina Inc., San Diego, CA, USA), pooled in equimolar ratios, and re-purified. Pooled libraries (12nM each) were sequenced (2 x 150 bp paired end reads) on an Illumina HiSeq2000 instrument at the Molecular Research LP next generation sequencing service, USA (http://www.mrdnalab.com).

*De-novo assembly, coverage analysis and gene prediction*

Paired-end sequences were quality filtered (reads with a mean Phred quality score < 25 and/or containing ambiguous bases) using Prinseq-lite [24]. Quality-filtered reads were *de novo* assembled into longer contiguous segments (i.e., contigs) using metaSPAdes v3.7.1 [25] and default settings. Contigs with length < 500bp were removed from each sample. To determine the percentage of assembled reads, quality-filtered sequences were mapped back to the contigs using Burrows-Wheeler Aligner (BWA) [26] and the resultant alignments were '*sorted*' and '*indexed*' using SAMtools version 1.3.1 [27]. Contig coverage was predicted using bamM (http://ecogenomics.github.io/BamM/), and open-reading frames (ORFs) were predicted using Prodigal v2.6.3 [28].

*Taxonomic and functional annotation of protein-coding ORFs*

Predicted ORFs were aligned against the NCBI-nr database using DIAMOND (BLASTP v0.7.1, E-value cut-off at $1e^{-5}$) [29]. Functional annotations based on Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathways, and taxonomic annotations of the protein-coding ORFs were performed in MEGAN CE v6.5.9 [30]. From these, we obtained two matrices that contained the taxonomic and functional ORF hit counts, which were used for statistical analysis. Normalization of ORF hits was performed in MEGAN CE during the comparison of samples.

*Genome reconstruction, refinement and quality checking*

Contigs with lengths of <1,500 bp were filtered from the assemblies to improve genome binning, which is highly reliant on the proportion of long contigs [31]. Binning of the long contigs was performed using CONCOCT v0.4.1 [31]. The recovered composite genomes were imported into the Anvi'o v2.2.2 and refined using the script "*anvio-refine"* as described previously [32]. The quality of the refined genome bins (i.e., completeness, contamination and strain heterogeneity) was assessed using CheckM v1.0.7 [33], an automated method used to estimate contamination and completeness of genome bins using marker genes that are specific to the respective lineage within a reference genome tree [33]. BamM (http://ecogenomics.github.io/BamM/) was used to assess coverage in composite genomes with ≥70% completeness.

*Composite genome annotation*

Composite genomes showing ≥ 70% completeness were submitted to the Rapid Annotation Subsystem Technology (RAST) server v2.0 [34, 35]. The predicted gene-encoded proteins (amino acid sequences) for each genome bin obtained after RAST annotations were submitted to the KEGG Annotation Server (KAAS) v2.1 [36].

*Statistical analyses*

The matrices containing taxonomic and functional ORF hit counts were Hellinger-transformed and the distances between samples calculated using Bray-Curtis. Dissimilarities between samples were depicted using PCoA plots. A permutational analysis of variance (PERMANOVA; [37]) was used to test for differences in composition and function between habitats (rhizosphere *vs.* bulk soil). KEGG orthologue groups shared between the rhizosphere

and bulk soil metagenomes were identified using Venn diagrams. Statistical Analysis of Metagenomic Profiles (STAMP) package was used to discriminate between important taxa and functions between the rhizosphere and the bulk soil samples [38]. Statistical significance was assessed using a two-sided White's non-parametric t-test with Benjamini-Hochberg false discovery rate (p-value <0.05) [39].

## RESULTS

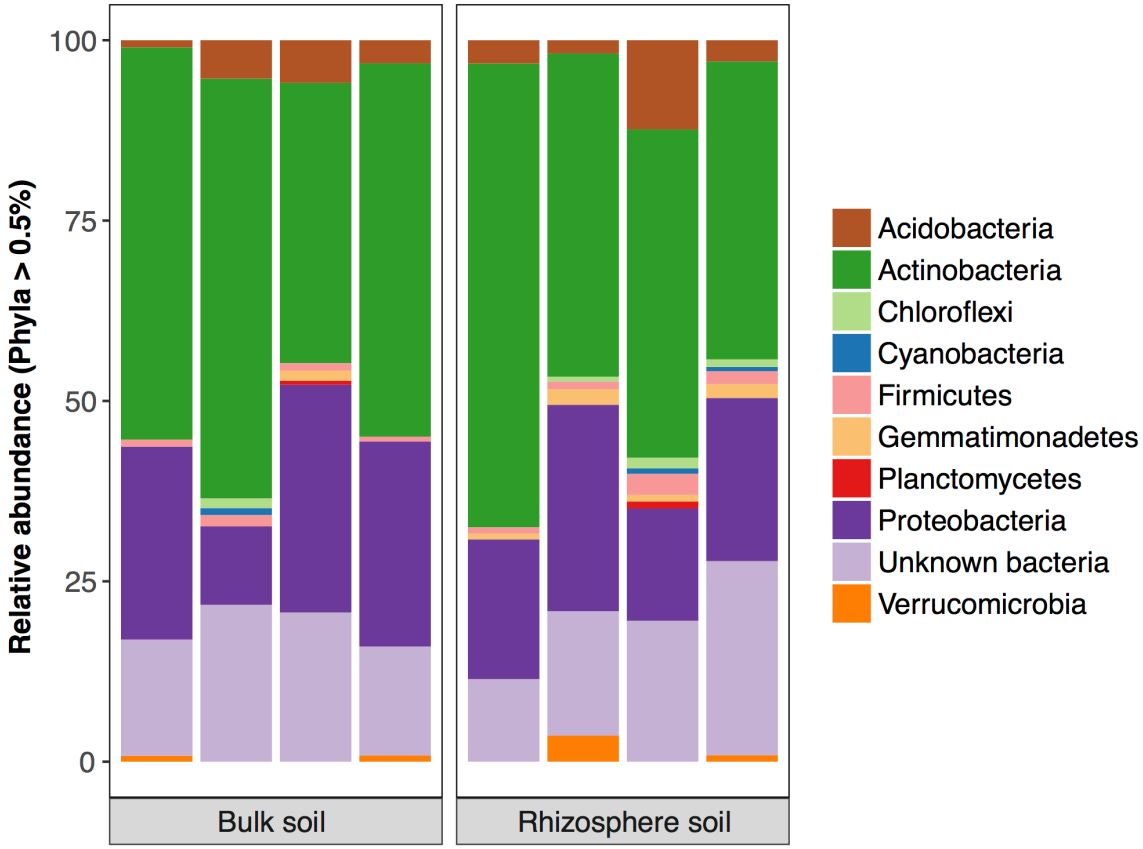
### *Soil metagenomes associated with invasive* A. dealbata

