



**University of Dundee**

### **Biogeographical patterns of legume-nodulating *Burkholderia* spp.**

Lemaire, Benny; Chiphango, Samson; Stirton, Charles; Rafudeen, Suhail; Honnay, Olivier; Smets, Erik; Chen, Wen-Ming; Sprent, Janet; James, Euan K.; Muasya, A. Muthama

*Published in:*  
Applied and Environmental Microbiology

*DOI:*  
[10.1128/AEM.00591-16](https://doi.org/10.1128/AEM.00591-16)

*Publication date:*  
2016

*Document Version*  
Peer reviewed version

[Link to publication in Discovery Research Portal](#)

#### *Citation for published version (APA):*

Lemaire, B., Chiphango, S., Stirton, C., Rafudeen, S., Honnay, O., Smets, E., ... Muasya, A. M. (2016). Biogeographical patterns of legume-nodulating *Burkholderia* spp. from African Fynbos to continental scales. *Applied and Environmental Microbiology*, 82(17), 5099-5115. <https://doi.org/10.1128/AEM.00591-16>

#### **General rights**

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

1 **Biogeographical patterns of legume-nodulating *Burkholderia*: from**  
2 **African Fynbos to continental scales**

3  
4  
5  
6 **Benny Lemaire<sup>#1,2</sup>, Samson Chimphango<sup>1</sup>, Charles Stirton<sup>1</sup>, Suhail Rafudeen<sup>3</sup>,**  
7 **Olivier Honnay<sup>2</sup>, Erik Smets<sup>2,4</sup>, Wen-Ming Chen<sup>5</sup>, Janet Sprent<sup>6</sup>, Euan K.**  
8 **James<sup>7</sup>, A. Muthama Muasya<sup>1</sup>**

9  
10  
11 Running title: Biogeography of the *Burkholderia*-legume interaction

12 Keywords: beta-rhizobia, biogeography, *Burkholderia*, host range, legume nodulation

13  
14 <sup>1</sup>*Department of Biological Sciences, University of Cape Town, Private Bag X3, Rondebosch*  
15 *7701, Cape Town, South Africa*

16  
17 <sup>2</sup>*Plant Conservation and Population Biology, KU Leuven, Kasteelpark Arenberg 31, PO Box*  
18 *2435, 3001 Heverlee, Belgium*

19  
20 <sup>3</sup>*Department of Molecular and Cell Biology, University Cape Town, Cape Town, South Africa*

21  
22 <sup>4</sup>*Naturalis Biodiversity Center, Leiden University, 2300 RA Leiden, The Netherlands*

23  
24 <sup>5</sup>*Laboratory of Microbiology, Department of Seafood Science, National Kaohsiung Marine*  
25 *University, No. 142, Hai-Chuan Rd. Nan-Tzu, Kaohsiung City 811, Taiwan*

26  
27 <sup>6</sup>*Division of Plant Sciences, University of Dundee at JHI, Dundee, DD2 5DA UK*

28  
29 <sup>7</sup>*The James Hutton Institute, Invergowrie, Dundee, DD2 5DA, UK*

30  
31 **#Corresponding author**  
32 **Department of Biological Sciences**  
33 **University of Cape Town, Private Bag X3**  
34 **Rondebosch 7701**  
35 **South Africa**  
36 **E-mail address: [benny.lemaire@uct.ac.za](mailto:benny.lemaire@uct.ac.za)**

37  
38 **Plant Conservation and Population Biology**  
39 **KU Leuven, Kasteelpark Arenberg 31, box 02435**  
40 **3001 Heverlee**  
41 **Belgium**  
42 **TEL. + 32 16 32 86 36 FAX + 32 16 32 19 68**  
43 **E-mail address: [benny.lemaire@bio.kuleuven.be](mailto:benny.lemaire@bio.kuleuven.be)**

