Plants of the Fynbos biome harbour host species-specific bacterial communities

Tsakani Miyambo, Thulani P. Makhalanyane, Don A. Cowan and Angel Valverde*

Department of Genetics, Centre for Microbial Ecology and Genomics, Genomics Research Institute (GRI), University of Pretoria, Pretoria, South Africa.

*Correspondence:
Angel Valverde, Department of Genetics, Centre for Microbial Ecology and Genomics, Genomics Research Institute (GRI), University of Pretoria, Lynnwood Road, Pretoria 0028, South Africa
Phone: (+27) 012 420 6944
Fax: (+27) 012 420 6870
Email: angel.valverde@up.ac.za

Running title: Bacterial endophytic diversity in fynbos plants

Keywords: bacterial endophytes, community phylogenetics, core microbiome, diversity, Fynbos, 16S rRNA amplicon sequencing,

Abstract

The Fynbos biome in South Africa is globally recognised as a plant biodiversity hotspot. However, very little is known about the bacterial communities associated with Fynbos plants, despite interactions between primary producers and bacteria impact the physiology of both partners and shape ecosystem diversity. This study reports on the structure, phylogenetic composition and potential roles of the endophytic bacterial communities located in the stems of three fynbos plants (*Erepsia anceps*, *Phaenocoma prolifera* and *Leucadendron laureolum*). Using Illumina MiSeq 16S rRNA sequencing we found that different subpopulations of Deinococcus-Thermus, Alphaproteobacteria, Acidobacteria and Firmicutes dominated the endophytic bacterial communities. Alphaproteobacteria and Actinobacteria were prevalent in *P. prolifera,* whereas Deinococcus-Thermus dominated in *L. laureolum*, revealing species-specific host-bacteria associations. Although a high degree of variability in the endophytic bacterial communities within hosts was observed. We also detected a core microbiome across the stems of the three plant species, which accounted for 72% of the sequences. Altogether, it seems that both deterministic and stochastic processes shaped microbial communities. Endophytic bacterial communities harboured putative plant growth-promoting bacteria, thus having the potential to influence host health and growth.

Introduction

Endophytic bacteria are ubiquitous in plants, where they play a pivotal role in plant health and performance (Partida-Martínez and Heil 2011). At the individual level they do so, for example, by producing phytohormones, enhancing nutrient acquisition and protecting plants against pathogens (reviewed in Sturz *et al.,* 2000). Bacterial endophytes have also been found to protect against abiotic stress such as salinity, heat and drought (Rosenblueth and Martínez-Romero, 2006; Lundberg *et al*., 2012 and references therein). At the community level, through their influence on plant diversity, performance and evolution, bacterial endophytes have the potential to influence plant biogeography and ecosystem functioning. For example, they facilitate the persistence and spread of invasive plant species (Rout *et al.,* 2013).

Plant endophytic bacteria are thought to be a subset of the rhizosphere microbiome (Bulgarelli *et al.,* 2013) and to be shaped by abiotic (e.g., soil chemistry) and biotic (e.g., plant species) factors (Philippot *et al.,* 2013). For instance, whereas deep profiling of the root microbiota of *Arabidopsis thaliana* ecotypes confirmed soil type as major source of variation in endophytic bacterial communities (Bulgarelli *et al.,* 2012; Lundberg *et al.,* 2012), plant species was more important than soil type in explaining the distribution of endophytic bacterial communities (Weber *et al.,* 1999). In fact, within a single plant species, different genotypes can host different endophytic bacterial communities (Hardoim *et al.,* 2011). Stochastic events, such as open wounds along the root hairs can also shape endophytic bacterial communities (Hardoim *et al.,* 2008). However, as most studies on the structure and composition of endophytic bacterial communities have focused on the roots of agricultural plants (e.g., Peiffer *et al.,* 2013) and model species such as *Arabidopsis thaliana* (e.g., Bulgarelli *et al.,* 2012; Lundberg *et al.,* 2012), the presence and relevance of endophytic bacterial communities in natural ecosystems, as well as the way they assemble, remain largely unexplored (but see Nissinen *et al.,* 2012; Carrell and Frank 2014).

The Fynbos biome, within the South African Cape Floristic Region (CFR), comprises three quite different, naturally fragmented vegetation types: fynbos, renosterveld and strandveld (Mucina and Rutherford, 2011). Fynbos is the predominant vegetation type, which is described as an evergreen, fire-prone shrubland characterised by the presence of restios (evergreen graminoids of the *Restioaceae*), a high cover of ericoid shrubs (principally in the families *Ericaceae*, *Asteraceae*, *Rhamnaceae*, *Thymelaeaceae* and *Rutaceae*), and the common occurrence of proteoid shrubs

(exclusively *Proteaceae*) (Mucina and Rutherford, 2011). The CFR contains over 8,000 plant species, most of which (69%) are endemic, making the region a biodiversity hotspot (Myers *et al.,* 2000). This is in spite of Fynbos plants being challenged by plant invasions (e.g., *Acacia* spp.; Sprent and Parsons 2000) and natural edaphic conditions such as low nutrient levels and water stress in the form of drought (Mucina and Rutherford, 2011). Fynbos plant diversity and endemicity is attributed to low migration rates and high speciation rates (Latimer *et al.,* 2005), but several studies suggest that plant-associated microorganisms play a role in sustaining plant communities (reviewed in van der Heijden *et al.,* 2008). Indeed, studies have shown that microbial diversity can be high in fynbos soils (Slabbert *et al.,* 2010), and that the rhizospheres of *Proteaceae* species are extensively colonised by members of the Proteobacteria, Firmicutes, Actinobacteria and Acidobacteria (Stafford *et al.,* 2005). Nevertheless, nothing is currently known on the extent to which plants in the Cape Floristic Region contain bacterial endophytes and the factors that shape those communities. This is of relevance because, for example, the restoration and conservation of plant diversity, as well as the prevention of exotic plant invasions, require a better understanding of the association between microorganisms and plants (Philippot *et al.,* 2013).

In this study, we used terminal restriction fragment length polymorphisms (T-RFLP) and amplicon sequencing of the 16S rRNA genes, to investigate endophytic bacterial communities in the stems of three different fynbos plant species (*Erepsia anceps*, *Phaenocoma prolifera* and *Leucadendron laureolum*) belonging to three different families (*Mesembryanthemaceae*, *Asteraceae* and *Proteaceae*, respectively). The three plant species were selected on the basis that they were at the same stage of growth (early flowering) and co-existed in several locales within a geographic area (Fernkloof Nature Reserve) of the fynbos biome, South Africa. This allowed us to test whether or not co-occurring plant species display different microbiota. As in natural ecosystems plants are growing in soils with long-term co-evolving microorganisms (Philippot *et al.,* 2013), we expect plant species to be a strong determinant of the endophytic bacterial community. Consequently, we predict plants should harbour species-specific bacterial communities. We explain patterns in community composition by applying community phylogenetics and modern ecological theory.