Illumina Hiseq2000 sequencing produced a total of 73,292,578 reads (10.87 Gbp, average length of 148 bp) for the four bulk soil metagenomes and 58,020,031 reads (8.55 Gbp, average length of 147 bp) for the four rhizosphere soil metagenomes. After quality filtering, approximately 70 million reads (10.29 Gbp) for bulk soil samples, and 54.8 million reads (7.98 Gbp) for rhizosphere soil samples, were obtained. Assembled contigs from bulk soil metagenomes showed an average sequence coverage of 55.06%, while 63.77% coverage was observed in rhizospheric metagenomes (Supplementary Table S1).

### *Taxonomic profiling*

Taxonomic annotation of protein-coding ORFs resulted in identification of 1,390 taxa (species level), with a total of 256,005 ORF hits. Bacteria accounted for about 99% of the ORFs, while 0.24% were assigned to archaea, 0.15% to eukaryotes and 0.01% to viruses. 0.05% of the ORFs remained unassigned (Supplementary Table S2). A PCoA plot of Bray-Curtis dissimilarities showed no clear differences in microbial structure and composition between the rhizosphere and the bulk soil metagenomes (Supplementary Fig. S1; PERMANOVA $F_{1,6}$=1.07; P = 0.39). At the phylum level, *Actinobacteria*, *Proteobacteria*, and *Acidobacteria* accounted for the

**Fig. 1** Mean relative abundance of taxa (phylum level) within each sample. The relative abundance of each taxon was calculated as the percentage of protein-coding ORF hits per sample.



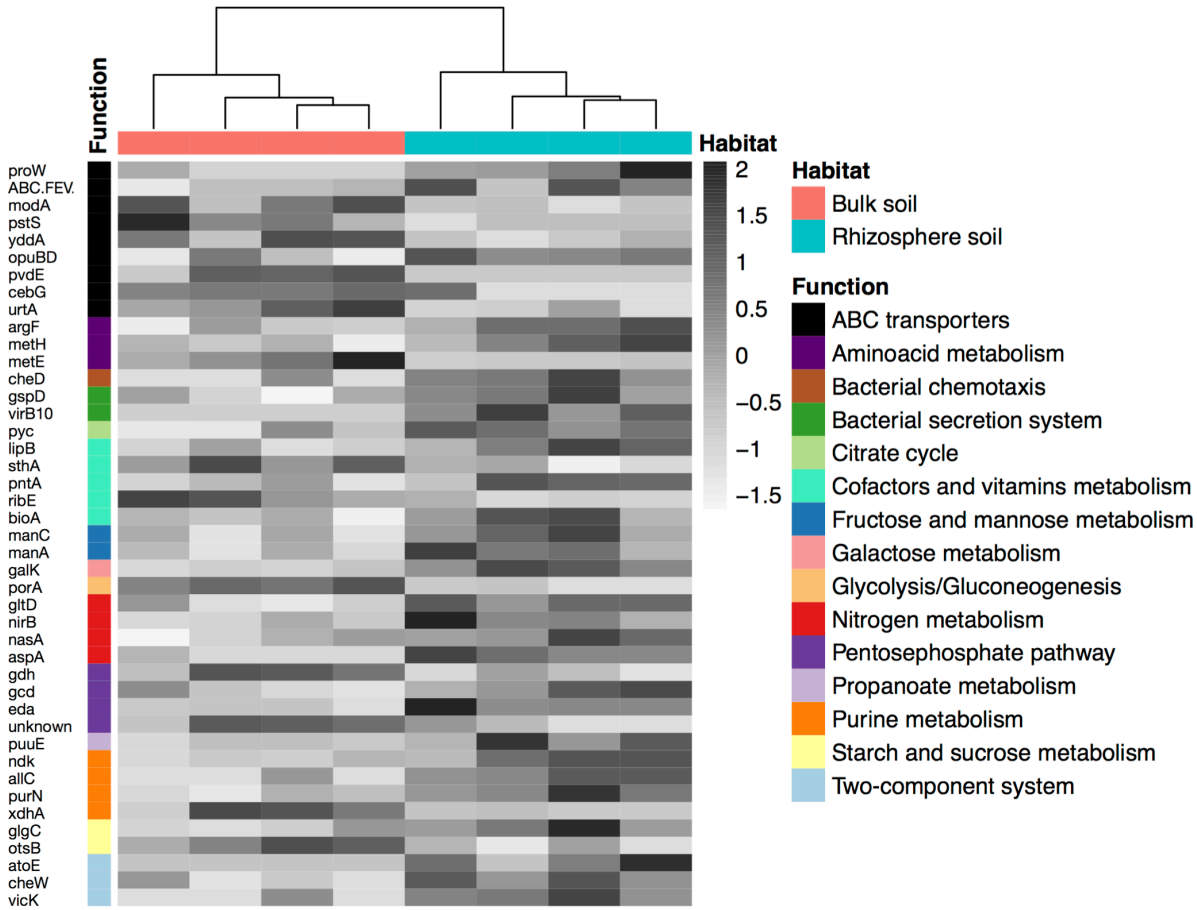
majority of the ORFs (Fig. 1; Supplementary Table S2). The relative abundance of these phyla did not differ significantly between the rhizosphere and bulk soil samples (White's non-parametric t-test, $P > 0.05$). However, 25 and 27 taxa (species level) were overrepresented in bulk soil and rhizospheric metagenomes, respectively (Supplementary Table S3), including, for example, *Bradyrhizobium* and *Sphingomonas* species, which are known to contain strains with PGP activities [40, 41].