44

45 **Abstract**

46 Rhizobia of the genus *Burkholderia* have large-scale distribution ranges, and are usually  
47 associated with South African papilionoid and South American mimosoid legumes, yet little  
48 is known about their genetic structuring at either local or global geographical scales. To  
49 understand variation at different spatial scales, from individual legumes in the Fynbos (South  
50 Africa) to a global context, we conducted analyses of chromosomal (16S rRNA, *recA*) and  
51 symbiosis (*nifH*, *nodA*, *nodC*) gene sequences. We showed that the global diversity of  
52 nodulation genes is generally grouped according to the South African papilionoid or South  
53 American mimosoid subfamilies, whereas chromosomal sequence data were unrelated to  
54 biogeography. While nodulation genes are structured on a continental scale, a geographical  
55 or host specific distribution pattern was not detected in the Fynbos region. In host range  
56 experiments, symbiotic promiscuity of *Burkholderia tuberum* STM678<sup>T</sup> and *B. phymatum*  
57 STM815<sup>T</sup> was discovered in selected Fynbos species. Finally, a greenhouse experiment was  
58 undertaken to assess the ability of mimosoid (*Mimosa pudica*) and papilionoid (*Dipogon*  
59 *lignosus*, *Indigofera filifolia*, *Macroptilium atropurpureum* and *Podalyria calyprata*) species  
60 to nodulate in South African (Fynbos) and Malawian (Savanna) soils. While the  
61 *Burkholderia*-philous Fynbos legumes (*D. lignosus*, *I. filifolia* and *P. calyprata*) only  
62 nodulated in their native soils, the invasive neotropical species *M. pudica* did not develop  
63 nodules in the African soils. The Fynbos soil, notably rich in *Burkholderia*, seems to retain  
64 nodulation genes compatible with the local papilionoid legume flora, but is incapable of  
65 nodulating mimosoid legumes which have their center of diversity in the South American  
66 continent.

67

68

69

70 **Importance section**

71 This study is the most comprehensive phylogenetic assessment of root-nodulating  
72 *Burkholderia* and investigates biogeographic and host-related patterns of the legume-  
73 rhizobial symbiosis in the South African Fynbos biome, as well as at global scales, including  
74 native species from the South American Caatinga and Cerrado biomes. While a global  
75 investigation of the rhizobial diversity revealed distinct nodulation and nitrogen fixation  
76 genes among South African and South American legumes, regionally distributed species in  
77 the Cape region were unrelated to geographical and host factors.

78

79 **Introduction**

80 Microorganisms have been observed to vary in distribution, diversity and species  
81 composition across spatial scales (1), challenging the long-held perception of a microbial  
82 cosmopolitanism driven by their high dispersal capacities (2). Although microorganisms can  
83 disperse over lengthy distances, dispersal limitations have revealed spatially isolated  
84 microbial populations over multiple spatial scales (1,3-5). For rhizobia (both alpha- and beta-  
85 subclasses of the Proteobacteria), similar geographical distribution patterns have been  
86 detected in different bacterial groups and over various spatial scales, showing a geographical  
87 structure preserved in phylogenies of both chromosomal and nodulation genes (6-11).  
88 Root-nodulating species of the genus *Burkholderia* (Betaproteobacteria), have been described  
89 from different regions in the world, including parts of the Americas, Africa, Asia and  
90 Australasia. The highest level of diversity has been reported from the South American  
91 Cerrado/Caatinga and South African Fynbos biomes (12), together with Asian and  
92 Australian/New Zealand representatives so far described exclusively from non-native  
93 invasive species, such as the weeds *Mimosa diplotricha*, *M. pigra*, *M. pudica* (13-18) and  
94 *Dipogon lignosus* (49,79). *Burkholderia* species isolated from native legumes from  
95 neotropical and African regions, which are dominated by distinct legume floras (South  
96 American Mimosoideae versus South African Papilionoideae), differ genetically in their  
97 nodulation genes (12,19), suggesting that the legume host is shaping symbiotic diversity and  
98 that the biogeography of rhizobia is linked to the distribution of compatible legume hosts  
99 ((20) and references therein). Despite many local surveys of *Burkholderia* interactions with  
100 papilionoids and mimosoids across the globe, our knowledge of the global distribution  
101 pattern is still fragmented and a spatial survey of the genus *Burkholderia* has never been  
102 conducted in a global context and across biomes.

103 In South Africa, *Burkholderia* symbionts are widespread and associated with diverse lineages  
104 of the tribes Crotalarieae (21-23,26), Hypocalypeteae (24,25), Indigofereae (26), Phaseoleae  
105 (26,27,28) and Podalyrieae (24,26,29,30), indicating that the South African soils are an  
106 important reservoir for nodulating *Burkholderia*, and thus this needs to be explored further  
107 for new candidate species. With the exception of *B. phymatum* strains nodulating the non-  
108 native crop species *Phaseolus vulgaris* (common bean) in Moroccan soils (31), the legume-  
109 *Burkholderia* symbiosis in Africa has only been reported in a range of sites within the  
110 Fynbos region, supporting the idea of the Cape region as an exclusive biodiversity hotspot  
111 for the symbiosis (12).