Materials and methods

Sample collection and sterilization

Samples were collected in the Fernkloof Nature Reserve (FNR; Supplementary Figure S1), Hermanus, Western Cape (South Africa) on the 25th of March 2014. The area covers 1383 hectares and is host to 1474 different plant species (http://fernkloof.com). The local climate is warm-temperate with warm, dry summers and mild, wet winters. Stems (ca. 15 cm) of side-by-side growing *E. anceps*, *P. prolifera* and *L. laureolum* (fifteen each) were collected at 3 sites with similar macro-environmental conditions (that is, slope, aspect, elevation) along a ca. 3-km transect using sterile scissors and gloves. The material was stored at 4 °C and processed within 48 h after harvesting. Samples (n=45; 5 stems x 3 plant species x 3 locations) were thoroughly washed with tap water, cut into sections (~2 cm long), surface sterilized by immersion into 2% NaOCl (3 min) and 2% sodium thiosulphate (3 min), and rinsed three times with sterile distilled water (3 x 3 min). Surface-sterilized stem fragments were put on R2A plates and checked for sterility after 4-60 days of incubation at 25ºC.

DNA extraction and Terminal-restriction fragment length polymorphism (T-RFLP) analysis

Stem fragments (n=45) were ground to a fine powder in a sterile mortar with liquid nitrogen and DNA extracted by using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions. Bacterial 16S rRNA gene amplification was generated using primers 799F2 (5'-AACMGGATTAGATACCCGG-3') and 1193R (5'-ACGTCATCCCCACCTTCC-3') following the PCR conditions previously described (Bulgarelli *et al.,* 2012). The primers amplify the hypervariable regions V5-V6-V7 of the 16S rRNA gene. Primer 799F2 was shown to not compete with plant organellar 18S RNA gene templates, allowing the preferential amplification of the bacterial 16S rRNA gene (Bulgarelli *et al.,* 2012). The forward primer was labelled with 6' carboxyfluorescein (6-FAM). Purification, digestion with *Hae*III, separation of fragments, evaluation of electrophoretic signals and subsequent binning into operational taxonomic units (OTUs) were performed as reported elsewhere (Valverde *et al.,* 2012).

16S rRNA amplicon sequencing and analysis

Bacterial amplicons were generated as described above but using unlabelled primers. All amplicon products from different samples (n=18, 2 randomly selected stems x 3 plant species x 3 sites) were mixed in equal concentrations and purified using Agencourt Ampure beads (Agencourt Bioscience Corporation, MA, USA). Sequencing was carried out on an Illumina MiSeq instrument (Illumina Inc., CA, USA) using paired-end sequencing at the Molecular Research LP next generation sequencing service (http://www.mrdnalab.com, Shallowater, TX, USA). Sequences were analysed in MOTHUR (Schloss *et al.,* 2009), following a previously established pipeline (http://www.mothur.org/wiki/MiSeq_SOP; Kozich *et al.,* 2013), using the Silva core set for alignment and MEGA5 (Tamura *et al.,* 2011) to estimate a phylogeny containing all OTUs observed across all samples. Reads were removed from further analysis if at least one of the following criteria was met: (i) reads shorter than 200 bp, (ii) presence of ambiguous bases, and (iii) presence of homopolymers with more than 8 bp. Chimeras, chloroplast- and mitochondria-related reads were also removed. Sequences were grouped into OTUs, defined using a 97% sequence similarity cut-off (Schloss and Handelsman 2004), and taxonomic identities were assigned to phylotypes using the RDP classifier (Wang *et al.,* 2007) with a confidence threshold of 0.8. Singletons were removed, as they are generally assumed to be due to sequencing errors/artefacts (Quince *et al.,* 2008). Each sample was rarefied to 1224 sequences, the lowest number of sequences in any sample.

The sequence data generated in this study were deposited in the NCBI Sequence Read Archive and are available under the project number SRP059346. A table containing sample metadata is provided as supplementary material (Supplementary Table S1).

Statistical analysis

All data were analysed using the vegan, labdsv, phia and picante packages for R (R Development Core Team 2013), together with custom scripts. Bacterial diversity metrics were compared by mixed ANOVA analysis. Data were log transformed to meet the assumptions of normality. Bacterial community structure was visualised using non-metric multidimensional scaling (nMDS) obtained with Jaccard (T-RFLP data) and unweighted UniFrac (sequencing data) dissimilarity matrices. A permutational analysis of variance (PERMANOVA) (Anderson 2001) was used to test for differences in bacterial composition between plant species. Site of sampling was included in the PERMANOVA and ANOVA models as a random factor. To identify

bacterial lineages as analogous to the concept of "indicator species" for each plant species, we used the indicator value (IndVal) index, which combines relative abundance and relative frequency of occurrence (Dufrene and Legendre 1997). The phylogenetic structure of the bacterial communities was quantified using null models (Kembel 2009). We estimated the abundance-weighted mean pairwise distance (MPD) and the mean nearest phylogenetic taxon distance (MNTD) among sequences in each sample and calculated a standardized effect size (SES.MPD and SES.MNTD) (Kembel 2009). The null model algorithm used was "independentswap" with 999 randomized null communities. Independentswap retains column and row totals for null model analysis of species co-occurrence (Gotelli 2000). This approach is particularly suited as it accounts for variations in diversity and richness between communities.

Results and discussion

Plants harbour endophytic bacterial communities, members of which have been shown to positively affect both the health and diversity of their hosts (Berendsen *et al.,* 2012). However, most of this research has concentrated on the rhizo- and phyllo-endosphere (but see Nissinen *et al.,* 2012). Here we describe the endophytic bacterial communities inhabiting the stems of three plant species co-occurring in a natural habitat, the fynbos biome.

Using T-RFLP analysis a total of 34 bacterial OTUs were detected among the 45 samples, of which 11 (32%) were shared between the three different plant species and 18 (53%) were unique to the respective plant species (Supplementary Figure S2). The number of T-RFLPs-derived OTUs per sample (alpha-diversity) ranged from 4 to 20. Overall, *L. laureolum* contained lower bacterial OTU numbers than *E. anceps* and *P. prolifera*: 23, 23 and 15 OTUs were observed in *E. anceps*, *P. prolifera* and *L. laureolum*, respectively (Supplementary Figure S2). It could be argued that fingerprinting techniques are not well suited for diversity analyses, as they are known to underestimate community diversity. However, the reason for the use of T-RFLPs was not to assess the diversity of the samples (see below), but rather to visualize differences in the dominant members of the endophytic bacterial community composition (beta-diversity) between the three different plant species in a cost-effective manner. Several studies have shown that fingerprinting methods lead to reliable beta-diversity patterns (Gobet *et al.,* 2014; Valverde *et al.,* 2014). We did not