*Functional profiles*

A total of 4,125 KEGG Orthologues (KOs) were obtained, of which 1,969 (48%) were assigned to functional categories, while 2,156 (52%) did not have known functional roles. 3,182 (77.14%) KOs were shared between the rhizosphere and bulk soil metagenomes, where 623 (15.1%) were unique to the rhizosphere samples and 320 (7.76%) were observed only in bulk soil samples (Supplementary Fig. S3a). A PCoA plot based on Bray-Curtis dissimilarities derived from KO counts (104,372 ORF hits) showed no significant differences between rhizosphere and bulk soil communities (Supplementary Fig. S1, PERMANOVA $F_{1,6}=1.09$; $P = 0.32$). Nevertheless, STAMP analysis identified 173 KOs which were significantly enriched in one or other of the two habitats: 94 KOs were over-represented in the rhizosphere, while 79 KOs were over-represented in bulk soil metagenomes (Supplementary Table S4). Of the 173 KOs, 152 (87.86 %) were present in both habitats, 14 (8.09%) KOs were unique to the rhizosphere and 7 (4.05%) KOs were unique to the bulk soils (Supplementary Fig. S3b).

The differentially abundant genes (KOs) of the rhizosphere (94 in total) were associated with amino acid metabolism (e.g., *OTC* and *metH*), bacterial chemotaxis (e.g., *cheD*), bacterial secretion systems (e.g., *virB10*), citrate acid cycle (e.g., *PC*), metabolism of cofactors and vitamins (e.g., *lipB* and *bioA*), metabolism of fructose, galactose and starch (e.g., *galK* and *manA*), nitrogen metabolism (e.g., *nirB* and *nasA*), purine metabolism (e.g., *purN*) and the two-

**Fig. 2** Heat map of differentially abundant plant growth-beneficial functions, as revealed by STAMP statistical analysis. Genes associated with a similar function are represented by one colour; for example, black denotes genes associated with ABC transporters. Each row was scaled so that the mean of each gene across samples was calculated and coloured by the corresponding z-score of each cell. Clustering of the samples was done using the UPGMA method with correlation distances. Individual genes associated to a specific function are shown on the left side of the figure.



component system (e.g., *vicK*) in the rhizosphere samples (Fig. 2). The analysis of the data revealed that the genera *Bradyrhizobium* (105 genes), *Amycolatopsis* (81 genes), *Mycobacterium* (72 genes) and *Rubrobacter* (61 genes) were the bacterial taxa carrying most of the putative plant growth-beneficial functions in the rhizosphere of *A. dealbata*. Other bacteria with known plant growth-promoting potential such as *Rhizobium* and *Streptomyces* were also shown to be present in rhizosphere soils (Supplementary Table S5).

*Composite genomes reconstructed from soil metagenomes of invasive A. dealbata*

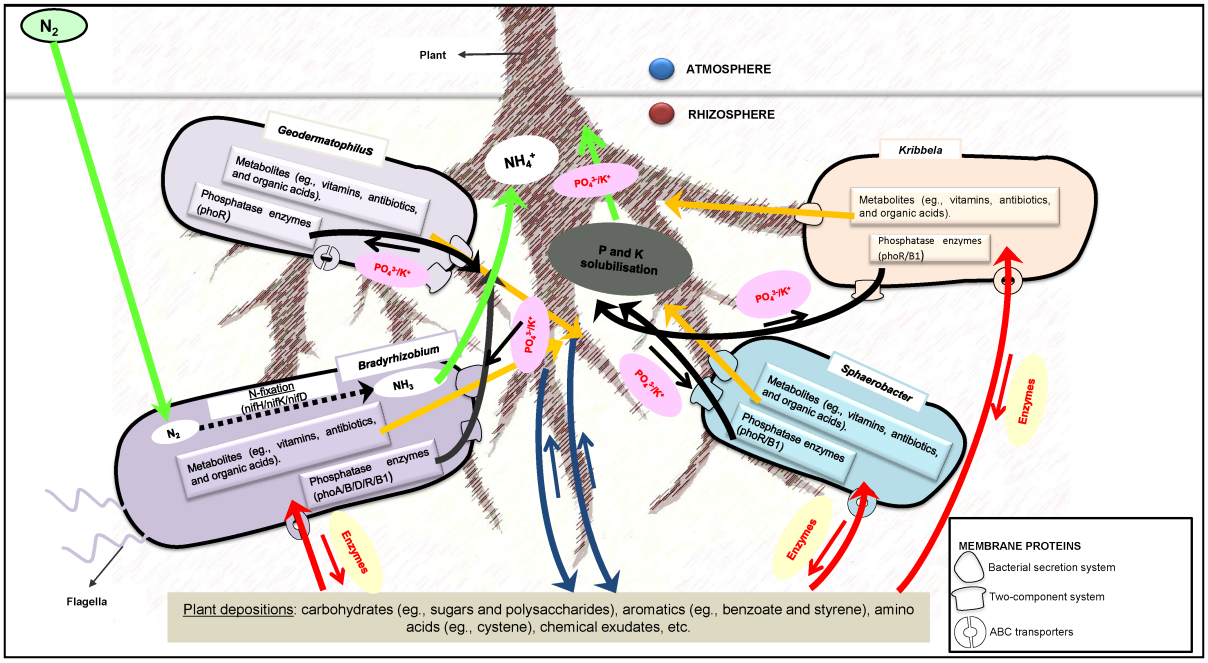
The binning of metagenomic contigs (length ≥ 1,500bp) yielded 46 bins (Supplementary Table S6). Bin refinement, using Anvio, resulted in five composite genomes that showed ≥ 70% completeness, including Bin12.2 (80.17% complete, 4.55% contamination), Bin15.1 (89.56% complete, 0.74% contamination), Bin17.1 (83.64% complete, 1.64% contamination), Bin20.1 (85.11% complete, 3.02 contamination) and Bin21 (73.49% complete, 0% contamination) (Supplementary Table S7). Other genome bins showed less than 50% completeness and high contamination (Supplementary Table S7). Genome annotation identified those five composite genomes as *Koribacter* (Bin12.2, 4.9 Mbp), *Bradyrhizobium* (Bin 15.1, 7.5 Mbp), *Geodermatophilus* (Bin17.1, 3.91 Mbp), *Kribella* (Bin 20.1, 6.6 Mbp) and *Sphaerobacter* (Bin21, 2.4 Mbp). *Bradyrhizobium*, *Geodermatophilus*, *Kribbela* and *Sphaerobacter* were present in high abundance in the rhizosphere, as compared to *Koribacter* which was more abundant in the bulk soil. The highest number of coding sequences and subsystems were predicted in the composite genome of *Bradyrhizobium*, in relation to the other composite genomes (Table 1).