112 The general aim of the present study is to provide novel insights into the biogeography of  
113 *Burkholderia* and to elucidate the extent to which it exhibits a geographical pattern in relation  
114 to the distribution of its hosts. Because lineages vary in distribution and diversity over  
115 various spatial scales, and spatial factors play a significant role in shaping microbial  
116 communities, it is clear that geographical patterning should be analyzed across multiple  
117 spatial scales (from local to broad geographical regions). We took advantage of the large  
118 record of root-nodulating *Burkholderia* established since the first reports of its nodulation  
119 ability ((12) and references therein), supplemented with new sequence data of Fynbos  
120 *Burkholderia*. Available sequence data for chromosomal 16S rRNA and the symbiosis-  
121 related *nodA*, *nodC* and *nifH* genes were analyzed in a world-wide perspective to assess  
122 geographic patterns at a continental scale, as well as the host specific interactions with the  
123 legume subfamilies Mimosoideae and Papilionoideae.

124 The diversity, geographic distribution and host associations were further investigated at a  
125 regional scale in the South African (Cape) Fynbos biome. The *Burkholderia* symbionts from  
126 five Cape legume tribes and 11 genera of the Papilionoideae were investigated by

127 phylogenetic analyses of two chromosomal genes (16S rRNA and *recA*) and one nodulation  
128 gene (*nodA*) in relation to their geography and host phylogeny.

129 We hypothesize that the *Burkholderia* symbionts of native and invasive legume species  
130 reported from Africa, America, Asia and Australasia exhibit a geographical distribution  
131 pattern with continents having their own subset of symbionts. We also expect a geographical  
132 effect on the genetic variation of rhizobia at a regional scale within the Fynbos. Our specific  
133 objectives were (1) to determine and compare the *Burkholderia* types for housekeeping and  
134 symbiosis loci recorded from mimosoids and papilionoids reported from four different  
135 continents; (2) to investigate the distribution pattern of *Burkholderia* and its host-associations  
136 within the Fynbos biome, using field-collected nodules of indigenous papilionoids; (3) to  
137 investigate the ability of South African papilionoid legume species (*Indigofera filifolia*,  
138 *Dipogon lignosus*, *Podalyria calytrata*, *Psoralea pinnata*) and the South American species  
139 *Mimosa pudica* (subfamily Mimosoideae) to form nodules in South African (Fynbos) and  
140 Malawian (Savanna) soil; and (4) to test and evaluate the host range of the *Burkholderia*  
141 *tubercum* STM678<sup>T</sup> and *B. phymatum* STM815<sup>T</sup> type strains on selected Fynbos species,  
142 which are known to exhibit different host affinities as dictated by their genetically distinct  
143 nodulation genes. We expect that the tested papilionoid legumes from the Fynbos are  
144 exclusively nodulated by the common and native symbiont *B. tubercum* STM678<sup>T</sup>.

145

146 **Material and Methods**

147 *Burkholderia* datasets and OTU-based analyses

148 Analyses of Operational Taxonomic Units (OTUs) were used to cluster the 16S rRNA  
149 sequence data. A large 16S rRNA data set was constructed, comprising 1121 sequences and  
150 75 validly named *Burkholderia* species with multiple accessions per species. Sequences  
151 were aligned with available bacterial reference sequences via the Ribosomal Database  
152 Project (RDP pyrosequencing pipeline; <http://pyro.cme.msu.edu>). An uncorrected pairwise  
153 distance matrix was calculated and the number of OTUs and rarefaction curves at various  
154 cut-off values (0.030 to 0.010) were calculated in Mothur v.1.31.2 (33).

155 Four other *Burkholderia* datasets were obtained from available 16S rRNA (365 sequences),  
156 *nifH* (246 sequences), *nodA* (152 sequences) and *nodC* (199 sequences), assigning all  
157 rhizobia to four geographical regions (Africa, America, Asia and Australasia and two  
158 legume subfamilies (Papilionoideae and Mimosoideae). The alignments were created with  
159 Muscle (32) using Geneious v.5.1.7 (<http://www.geneious.com>). The diversity of 16S rRNA  
160 sequences was clustered into OTUs, using the previous estimated cut-off value to delineate  
161 taxonomic identities at species level. For the data sets of *nifH*, *nodA* and *nodC*, we applied a  
162 similar conservative similarity cut-off value in order to classify genetic groups of the more  
163 variable symbiosis genes. Unique and shared types among different continents and  
164 subfamilies were identified in Mothur.

165 Alignments for the NeighborNet analyses were compiled based on the previous 16S rRNA,  
166 *nifH*, *nodA* and *nodC* rhizobial datasets: one sequence representative per sequence cluster  
167 was manually selected from the original alignments and these were imported into SplitsTree  
168 v.4.12.8 (34) to display the phylogenetic relatedness among the clusters as a NeighborNet  
169 network (35), using the most complex model of nucleotide substitution (GTR) available.  
170 Bootstrap confidence values were generated using 1,000 permutations.