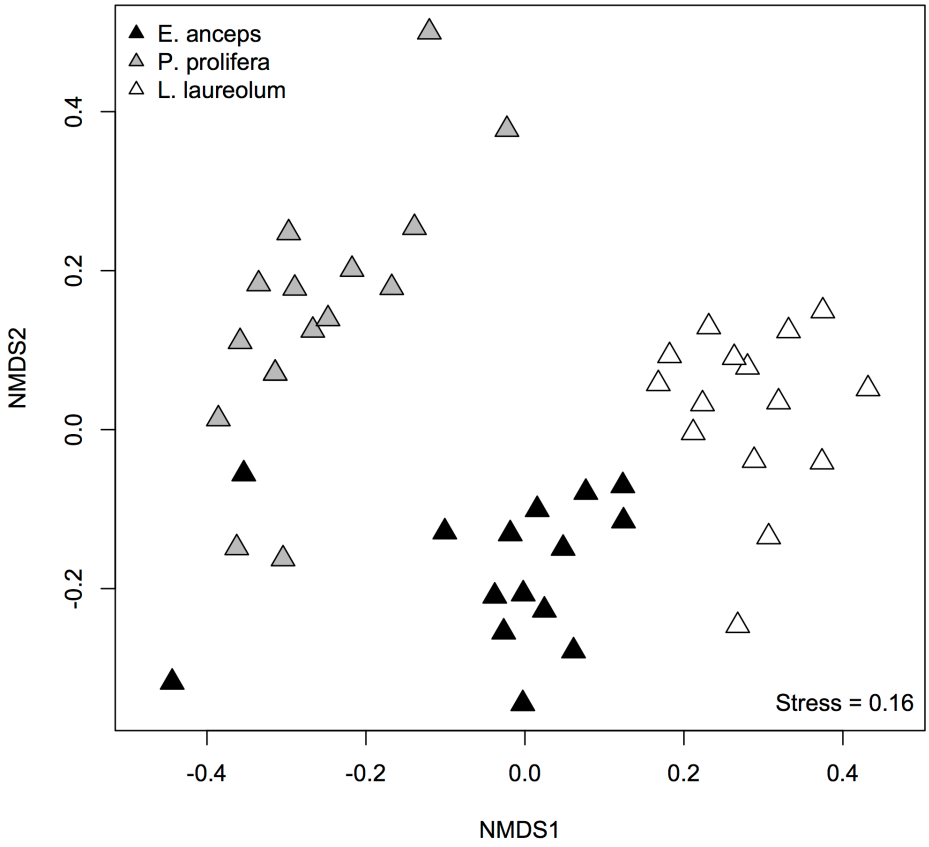


Figure 1. NMDS ordination plot of the endophytic bacterial communities based on T- RFLP-based Jaccard distances among samples. Different plant species showed distinct bacterial communities (PERMANOVA: $F_{1,41} = 11.34$, $P = 0.001$). Pairwise comparisons: *E. anceps* vs. *P. prolifera* $R^2 = 0.22$, *E. anceps* vs. *L. laureolum* $R^2 = 0.22$, *L. laureolum* vs. *P. prolifera* $R^2 = 0.25$ (all $P < 0.05$).

observe any effect of 'site' in shaping the bacterial communities ($PERMANOVA_{jaccard}$ $P>0.05$). In contrast, as hypothesized, 'plant species' was an important factor explaining bacterial community composition (Figure 1; $PERMANOVA_{jaccard}$: coefficient of determination $R^2 = 0.17$; $P < 0.01$). Different plant species likely represent different ecological niches for endophytic bacteria, each of which seems to maintain a specific microbial community (Truyens *et al.,* 2015).

To further investigate the diversity and phylogenetic structure of the bacterial communities, a subset of 18 non-pooled samples was randomly selected (2 stems x 3 plants x 3 sites) for amplicon sequencing analysis of the 16S rRNA gene using the Illumina platform. After quality control, removal of chimeras, chloroplast- and mitochondria-related reads, and singletons: 22032 sequences were retained, representing 78 OTUs. These low numbers of stem endophytic bacteria OTUs were not completely unexpected, as other studies have reported similar findings for bacterial communities in the interior of roots (Bulgarelli *et al.,* 2012; Schlaeppi *et al.,* 2014).

Bacterial diversity was generally higher for *P. prolifera* (ANOVA test, $P < 0.05$) using a variety of indices (Figure 2a), including richness. *E. anceps* presented 19.3 ± 2.7 (average ± SE) different OTUs per sample, while 22.0 ± 4.8 and 6 ± 0.7 different OTUs were found in *P. prolifera* and *L. laureolum*, respectively. At the plant level, 49, 56 and 16 OTUs were observed in *E. anceps*, *P. prolifera* and *L. laureolum*, respectively (Figure 3), in accordance with the trends observed in T-RFLP analysis. However, the Chao, Jackknife, and Bootstrapping estimates of OTU richness indicated that the percentages detected accounted for 70-97% of the total community richness for each plant species, depending of the particular estimated used (Figure 2b), and that additional bacterial taxa are likely to appear in further samples.

Of these 78 OTUs, 58 could be classified at the phylum level (making up 98.5% of all sequences) (Figure 4): 38 belonged to Proteobacteria [28 OTUs were Alphaproteobacteria (23.4%), 5 Betaproteobacteria (0.7%) and 5 Gammaproteobacteria (0.7%)], 7 belonged to Deinococcus-Thermus (59.5%), 7 to Actinobacteria (5.7%), 3 to Acidobacteria (9.1%) and 3 to Firmicutes (0.1%). All these phyla have been found as plant endophytes (Sun *et al.,* 2008; Bulgarelli *et al.,* 2012; Nissinen *et al.,* 2012; Bodenhausen *et al.,* 2013; Truyens *et al.,* 2015), in the phylosphere (Knief *et al.,* 2010; Leff *et al.,* 2015) and in soil (Bulgarelli *et al.,* 2012;