The composite genomes overrepresented in the rhizosphere (i.e., *Bradyrhizobium*, *Geodermatophilus*, *Kribella* and *Sphaerobacter*) showed potential for the metabolism of compounds that can directly or indirectly promote plant growth (Fig. 3; Supplementary Tables

**Table 1** Characteristics of the five near-complete composite genomes obtained from *Acacia dealbata*-invaded soil metagenomes.

| **Parameters** | **Bin12.2** | **Bin15.1** | **B17.1** | **Bin20.1** | **Bin21** |
|---|---|---|---|---|---|
| Closest taxonomic neighbour | *Koribacter* | *Bradyrhizobium* | *Geodermatophilus* | *Kribbela* | *Sphaerobacter* |
| Genome size (Mbp) | 4.9 | 7.5 | 3.91 | 6.6 | 2.4 |
| Sequence size (bp) | 4,912,472 | 7,476,157 | 3,911,194 | 6,607,702 | 2,363,288 |
| Shortest contig size | 1501 | 1530 | 2523 | 2203 | 1662 |
| Median sequence size | 4644 | 7822 | 14308 | 19673 | 4680 |
| Mean sequence size | 5918.6 | 9785.5 | 17778.2 | 26751.8 | 5908.2 |
| Longest contig size | 30361 | 49996 | 84287 | 122199 | 26027 |
| No. of contigs with protein-encoding genes (PEGs) | 830 | 764 | 220 | 247 | 400 |
| No. of subsystems | 299 | 476 | 345 | 390 | 229 |
| No. of coding sequences | 4417 | 7052 | 3796 | 6324 | 2442 |
| No. of RNAs | 57 | 47 | 49 | 33 | 32 |
| GC content (%) | 53.7 | 63.9 | 68.5 | 69.5 | 66.3 |
| N50 value | 5926 | 8817 | 21074 | 40710 | 4887 |
| L50 value | 274 | 266 | 62 | 51 | 154 |
| Genome completeness (%) | 80.17 | 89.56 | 83.64 | 85.11 | 73.49 |
| Genome contamination (%) | 4.55 | 0.74 | 1.64 | 3.02 | 0 |
| **Sequence coverage values (no. of reads mapped back to the composite genomes)** | | | | | |
| 34R (Rhizosphere) | 4 | 346699 | 48 | 906 | 0 |
| 43R (Rhizosphere) | 10 | 3342 | 114 | 117 | 94458 |
| 21R (Rhizosphere) | 15 | 21915 | 334833 | 1043984 | 8 |
| 51R (Rhizosphere) | 17 | 12035 | 57 | 1569 | 107 |
| **Rhizosphere soil total** | **46** | **383991** | **335052** | **1046576** | **94573** |
| 14C (Bulk soil) | 51 | 2989 | 82 | 638 | 6 |
| 23C (Bulk soil) | 214753 | 1013 | 166 | 582 | 21 |
| 84C4 (Bulk soil) | 43 | 3006 | 146 | 1081 | 49 |
| 43C (Bulk soil) | 27 | 4256 | 5517 | 1334 | 82 |
| **Bulk soil total** | **214874** | **11264** | **5911** | **3635** | **158** |

**Fig. 3** Summary of putative PGP roles of the genomes reconstructed from *A. dealbata* rhizosphere metagenomes. Results are based on RAST (Overbeek et al., 2014) and KAAS (Moriya et al., 2007) annotations of genome bins (≥ 70% completeness). Black arrows illustrate transference of phosphorus solubilizing enzymes from bacteria to the soil and the uptake of phosphate and potassium by the bacteria. Orange arrows denote conveyance of bacterial metabolites into the soil. Dark blue arrows depict plant depositions into the rhizosphere and re-uptake of organic nutrients by plant roots. Red arrows illustrate secretion of various enzymes involved in decomposition of plant deposits and uptake of organic nutrients by the bacteria. Light green arrows elucidate assimilation of nitrogen, phosphorus and potassium into the plant host.



S8-S11). For example, we identified genes involved in metabolism of B-vitamins such as thiamine (e.g., *thiL* and *thiG* genes), riboflavin (e.g., *ribA* gene), pyridoxine (e.g., *pdxA* and *pdxJ* genes), biotin (e.g., *bioA* and *bioB* genes), lipoic acid (e.g., *lipA* and *lipB* genes) and folate (e.g., *folA* and *folC* genes). Genes encoding for metabolism of secondary metabolites such as antibiotics, for example, monobactam (i.e., *dapB*), streptomycin (i.e., *rfbC*), carbapenem (i.e., *proA*) and penicillin (i.e., *penP*) were also identified in the genomes.

Most of the genes associated with the cycling of nitrogen and sulphur were predicted in the *Bradyrhizobium* genome, compared to the other composite genomes. For instance, genes involved in the conversion of dinitrogen into ammonia (e.g., *nifA*, *nifD* and *nifH*), dissimilatory nitrogen reduction (e.g., *napB* and *nirD*), assimilatory nitrogen reduction (e.g., *nirA*) and denitrification (e.g., *napA* and *nirK*). The capacity for assimilatory sulfate reduction (e.g., *cysN* and *cysH* genes) and the sox system (e.g., *soxA* and *soxB* genes) were also found in the *Bradyrhizobium* genome. All four composite genomes revealed the potential for carbon cycling processes. For example, genes encoding the enzymes responsible for the metabolism of starch (e.g., *AMY*), mannose (*manA*), fructose (*fucA*) and galactose (*galK*) were predicted.