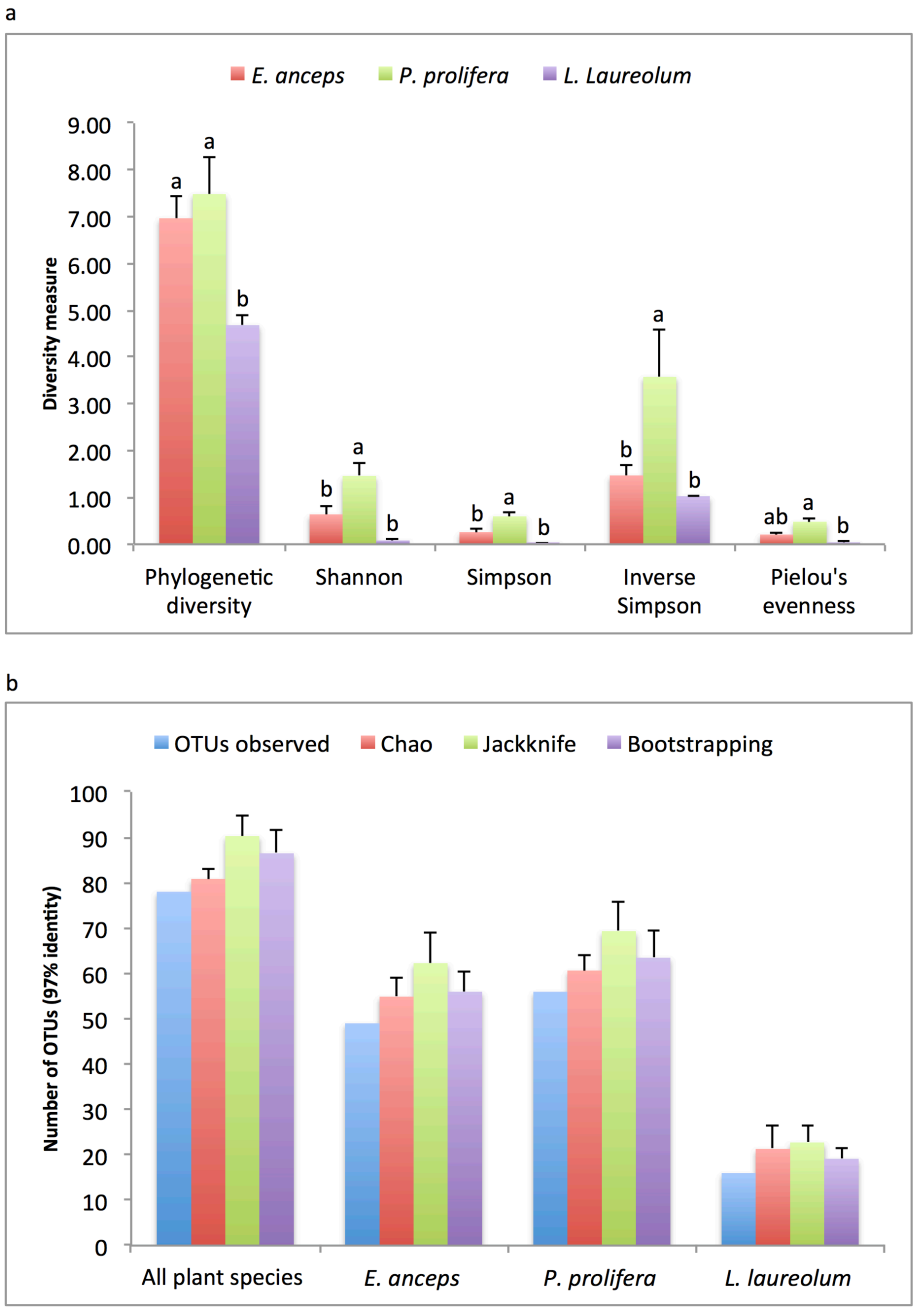


Figure 2. Diversity metrics and richness estimation, based on sequencing data, of endophytic bacterial communities for the three plant species. a) Different diversity metrics (average ± SE). Different letters indicate significant differences ($P < 0.05$). b) Number of total bacterial OTUs detected in each plant species together with the total number of OTUs (average ± SE) estimated by different approaches.

Zarraonaindia *et al.,* 2015), although at different relative abundances. Strikingly, the phyla Deinococcus-Thermus and Acidobacteria seem to be overrepresented here relative to other studies (Sun *et al.,* 2008; Nissinen *et al.,* 2012). This is probably not caused by primer bias, as the primer pair used in this study has been used with Arabidopsis plants resulting in the dominance of Chloroflexi, Proteobacteria, Actinobacteria and Bacteroidetes (Bulgarelli *et al.,* 2012). Fynbos soils are nutrient-poor, particularly with regard to N and P (Richards *et al.,* 1997), which might favour oligotrophic bacteria such as those typically found within the Deinococcus-Thermus and Acidobacteria phyla (Fierer *et al.,* 2007; Meola *et al.,* 2015). The fact that these soils are also typically acid (Richards *et al.,* 1997), might further explain the relatively high proportion of Acidobacteria (Jones *et al.,* 2009).

We noticed that the predominance of Deinococcu-Thermus and Alphaproteobacteria was driven by the high abundance of only 6 OTUs (representing 80% of all sequences) (Supplementary Figure S3). These OTUs were classified as *Deinococcus* (52.0%), Rhizobiales (11.2%), Deinococcales (5.0%), *Sphingomonas* (4.6%), Acetobacteraceae (3.8%) and *Truepera* (2.6%), respectively. Bacteria from the phylum Deinococcus-Thermus are known to possess important adaptations to environmental stresses, such as desiccation, ultraviolet radiation, high salinity and high temperatures (Battistuzzi and Hedges 2009). Members of *Truepera* may play an essential role in plant development and physiology, due to their ability to produce brassinosteroids, which have been shown to control seed germination, stem and root elongation, vascular differentiation, leaf expansion, and stress protection in plants (Bajguz and Hayat 2009). Some *Sphingomonas* strains have a protective effect against plant pathogens (Innerebner *et al.,* 2011). The Rhizobiales and Acetobacteraceae include various well-known nitrogen fixing plant symbionts (Lugtenberg and Kamilova 2009). Therefore it seems that a significant proportion of the phylotypes identified as members of the endophytic community has the potential to positively influence the fitness of their host. Similar results have been found in the endosphere of pine needles (Carrell and Frank 2014), supporting the growing concept that PGPB are important for plant health in natural ecosystems (Compant *et al.,* 2010). However, more research is needed to elucidate the role of these bacterial endophytes, as it is well known that plant growth promoting characteristics are strain dependent.

Non-metric multidimensional scaling (nMDS) plots and PERMANOVA analysis, based on unweighted UniFrac (Supplementary Figure S4) distances obtained with the

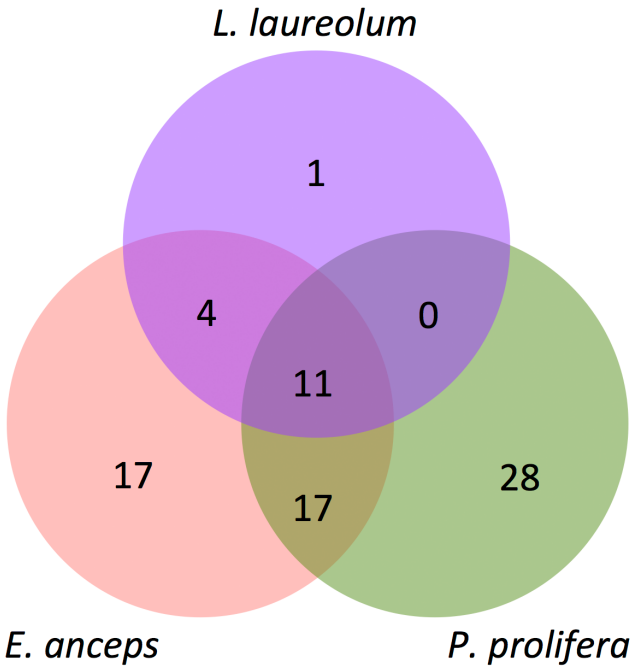


Figure 3. Venn diagram of the shared and specific bacterial OTUs (97% identity) found in *E. anceps, P. prolifera* and *L. laureolum.*