## DISCUSSION

We investigated the microbial functional diversity within the rhizosphere of a widespread invasive Australian tree species, *A. dealbata,* in South Africa. Numerous studies have illustrated how invasive plants can change soil microbial community diversity and structure (e.g. [3, 7]). Such changes often result in positive feedbacks between invasive plants and altered soil microbial communities (e.g., [42]), but the precise functional changes underlying these phenomena remains unexplored. Here we report that invasive plants do indeed change soil microbial functionality, assessed using shotgun metagenomics. Specifically, we found different microbial functional traits, known to promote plant growth, to be overrepresented in the

rhizospheres of *A. dealbata* compared to uninvaded soils. Moreover, *Bradyrhizobium* species, the preferred rhizobial mutualists of Australian acacias [17, 18], harboured most of these traits. We were also able to assemble a composite genome of *Bradyrhizobium* and several other genomes.

Previous studies have shown that the "rhizosphere effect"; that is, where the rhizosphere is enriched for microbial communities that differ (both taxonomically and functionally) from those of the surrounding soil, seems to be primarily host plant-driven. For example, the rhizosphere effect in barley [6] and soybean [3] appears to be stronger than those of *Arabidopsis thaliana* and *A. thaliana* relatives [43, 44], at least on the basis of bacterial community composition. The stronger rhizosphere effect may be related to a stronger investment of some plant species in the promotion of a beneficial microbiome, potentially through differential patterns in root exudation. Our results are in agreement with these studies, where only a limited number of bacterial species and functions were overrepresented in the rhizosphere of *A. dealbata* compared to the bulk soil metagenomes.

It is particularly notable that genes associated with PGP traits, such as bacterial chemotaxis, bacterial secretion system, two-component system, and metabolism of cofactors, vitamins, carbohydrates and nitrogen, were overrepresented in the rhizospheric metagenomes (Figure 2). Many of these genes have been previously associated to key traits in bacteria colonizing rhizosphere soils of model plants such as *Arabidopsis thaliana* [45-47], agricultural crops [3, 4] and forage plants [48].

Studies on crops such as wheat, cucumber and soybean revealed the potential of particular plants in selecting for a relatively low number of rhizospheric microorganisms, which were postulated to harbour important growth-promoting activities [3, 4]. Our results indicate that these findings may also hold up for invasive plants. In mapping PGP traits to rhizosphere-associated taxa, we found most of the genes overrepresented in the rhizosphere to be associated

with *Bradyrhizobium*, *Amycolatoptis* and *Mycobacterium* species (Supplementary Table S5). *Bradyrhizobium* is a nitrogen fixer, commonly found in association with acacias [49-51], which is likely enhancing the competitive ability of these trees in invaded habitats. Un-invaded natural ecosystems often have depleted nitrogen levels, hence bacteria that fix nitrogen may improve the competitive ability of these trees when growing in areas with poor nitrogen nutrition [52], thereby promoting their invasiveness [50, 53]. It is therefore expected for such beneficial microbes to be enriched by their invasive host plants, as has been shown for *A. dealbata* [14]. Recent work illustrated that several invasive acacias homogenize rhizobial bacterial communities, even across wide geographic ranges [42]. Acacias and their Bradyrhizobial associates have been co-introduced in numerous regions around the globe, including South Africa (S Warrington, personal communication). Therefore, Bradyrhizobial enrichment by invasive *A. dealbata* in South Africa may represent multiple and co-invading organisms, possibly with strong positive feedbacks [42].

We also found several members of the genus *Amycolatopsis*, known to prolifically degrade lignin and low molecular weight aromatics compounds [54, 55], to be dominant in the rhizosphere of invasive *A. dealbata* trees. This may reflect the large levels of detritus accumulated by this plant [56]. Also, *Amycolatopsis* species have previously been reported to contribute to plant health by antibiotic production [57]. The genus *Mycobacterium*, which contains important human pathogens [58], was also prevalent in rhizosphere soils. This is not surprising, as most of the species of the genus are ubiquitous saprophytes in many environments [59]. Some bacterial species of the genera *Burkholderia*, *Enterobacter*, *Herbaspirillum*, *Ochrobactrum*, *Pseudomonas*, *Ralstonia*, *Serratia*, *Staphylococcus*, *Stenotrophomonas* and *Xylella* can enter bivalent interactions with plant and human hosts [60, 61]. This seems to indicate that the molecular basis for bacterial pathogenicity is both conserved and independent of host phylogeny. *Mycobacterium* belongs to the order

*Actinomycetales*, which includes many species that produce different types of antibiotics [62]. Antibiotic production may help these bacteria to colonize and proliferate on plant surfaces by outcompeting other microbes, including plant soil-borne pathogens [63], thereby protecting the host plant from soil borne diseases [64]. Members of the genera *Rhizobium* and *Streptomyces*, with well-known beneficial effects on plant growth [65, 66], were also linked to genes associated with PGP traits in the rhizosphere metagenomes. Overall, our results suggest that in addition to *Bradyrhizobium*, which may be the primary driver of PGP processes in invaded soils, *A. dealbata* associate with several other taxa of beneficial bacteria.

Because most of the genes linked to key PGP traits could be assigned to the genus *Bradyrhizobium*, we reconstructed the *Bradyrhizobium* genome from the rhizosphere metagenomes. Using this strategy, we confirmed that most of the genes associated with PGP traits, such as metabolism of nitrogen, carbohydrates, sulphur, antibiotics and vitamins, as well as various two-component systems, bacterial secretion systems and bacterial chemotaxis, were indeed found in the *Bradyrhizobium* genome. Some of these traits may have relevance in invasion ecology, for example, the ability to modify the nitrogen cycle. The ability to control nitrification rates seems to influence species coexistence and dominance [67]. This was shown for the invasive grass *Andropogon gayanus* in Australian woodlands, which prefers ammonium over nitrate as a nitrogen source, thereby inhibiting nitrification but stimulating ammonification [68]. Other studies have reported that invasive plants, such as *Fallopia* spp., could modify denitrification and nitrification processes [69]. Modification of these processes in invaded soils could lead to reduced nitrogen loss from the ecosystem through nitrate leaching or volatilization of nitrous oxides and dinitrogen, promoting the productivity of the invader.

The potential for production of soluble vitamins and antibiotics detected in the rhizosphere metagenomes and genomes of *Bradyrhizobium*, *Sphaerobacter*, *Geodermatophilus and Kribbela* may also be important in the context of plant invasions [62, 70]. Soluble vitamins

show the potential to protect plants against phytopathogens [71, 72], to antagonize the effects of antioxidants (e.g., singlet oxygen species) [73, 74] and to act as cofactors in metabolic reactions [75, 76]. Recently, pyridoxine (vitamin B6) was shown to influence nitrogen balance in plants [77], an effect which can directly impact on the rate of plant development and overall plant fitness. The secretion of antibiotics may promote the competitive ability of bacteria by suppressing competitive microbes [78].

Although our results suggest that *A. dealbata* plants enrich the rhizosphere with beneficial bacterial taxa for *A. dealbata*, it can still be argued that the presence of genes in a genome does not necessarily mean that these genes are being expressed. Nevertheless, Ofek-Lalzar et al. [4] combined both metagenomics (functional potential) and metatranscriptomics (functional expression) to show that a large proportion of genes involved in PGP activities were indeed expressed. Thus, we can postulate that at least some of the important genes detected in the rhizosphere metagenomes may be expressed in soil environments invaded by *A. dealbata*. However, a more in-depth investigation is needed to validate this hypothesis.

In conclusion, using shotgun metagenomics, this study has demonstrated that several bacterial traits such as those related to motility towards the plant host, antibiotics production, nutrient cycling and membrane transport systems, among others, are overrepresented in the rhizosphere microbiome of *A. dealbata*. We also show that, although this leguminous tree seems to associate with different bacterial taxa, *Bradyrhizobium* strains, possibly co-introduced with *A. dealbata*, harbour most of these genes and therefore probably play an integral role in promoting *A. dealbata* growth and productivity. We argue that using samples from a variety of soils, as opposed to samples from similar soils, improves progress towards the identification of taxa with a greater likelihood of influencing the host phenotype, and therefore with the potential to promote the invasiveness of *A. dealbata*.

**References**


1. Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH (2013) Going back to the roots: the microbial ecology of the rhizosphere. Nat Rev Microbiol 11: 789-799.
2. Hacquard S, Garrido-Oter R, González A, Spaepen S, Ackermann G, Lebeis S, McHardy AC, Dangl JL, Knight R, Ley R (2015) Microbiota and host nutrition across plant and animal kingdoms. Cell Host Microbe 17: 603-616.
3. Mendes LW, Kuramae EE, Navarrete AA, Van Veen JA, Tsai SM (2014) Taxonomical and functional microbial community selection in soybean rhizosphere. ISME J 8: 1577-1587.
4. Ofek-Lalzar M, Sela N, Goldman-Voronov M, Green SJ, Hadar Y, Minz D (2014) Niche and host-associated functional signatures of the root surface microbiome. Nat Commun 5(4950): 1-9.
5. Zarraonaindia I, Owens SM, Weisenhorn P, West K, Hampton-Marcell J, Lax S, Bokulich NA, Mills DA, Martin G, Taghavi S (2015) The soil microbiome influences grapevine-associated microbiota. mBio 6: 1-10.
6. Bulgarelli D, Garrido-Oter R, Münch P.C, Weiman A, Dröge J, Pan Y, McHardy AC, Schulze-Lefert P (2015) Structure and function of the bacterial root microbiota in wild and domesticated barley. Cell Host Microbe 17: 392-403.
7. Yan Y, Kuramae EE, de Hollander M, Klinkhamer PG, van Veen JA (2016) Functional traits dominate the diversity-related selection of bacterial communities in the rhizosphere. ISME J 11: 56-66.
8. Van der Putten WH, Klironomos JN, Wardle DA (2007) Microbial ecology of biological invasions. ISME J 1: 28-37.

9. Inderjit, Cahill JF (2015) Linkages of plant-soil feedbacks and underlying invasion mechanisms. AoB Plants 7: pvl022.

10. Le Roux JJ, Hui C, Keet JH, Ellis AG (2017) Co-introduction vs ecological fitting as pathways to the establishment of effective mutualisms during biological invasions. New Phytologist. New Phytol 215: 1354-1360.

11. Poynton R (2009) Tree planting in Southern Africa. Volume 3: Other genera. Department of Agriculture. Forestry and Fisheries, Pretoria.

12. Lazzaro L, Giuliani C, Fabiani A, Agnelli AE, Pastorelli R, Lagomarsino A, Benesperi R, Calamassi R, Foggi B (2014) Soil and plant changing after invasion: The case of *Acacia dealbata* in a Mediterranean ecosystem. Sci Total Environ 497: 491-498.

13. May B, Attiwill P (2003) Nitrogen-fixation by *Acacia dealbata* and changes in soil properties 5 years after mechanical disturbance or slash-burning following timber harvest. For Ecol Manag 181: 339-355.

14. Kamutando CN, Vikram S, Kamgan-Nkuekam G, Makhalanyane TP, Greve M, Roux JJL, Richardson DM, Cowan D, Valverde A (2017) Soil nutritional status and biogeography influence rhizosphere microbial communities associated with the invasive tree *Acacia dealbata*. Sci Rep 7: 6472.

15. Vacheron J, Desbrosses G, Bouffaud M-L., Touraine B, Moënne-Loccoz Y, Muller D, Legendre L, Wisniewski-Dyé F, Prigent-Combaret C (2013) Plant growth-promoting rhizobacteria and root system functioning. Front Plant Sci 4: 1-19.

16. Le Roux JJ, Mavengere NR, Ellis AG (2016) The structure of legume–rhizobium interaction networks and their response to tree invasions. AoB plants 8: plw038.

17. Crisóstomo JA, Rodríguez-Echeverría S, Freitas H (2013) Co-introduction of exotic rhizobia to the rhizosphere of the invasive legume *Acacia saligna*, an intercontinental study. Appl Soil Ecol 64: 118-126.

18. Rodríguez-Echeverría S (2010) Rhizobial hitchhikers from down under: invasional meltdown in a plant–bacteria mutualism? J Biogeogr 37: 1611-1622.