sequencing data, confirmed that the three endophytic bacterial communities were taxonomically and phylogenetically distinct. More in depth analysis showed that 28 OTUs were unique to *P. prolifera*, 17 unique to *E. anceps* and 1 unique to *L. laureolum* (Figure 3), a greater number (14 OTUs) being Alphaproteobacteria.
We found 16 OTUs to be indicators for the three plant species (6 for *E. anceps*, 9 for *P. prolifera* and 1 for *L. laureolum*) (Figure 5). If hosts select these bacteria on the basis of their functional capacities (Burke *et al.,* 2011), indicator lineages should be included in the pivotal ecological functions within the specific plant species. Members of the Alphaproteobacteria and Actinobacteria were enriched in *P. prolifera* and Deinococcus-Thermus in *L. laureolum* (Figure 4). This suggests that the environment created by the host (host filtering), for example due to differences in immune systems (Van der Heijden and Schlaeppi 2015), may play an important role in the assembly of these communities. Noteworthy, it has recently been shown experimentally that the foliar defence phytohormone salicylic acid directly shapes the root endophyte microbiome of *Arabidopsis thaliana* plants (Lebeis *et al.,* 2015).
However, in spite of the observed host species-specific effects, communities were highly variable from plant to plant. On average, 18% of the OTUs were shared among *E. anceps* samples, 17% among *P. prolifera* samples and 12% among *L. laureolum* samples. Moreover, we detected a set of 11 (14% of all OTUs) common and, in most cases, abundant (they accounted for 72.5% of the sequences) OTUs between the three plant species (Figure 3). Source/soil-sink/endosphere dynamics can generate this pattern, as species are not excluded from communities where they are bad competitors because they immigrate from other communities where they are good competitors (Mouquet and Loreau 2003). Together, these results appear to indicate that stochastic variation is important in explaining bacterial community assembly in the interior of fynbos plants.
As both deterministic (niche/host-based) and stochastic (neutral) processes appeared to have shaped endophytic bacterial communities, we assessed the degree of phylogenetic relatedness using null models (Webb *et al.,* 2002). Community phylogenetics has been used successfully to infer mechanisms of community assembly in plant host-associated bacteria (Kembel *et al.,* 2014). Most samples showed evidence of neutral processes, with five showing evidence of phylogenetic clustering or overdispersion (Figure 6). This result suggests that these communities are shaped by stochastic factors. At least three non-exclusive explanations can be

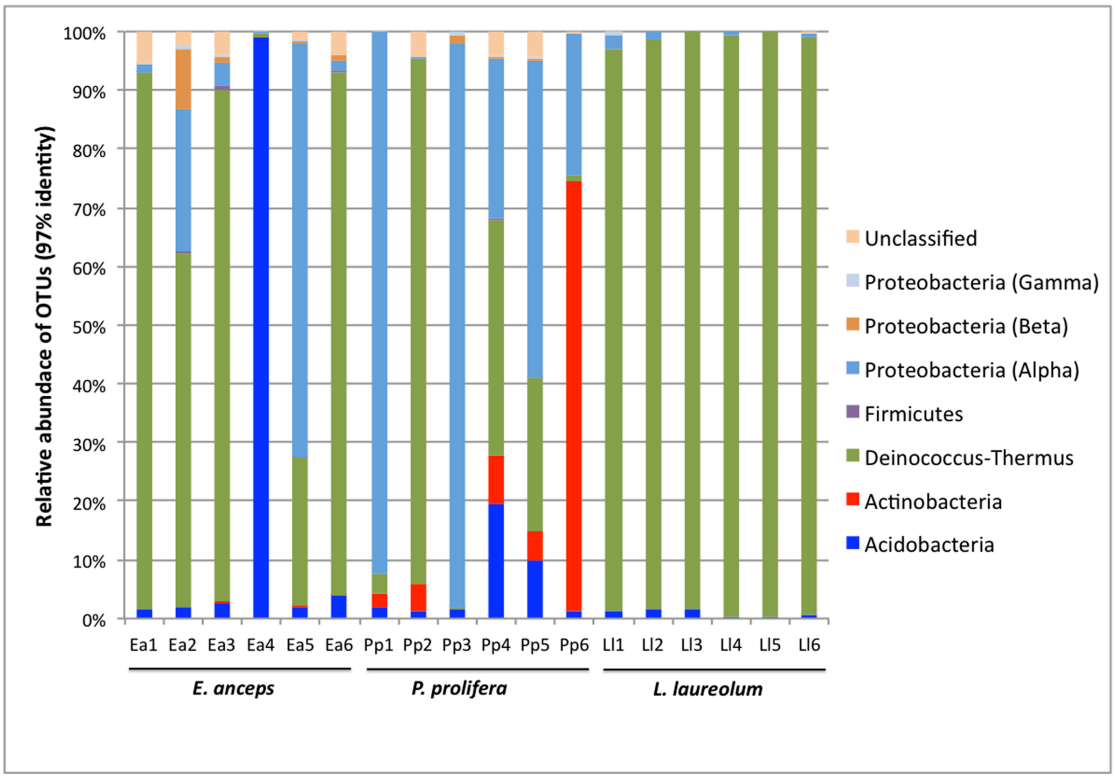


Figure 4. Taxonomic distribution of endophytic bacterial OTUs (97% identity). Affiliation was performed using the Ribosomal Database Project Classifier with a confidence threshold of 80%.

OTU	*E. anceps*	*P. prolifera*	*L. laureolum*	Indicator for	Indicator value	Probability	Taxonomic affiliation
31	0.12	0.02	0.00	*E. anceps*	0.71	0.01	Proteobacteria
16	0.27	0.11	0.03	*E. anceps*	0.67	0.00	Proteobacteria
52	0.04	0.00	0.00	*E. anceps*	0.67	0.02	Proteobacteria
68	0.06	0.00	0.00	*E. anceps*	0.67	0.02	Proteobacteria
37	0.05	0.01	0.00	*E. anceps*	0.55	0.04	Proteobacteria
60	0.05	0.01	0.00	*E. anceps*	0.55	0.05	Betaproteobacteria;Burkholderiales
04	0.00	5.01	0.00	*P. prolifera*	1.00	0.001	Deinococcus-Thermus;Deinococcales
02	0.90	10.08	0.17	*P. prolifera*	0.90	0.002	Alphaproteobacteria;Rhizobiales
05	0.01	2.55	0.00	*P. prolifera*	0.83	0.00	Deinococcus-Thermus;*Truepera*
08	0.06	0.65	0.00	*P. prolifera*	0.76	0.01	Alphaproteobacteria;Rhizobiales
10	0.00	0.49	0.00	*P. prolifera*	0.67	0.02	Actinobacteria
11	0.00	0.28	0.00	*P. prolifera*	0.67	0.02	Deinococcus-Thermus;Deinococcales
14	0.00	0.24	0.00	*P. prolifera*	0.67	0.02	Deinococcus-Thermus;Deinococcales
07	0.02	4.25	0.00	*P. prolifera*	0.66	0.02	Actinobacteria;*Frondihabitans*
09	0.02	4.53	0.01	*P. prolifera*	0.66	0.04	Alphaproteobacteria;*Sphingomonas*
01	13.17	1.84	37.02	*L. laureolum*	0.71	0.001	Deinococcus-Thermus;*Deinococcus*

Figure 5. Heat map showing indicator OTUs (97% identity) and their relative abundances as percentages of all sequences. Taxonomic assignments are the finest level that passed the Ribosomal Database Project classifier's (80% confidence threshold).

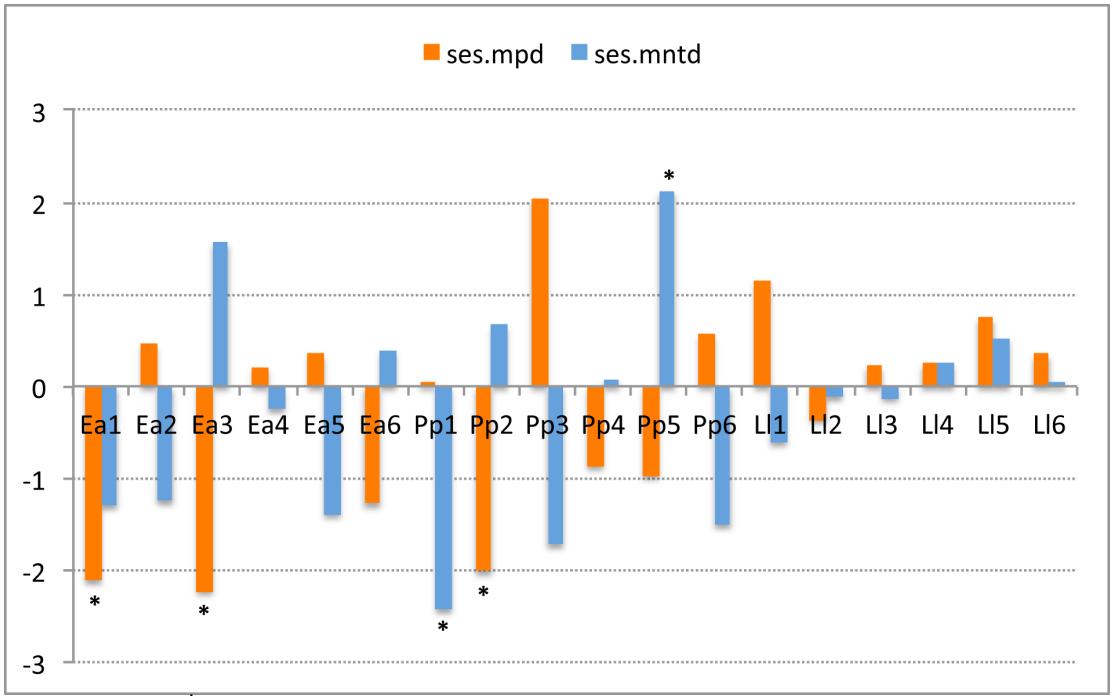


Figure 6. Standardized effect sizes of MPD (SES.MPD) and MNTD (SES.MNTD) for endophytic bacterial communities. Asterisks indicate significant results for phylogenetic clustering (negative values, mpd.obs.p < 0.05) or overdispersion (positive values, mpd.obs.p > 0.95) in bacterial communities.

offered for the relatively weak deterministic effect. Firstly, it could be a consequence of the methodological approach used, as it has been shown that the assembly and structure of plant-associated bacterial communities may be better explained by functional gene analysis than by the use of 16S rRNA genes (Ofek-Lalzar *et al.,* 2014). Nevertheless, a recent study has demonstrated that analyses using 16S rRNA-derived OTUs tend to indicate habitat filtering as the dominant driver of community assembly (Koeppel and Wu 2014). Secondly, critical deterministic traits may not differ enough between the three plant species we sampled. These could include, for instance, the type and availability of nutrients and the amount or type of antimicrobial compounds produced by the plants (Westoby *et al.,* 2002). Alternatively, and most likely, random colonization through cracks, could buffer the effect of host filtering by continued homogenization of host-associated communities (Costello *et al.,* 2012). Indeed, random colonization has been proposed to be important in explaining the diversity and community composition of endophytic microbial communities (Hardoim *et al.,* 2008).

The fact that some samples showed phylogenetic clustering or overdispersion may indicate localised host filtering and/or competitive exclusion. According to modern coexistence theory (Mayfield and Levine 2010), species coexistence is driven by the interaction of two types of species differences: niche differences and competitive ability differences. Under the Mayfield and Levine model, phylogenetic clustering can arise from either competition or habitat/host filtering, but overdispersion can only result from competition. Clearly, further research is needed to understand the relative contribution of the processes that structure plant endophytic bacterial communities in the fynbos biome.

In summary, we have revealed that fynbos plants host endophytic bacterial communities of relatively limited, but taxonomically diverse bacterial species. Several taxa were found to be specialists, while others were present in all host plants (generalists). It remains to be elucidated, however, whether these endophytes represent taxa living permanently in association with the host or transient taxa introduced from the soil and other environments. Using a culture-independent technique we showed that endophytic microbial communities contained potential plant-beneficial bacteria. Culture-dependent studies in conjunction with metagenomics and single-cell genomics will further elucidate the role of these communities in influencing host health and diversity.

Conflict of interest

The authors declare no conflict of interest

Funding

This work was supported by the National Research Foundation (South Africa).

Acknowledgements

We are grateful to the Sequencing Facility at the University of Pretoria for performing T-RFLP analysis, and the administrators of the Fernkloof Nature Reserve to allow us to collect the samples.

References

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

Bajguz A, Hayat S. Effects of brassinosteroids on the plant responses to environmental stresses. *Plant Physiol Biochem* 2009;**47**:1-8.

Battistuzzi FU, Hedges SB. A major clade of prokaryotes with ancient adaptations to life on land. *Mol Biol Evol* 2009;**26**:335-43.

Berendsen RL, Pieterse CMJ, Bakker PAHM. The rhizosphere microbiome and plant health. *Trends Plant Sci* 2012;**17**:478-86.

Bodenhausen N, Horton MW, Bergelson J. Bacterial communities associated with the leaves and the roots of *Arabidopsis thaliana*. *PLoS ONE* 2013;**8**.

Bulgarelli D, Rott M, Schlaeppi K *et al*. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* 2012;**488**:91-5.