19. Wurch L, Giannone RJ, Belisle BS, Swift C, Utturkar S, Hettich RL, Reysenbach A-L, Podar M (2016) Genomics-informed isolation and characterization of a symbiotic Nanoarchaeota system from a terrestrial geothermal environment. Nat Commun 7: 1-10.

20. Hugenholtz P, Pace NR (1996) Identifying microbial diversity in the natural environment: a molecular phylogenetic approach. Trends Biotechnol 14: 190-197.

21. Langille MG, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkepile DE, Thurber RLV, Knight R (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol 31: 814-821.

22. Jaspers E, Overmann J (2004) Ecological significance of microdiversity: identical 16s rRNA gene sequences can be found in bacteria with highly divergent genomes and ecophysiologies. Appl Environ Microbiol 70: 4831-4839.

23. Louca S, Parfrey LW, Doebeli M (2016) Decoupling function and taxonomy in the global ocean microbiome. Science 353: 1272-1277.

24. Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. Bioinformatics 27: 863-864.

25. Nurk S, Meleshko D, Korobeynikov A, Pevzner P (2016) metaSPAdes: a new versatile de novo metagenomics assembler. arXiv 27: 824-834.

26. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler Transform. Bioinformatics 25: 1754-1760.

27. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25: 2078-2079.

28. Hyatt D, Chen G-L., LoCascio PF, Land ML, Larimer FW, Hauser LJ (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11: 119.

29. Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. Nat Methods 12: 59-60.

30. Huson DH, Beier S, Flade I, Górska A, El-Hadidi M, Mitra S, Ruscheweyh H-J, Tappu R (2016) MEGAN community edition-interactive exploration and analysis of large-scale microbiome sequencing data. PLoS Comput Biol 12: e1004957.

31. Alneberg J., Bjarnason B.S., De Bruijn I., Schirmer M., Quick J., Ijaz U.Z., Lahti L., Loman N.J., Andersson A.F., Quince C. (2014) Binning metagenomic contigs by coverage and composition. Nat Methods 11: 1144-1146.

32. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO (2015) Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3: e1319.

33. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25: 1043-1055.

34. Aziz R.K., Bartels D., Best A.A., DeJongh M., Disz T., Edwards R.A., Formsma K., Gerdes S., Glass E.M., Kubal M. (2008) The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9: 1-15.

35. Overbeek R., Olson R., Pusch G.D., Olsen G.J., Davis J.J., Disz T., Edwards R.A., Gerdes S., Parrello B., Shukla M. (2014) The SEED and the Rapid Annotation of

microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42: D206-D214.

36. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res 35: W182-W185.

37. Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecol 26: 32-46.

38. Parks DH, Beiko RG (2010) Identifying biologically relevant differences between metagenomic communities. Bioinformatics 26: 715-721.

39. White JR Nagarajan N, Pop M (2009) Statistical methods for detecting differentially abundant features in clinical metagenomic samples. PLoS Comput Biol 5: e1000352.

40. White DC, Sutton SD, Ringelberg DB (1996) The genus *Sphingomonas*: physiology and ecology. Curr Opin Biotechnol 7: 301-306.

41. Figueiredo MdVB, Seldin L, de Araujo FF, Mariano RdLR (2010) Plant growth promoting rhizobacteria: fundamentals and applications, Plant growth and health promoting bacteria. Springer 21-43.

42. Le Roux JJ, Ellis AG, Zyl L-M, Hosking ND, Keet J-H, Yannelli FA (2018) Importance of soil legacy effects and successful mutualistic interactions during Australian acacia invasions in nutrient poor environments. J Ecol. in press doi:10.1111/1365-2745.12965

43. Bulgarelli D, Rott M, Schlaeppi K, van Themaat EVL, Ahmadinejad N, Assenza F, Rauf P, Huettel B, Reinhardt R, Schmelzer E (2012) Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature 488: 91-95.

44. Schlaeppi K, Dombrowski N, Oter RG, van Themaat EVL, Schulze-Lefert P (2014) Quantitative divergence of the bacterial root microbiota in *Arabidopsis thaliana* relatives. Proc Natl Acad Sci USA 111: 585-592.

45. Barret M, Morrissey JP, O'Gara F (2011) Functional genomics analysis of plant growth-promoting rhizobacterial traits involved in rhizosphere competence. Biol. Fertil Soils 47: 729-743.

46. Compant S, Clément C, Sessitsch A (2010) Plant growth-promoting bacteria in the rhizo- and endosphere of plants: their role, colonization, mechanisms involved and prospects for utilization. Soil Biol Biochem 42: 669-678.

47. Records AR (2011) The type VI secretion system: a multipurpose delivery system with a phage-like machinery. Mol Plant Microbe Interact 24: 751-757.

48. Zhou Y, Zhu H, Fu S, Yao Q (2017) Metagenomic evidence of stronger effect of stylo (legume) than bahiagrass (grass) on taxonomic and functional profiles of the soil microbial community. Sci Rep 7(10195): 1-11.

49. Birnbaum C, Bissett A, Thrall PH, Leishman MR (2016) Nitrogen-fixing bacterial communities in invasive legume nodules and associated soils are similar across introduced and native range populations in Australia. J Biogeogr 43: 1631-1644.

50. Hoque MS, Broadhurst LM, Thrall PH (2011) Genetic characterization of root-nodule bacteria associated with *Acacia salicina* and *A. stenophylla* (Mimosaceae) across south-eastern Australia. Int J Syst Evol Microbiol 61: 299-309.

51. Ndlovu J, Richardson DM, Wilson JR, Le Roux JJ (2013) Co-invasion of South African ecosystems by an Australian legume and its rhizobial symbionts. J Biogeogr 40: 1240-1251.

52. Klock MM, Barrett LG, Thrall P.. Harms KE (2015) Host promiscuity in symbiont associations can influence exotic legume establishment and colonization of novel ranges. Divers Distrib 21: 1193-1203.

53. Parker MA, Malek W, Parker IM (2006) Growth of an invasive legume is symbiont limited in newly occupied habitats. Divers Distrib 12: 563-571.