Bulgarelli D, Schlaeppi K, Spaepen S *et al* (2013). Structure and functions of the bacterial microbiota of plants. *Annu Revi Plant Biol*;**64**:807-38.

Burke C, Steinberg P, Rusch D *et al*. Bacterial community assembly based on functional genes rather than species. *Proc Natl Acad Sci USA* 2011;**108**:14288-93.

Carrell AA, Frank AC. *Pinus flexilis* and *Picea engelmannii* share a simple and consistent needle endophyte microbiota with a potential role in nitrogen fixation. *Front Microbiol* 2014;**5**.

Compant S, Clement C, Sessitsch A. Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. *Soil Biol Biochem* 2010;**42**:669-678.

Costello EK, Stagaman K, Dethlefsen L *et al*. The application of ecological theory toward an understanding of the human microbiome. *Science* 2012;**336**:1255-62.

Dufrene M, Legendre P. Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecol Monogr* 1997;**67**:345-66.

Fierer N, Bradford MA, Jackson RB. Toward an ecological classification of soil bacteria. *Ecology* 2007;**88**:1354-1364.

Gobet A, Boetius A, Ramette A. Ecological coherence of diversity patterns derived from classical fingerprinting and Next Generation Sequencing techniques. *Environ Microbiol* 2014;**16**:2672-81.

Gotelli NJ. Null model analysis of species co-occurrence patterns. *Ecology* 2000;**81**:2606-21.

Hardoim PR, van Overbeek LS, van Elsas JD. Properties of bacterial endophytes and their proposed role in plant growth. *Trends Microbiol* 2008;**16**:463-71.

Hardoim PR, Andreote FD, Reinhold-Hurek B *et al*. Rice root-associated bacteria: Insights into community structures across10 cultivars. *FEMS Microbiol Ecol* 2011;**77**:154-64.

Innerebner G, Knief C, Vorholt JA. Protection of *Arabidopsis thaliana* against leaf-pathogenic *Pseudomonas syringae* by *Sphingomonas* strains in a controlled model system. *Appl Environ Microbiol* 2011;**77**:3202-10.

Jones RT, Robeson MS, Lauber CL *et al*. A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *ISME J* 2009;**3**:442-453.

Kembel SW. Disentangling niche and neutral influences on community assembly: assessing the performance of community phylogenetic structure tests. *Ecol Lett* 2009;**12**:949-960.

Kembel SW, O'Connor TK, Arnold HK *et al*. Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. *Proc. Natl. Acad. Sci. USA* 2014;**111**:13715-20.

Knief C, Ramette A, Frances L *et al*. Site and plant species are important determinants of the *Methylobacterium* community composition in the plant phyllosphere. *ISME J* 2010;**4**:719-28.

Koeppel AF, Wu M. Species matter: The role of competition in the assembly of congeneric bacteria. *ISME J* 2014;**8**:531-40.

Kozich J, Westcott SL, Baxter NT *et al*. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq illumina sequencing platform. *Appl Environ Microbiol* 2013;**79**:5112-20.

Latimer AM, Silander JA, Cowling RM. Neutral ecological theory reveals isolation and rapid speciation in a biodiversity hot spot. *Science* 2005;**309**:1722-5.

Lebeis SL, Paredes SH, Lundberg DS *et al*. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* 2015;**349**:860-4.

Leff JW, Del Tredici P, Friedman WE *et al*. Spatial structuring of bacterial communities within individual Ginkgo biloba trees. *Environ Microbiol* 2015;**17**:2352-61.

Lugtenberg B, Kamilova F. Plant-Growth-Promoting Rhizobacteria. *Annu Rev Microbiol* 2009;**63**:541-56.

Lundberg DS, Lebeis SL, Paredes SH *et al*. Defining the core *Arabidopsis thaliana* root microbiome. *Nature* 2012;**488**:86-90.

Mayfield MM, Levine JM. Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecol Lett* 2010;**13**:1085-93.

Meola M, Lazzaro A, Zeyer J. Bacterial composition and survival on Sahara dust particles transported to the European Alps. *Front Microbiol* 2015;**6**.

Mouquet N, Loreau M. Community patterns in source-sink metacommunities. *Am Nat* 2003;**162**:544-7.

Mucina L, Rutherford MC. The vegetation of Souht Africa, Lesotho and Swaziland. Pretoria: South African National Biodiversity Institute, 2011.

Myers N, Mittermeier RA, Mittermeier CG *et al*. Biodiversity hotspots for conservation priorities. *Nature* 2000;**403**:853-8.

Nissinen RM, Männistö MK, van Elsas JD. Endophytic bacterial communities in three arctic plants from low arctic fell tundra are cold-adapted and host-plant specific. *FEMS Microbiol Ecol* 2012;**82**:510-22.

Ofek-Lalzar M, Sela N, Goldman-Voronov M *et al*. Niche and host-associated functional signatures of the root surface microbiome. *Nature Commun* 2014;**5**.

Partida-Martínez LP, Heil M. The microbe-free plant: Fact or artifact? *Front Plant Sci* 2011;**2**.

Peiffer JA, Spor A, Koren O *et al*. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc Natl Acad Sci USA* 2013;**110**:6548-53.

Philippot L, Raaijmakers JM, Lemanceau P *et al*. Going back to the roots: the microbial ecology of the rhizosphere. *Nature Rev Microbiol* 2013;**11**:789-99.

Quince C, Curtis TP, Sloan WT. The rational exploration of microbial diversity. *ISME J* 2008;**2**:997-1006.

R Development Core Team (2013). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing;* http://wwwR-projectorg/.

Richards MB, Stock WD, Cowling RM. Soil nutrient dynamics and community boundaries in the fynbos vegetation of South Africa. *Plant Ecol* 1997;**130**:143-153.

Rosenblueth M, Martinez-Romero E. Bacterial endophytes and their interactions with hosts. *MPMI* 2006;**19**:827-837.

Rout ME, Chrzanowski TH, Westlie TK *et al*. Bacterial endophytes enhance competition by invasive plants. *Am J Bot* 2013;**100**:1726-37.

Schlaeppi K, Dombrowski N, Oter RG *et al*. Quantitative divergence of the bacterial root microbiota in *Arabidopsis thaliana* relatives. *Proc Natl Acad Sci USA* 2014;**111**:585-92.

Schloss PD, Handelsman J. Status of the microbial census. *Microbil Mol Biol Rev* 2004;**68**:686-691.

Schloss PD, Westcott SL, Ryabin T *et al*. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 2009;**75**:7537-41.

Slabbert E, Kongor RY, Esler KJ *et al*. Microbial diversity and community structure in Fynbos soil. *Mol Ecol* 2010;**19**:1031-41.

Sprent JI, Parsons R. Nitrogen fixation in legume and non-legume trees. *Field Crop Res* 2000;**65**:183-96.

Stafford WHL, Baker GC, Brown SA *et al*. Bacterial diversity in the rhizosphere of Proteaceae species. *Environ Microbiol* 2005;**7**:1755-68.

Sturz AV, Christie BR, Nowak J. Bacterial endophytes: Potential role in developing sustainable systems of crop production. *Cr Rev Plant Sci* 2000;**19**:1-30.

Sun L, Qiu F, Zhang X *et al*. Endophytic bacterial diversity in rice (*Oryza sativa* L.) roots estimated by 16S rDNA sequence analysis. *Microb Ecol* 2008;**55**:415-24.

Tamura K, Peterson D, Peterson N *et al*. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011;**28**:2731-9.

Truyens S, Weyens N, Cuypers A *et al*. Bacterial seed endophytes: Genera, vertical transmission and interaction with plants. *Environ Microbiol Rep* 2015;**7**:40-50.

Valverde A, Tuffin M, Cowan DA. Biogeography of bacterial communities in hot springs: A focus on the actinobacteria. *Extremophiles* 2012;**16**:669-79.

Valverde A, Makhalanyane TP, Cowan DA. Contrasting assembly processes in a bacterial metacommunity along a desiccation gradient. *Front Microbiol* 2014;**5**.

Van Der Heijden MGA, Schlaeppi K. Root surface as a frontier for plant microbiome research. *Proc Natl Acad Sci USA* 2015;**112**:2299-2300.

van der Heijden MGA, Bardgett RD, van Straalen NM. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol Lett* 2008;**11**:296-310.

Wang Q, Garrity GM, Tiedje JM *et al*. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 2007;**73**:5261-7.

Webb CO, Ackerly DD, McPeek MA *et al*. Phylogenies and community ecology. *Annu Rev Ecol Syst* 2002;**33**:475-505.

Weber O, Baldani V, Teixeira K. Isolation and characterization of diazotrophic

bacteria from banana and pineapple plants. *Plant Soil* 1999;**210**:103-113.

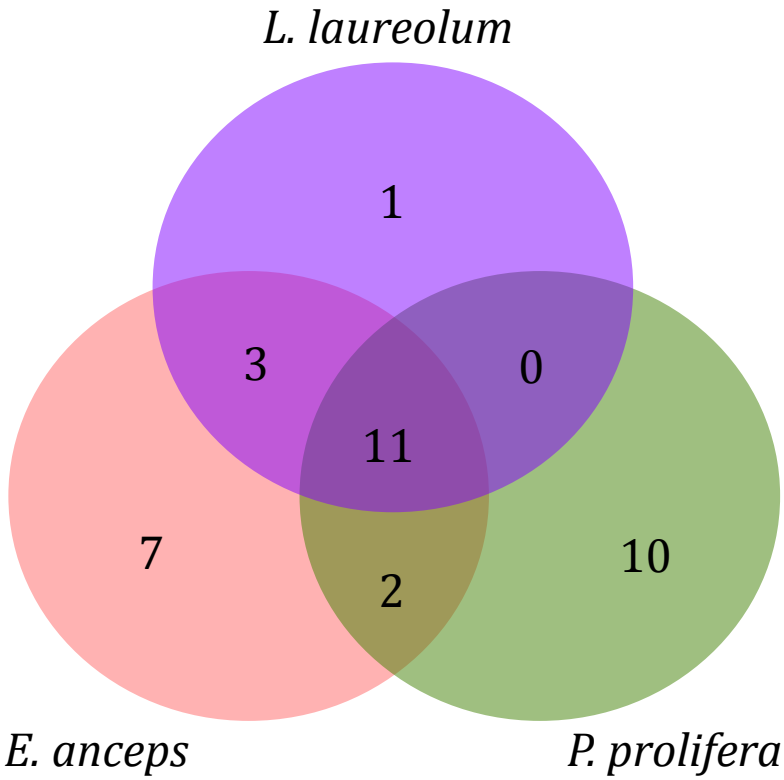
Westoby M, Falster DS, Moles AT *et al*. Plant ecological strategies: Some leading dimensions of variation between species. *Annu Rev Ecol Syst* 2002;**33**:125-59.

Zarraonaindia I, Owens SM, Weisenhorn P *et al*. The soil microbiome influences grapevine-associated microbiota. *mBio* 2015;**6**.

Supplementary material

Supplementary Figure S1. Aerial picture of the Fernkloof Nature Reserve (Google Maps).

Supplementary Figure S2. Venn diagram of the shared and specific bacterial T-RFLP-derived OTUs found in *E. anceps*, *P. prolifera* and *L. laureolum*.



Supplementary Figure S3. Relative abundance of predominant Deinococcus-Thermus and Alphaproteobacteria OTUs.

		Deinococcus-Thermus			Alphaproteobacteria			
Plant	**Samples**	***Deinococcus***	**Deinococcales**	***Truepera***	**Rhizobiales**	***Sphingomonas***	**Acetobacteraceae**	
		OTU 01	**OTU 04**	**OTU 05**	**OTU 02**	**OTU 09**	**OTU 17**	
E. anceps	**Ea1**	650	0	0	2	1	0	
E. anceps	**Ea2**	507	0	0	134	3	2	
E. anceps	**Ea3**	575	0	0	22	0	3	
E. anceps	**Ea4**	8	0	1	2	0	0	
E. anceps	**Ea5**	266	0	0	7	0	726	
E. anceps	**Ea6**	566	0	0	9	0	1	
P. prolifera	**Pp1**	18	13	8	107	830	0	
P. prolifera	**Pp2**	0	672	329	3	0	0	
P. prolifera	**Pp3**	3	1	0	1056	0	0	
P. prolifera	**Pp4**	202	173	98	269	32	1	
P. prolifera	**Pp5**	128	115	62	286	18	1	
P. prolifera	**Pp6**	8	4	1	247	4	0	
L. Laureolum	**Ll1**	1186	0	0	24	0	0	
L. Laureolum	**Ll2**	1195	0	0	5	1	1	
L. Laureolum	**Ll3**	1220	0	0	0	0	0	
L. Laureolum	**Ll4**	1211	0	0	1	0	1	
L. Laureolum	**Ll5**	1214	0	0	0	0	0	Total no. Sequences 19522
L. Laureolum	**Ll6**	1202	0	0	4	0	1	
		10159	**978**	**499**	**2178**	**889**	**737**	**Number of sequences**
		52.0	**5.0**	**2.6**	**11.2**	**4.6**	**3.8**	**Relative abundance (%)**

Supplementary Figure S4. NMDS ordination plot of the bacterial endophytic communities based on unweighted UniFrac distances among samples. Different plant species showed distinct bacterial communities (PERMANOVA: $F_{1,16}$ = 3.38, P = 0.001). Pairwise comparisons: *E. anceps* vs. *P. prolifera* R^2 = 0.22, *E. anceps* vs. *L. laureolum* R^2 = 0.27, *L. laureolum* vs. *P. prolifera* R^2 = 0.25 (all P < 0.05).