54. Davis JR, Goodwin LA, Woyke T, Teshima H, Bruce D, Detter C, Tapia R, Han S, Han J, Pitluck S (2012) Genome sequence of Amycolatopsis sp. strain ATCC 39116, a plant biomass-degrading actinomycete. J Bacteriol 194: 2396-2397.

55. Salvachúa D, Karp EM, Nimlos CT, Vardon DR, Beckham GT (2015) Towards lignin consolidated bioprocessing: simultaneous lignin depolymerization and product generation by bacteria. Green Chem 17: 4951-4967.

56. Lorenzo P González L, Reigosa MJ (2010) The genus Acacia as invader: the characteristic case of *Acacia dealbata* Link in Europe. Ann For Sci 67: 1-11.

57. Kumari R, Singh P, Lal R (2016) Genetics and genomics of the genus *Amycolatopsis*. Indian J Microbiol 56: 233-246.

58. Aro AO, Dzoyem JP, Hlokwe TM, Madoroba E, Eloff JN, McGaw LJ (2015) Some South African Rubiaceae tree leaf extracts have antimycobacterial activity against pathogenic and non-pathogenic *Mycobacterium* species. Phytother Res 29: 1004-1010.

59. Cosma CL, Sherman DR Ramakrishnan L (2003) The secret lives of the pathogenic Mycobacteria. Annu Rev Microbiol 57: 641-676.

60. Berg G, Eberl L, Hartmann A (2005) The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. Environ Microbiol 7: 1673-1685.

61. Berg G, Zachow C, Cardinale M, Müller H (2011) Ecology and human pathogenicity of plant-associated bacteria, regulation of biological control agents. Springer 175-189.

62. Morris RP, Nguyen L, Gatfield J, Visconti K, Nguyen K, Schnappinger D, Ehrt S, Liu Y, Heifets L, Pieters J (2005) Ancestral antibiotic resistance in *Mycobacterium tuberculosis*. Proc Natl Acad Sci USA 102: 12200-12205.

63. García-Fraile P, Menéndez E, Rivas R (2015) Role of bacterial biofertilizers in agriculture and forestry. AIMS Bioeng 2: 183-205.

64. Haansuu P, Vuorela P, Haahtela K (1999) Detection of antimicrobial and 45Ca2$^{+}$ transport blocking activity in Frankia culture broth extracts. Pharmaceutical and Pharmacological Letters 9: 1-4.

65. Barka EA, Vatsa P, Sanchez L, Gaveau-Vaillant N, Jacquard C, Klenk H-P, Clément C, Ouhdouch Y, van Wezel GP (2016) Taxonomy, physiology, and natural products of Actinobacteria. Microbiol. Mol Biol Rev 80: 1-43.

66. Dobbelaere S, Vanderleyden J, Okon Y (2003) Plant growth-promoting effects of diazotrophs in the rhizosphere. CRC Crit Rev Plant Sci 22: 107-149.

67. Aanderud ZT, Bledsoe CS (2009) Preferences for 15 N-ammonium, 15 N-nitrate, and 15 N-glycine differ among dominant exotic and subordinate native grasses from a California Oak woodland. Environ Exp Bot 65: 205-209.

68. Rossiter-Rachor N, Setterfield S, Douglas M, Hutley LB, Cook G, Schmidt S (2009) Invasive *Andropogon gayanus* (gamba grass) is an ecosystem transformer of nitrogen relations in Australian savanna. Ecol Appl 19: 1546-1560.

69. Dassonville N, Guillaumaud N, Piola F, Meerts P, Poly F (2011) Niche construction by the invasive Asian knotweeds (species complex Fallopia): impact on activity, abundance and community structure of denitrifiers and nitrifiers. Biol Invasions 13: 1115-1133.

70. Wardle D, Nicholson K, Ahmed M, Rahman A (1994) Interference effects of the invasive plant *Carduus nutans* L. against the nitrogen fixation ability of *Trifolium repens* L. Plant Soil 163: 287-297.

71. Dakora FD, Matiru VN, Kanu AS (2015) rhizosphere ecology of lumichrome and riboflavin, two bacterial signal molecules eliciting developmental changes in plants. Front Plant Sci 6: 1-11.

72. Palacios OA, Bashan Y, de-Bashan LE (2014) Proven and potential involvement of vitamins in interactions of plants with plant growth-promoting bacteria—an overview. Biol Fertil Soils 50: 415-432.

73. Gill SS, Tuteja N (2010) Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiol Biochem 48: 909-930.

74. Taheri P, Tarighi S (2010) riboflavin induces resistance in rice against *Rhizoctonia solani* via Jasmonate-mediated priming of phenylpropanoid pathway. J Plant Physiol 167: 201-208.

75. Brutinel ED, Dean AM, Gralnick JA (2013) Description of a riboflavin biosynthetic gene variant prevalent in the phylum *Proteobacteria*. J. Bacteriol 195: 5479-5486.

76. Hellmann H, Mooney S (2010) Vitamin B6: a molecule for human health? Molecules 15: 442-459.

77. Colinas M, Eisenhut M, Tohge T, Pesquera M, Fernie AR, Weber AP, Fitzpatrick TB (2016) Balancing of B6 vitamers is essential for plant development and metabolism in Arabidopsis. Plant Cell 28: 439-53.

78. Naamala J, Jaiswal SK, Dakora FD (2016) Antibiotics resistance in *Rhizobium*: Type, process, mechanism and benefit for agriculture. Curr Microbiol 72: 804-816.

## Data availability

The raw sequencing reads for this project were submitted to the National Centre for Biotechnology Information Short Read Archive under accession no. SRP098951.

## Acknowledgements

Funding for this research was provided by the National Research Foundation, South Africa (Grant No. CPRR14071676470 to AV). We thank Lesley Henderson, Fanus Venter and Emma Steenkamp for the information on *A. dealbata* populations.

## Author Contributions

A.V., T.P.M., J.J.LR., D.M.R. and D.A.C. designed research. C.N.K. and G.K.N performed the research. C.N.K., S.V. and A.V. analysed the data. C.N.K and A.V wrote the manuscript. All authors commented on the manuscript at all stages.