171 *Nodule sampling, DNA extraction, amplification, cloning and sequencing to identify*

172 *Fynbos rhizobia*

173 We investigated 20 root nodulated Fynbos species in this study, representing various  
174 localities (Fig. 1) and diverse host legumes (five legume tribes and 11 genera). Voucher  
175 information and GenBank accession numbers are listed in Table S2 and the geographical  
176 localities are shown in Figure 1. Nodules were collected in the field from a broad  
177 geographical range at different localities, covering diverse soil types ranging from limestone  
178 substrate (De Hoop Nature Reserve, Still Bay), granite substrate (Paarl Mountain Nature  
179 Reserve), sandstone mountain slopes (Bainskloof Nature Reserve) to coastal deep sand (Cape  
180 Point Nature Reserve). Three to five nodules were removed from each host plant for isolation  
181 of rhizobia.

182 Rhizobia were identified by both standard culturing techniques (36) and direct genomic DNA  
183 extraction from nodules. The latter method enabled the unequivocal assessment of the  
184 intranodular endophyte diversity, including unculturable endophytes that can be masked  
185 using culturing-based techniques due to the selective effects of growth media and an  
186 incomplete sampling of colony morphotypes. For the standard culturing technique, rhizobia  
187 were isolated on yeast extract mannitol agar (YEMA) from a single bacterial colony type,  
188 following standard procedures (36). Pure rhizobial cultures from single colonies were stored  
189 at -80°C in YEM broth containing 20% glycerol. Total DNA of the rhizobial cultures was  
190 obtained by the following thermal cell lysis procedure: A loopful of bacterial culture was  
191 suspended in 20 µl lysis buffer (10% SDS, 1M NaOH) followed by incubation for 15 minutes  
192 at 95°C. The lysate was centrifuged at 10,000 g for 45 s and 180 µl of sterile water was  
193 added. The DNA extract was centrifuged for another 5 minutes at 10,000 g at 4°C and  
194 preserved at -20°C. For the direct DNA extraction from root nodules, genomic DNA of

195 surface-sterilized nodules was obtained using the E.Z.N.A.<sup>TM</sup> HP Plant DNA Mini Kit  
196 (Omega bio-tek) as per manufacturer's instructions.

197 PCR amplification of 16S rRNA used universal bacterial primers (27f and 1492r) as  
198 previously described (37). Amplicons of nearly complete 16S rRNA were sequenced and  
199 subjected to BLAST analyses on GenBank as a first identification tool. Amplification of the  
200 *recA* housekeeping gene and the *nodA* nodulation gene was carried out with the primers  
201 *recA*-63F, *recA*-504R, *nodA*-1F, *nodA*-2R and PCR parameters as described by Gaunt *et al.*  
202 (38) and Haukka *et al.* (39). Amplification of the *nodC* nodulation gene was carried out for  
203 selected Fynbos isolates, using the primers *nodC*-540 and *nodC*-1160. All primer sequences  
204 are listed in Table S3.

205 Amplified 16S rRNA products from total genomic DNA extractions of the nodules were  
206 cloned into a pGEM-T vector (Promega), according to the manufacturer's instructions, and  
207 transformed into JM109 *E. coli* by heat shock (40). Purified plasmids and all PCR products  
208 were sent to Macrogen for sequencing (Macrogen Inc, Seoul, Korea). Sequencing primers for  
209 16S rRNA, *recA* and *nodA* were the same as for the initial PCR.

210

#### 211 *Authentication of cultured rhizobia from field nodules*

212 Nodulation capabilities of isolates from field nodules were tested on siratro (*Macroptilium*  
213 *atropurpureum*) (36). Table S2 lists the authenticated isolates in this study together with  
214 previously tested strains (26). Rhizobial isolates from nodules of legume species (*Dipogon*  
215 *lignosus*, *Indigofera filifolia*, *Podalyria calyptata* and *Psoralea pinnata*) grown in the  
216 greenhouse were authenticated on their respective host. Nodulation (three replicates) was  
217 assessed by either inoculating seedlings with a rhizobial culture (OD<sub>600</sub>) or leaving them  
218 uninoculated as negative controls. Authentication was confirmed if isolates nodulated the

219 roots of inoculated plants from all replicated pots, and the uninoculated plants remained  
220 nodule-free.

221

#### 222 *Phylogenetic analyses of the 16S rRNA, recA and nodA sequence data*

223 Sequence reads were assembled and sequence alignments were created with Muscle (32)  
224 using Geneious v.5.1.7 (<http://www.geneious.com>). For the combined phylogeny of 16S  
225 rRNA and *recA*, missing sequences due to the lack of amplification were treated as missing  
226 data. Phylogenetic relationships were conducted using Bayesian Inference (BI) and  
227 Maximum Likelihood (ML) optimality criteria. Bayesian analyses were carried out in  
228 MrBayes v.3.1 (41) after determining the appropriate model of evolution with MrModeltest  
229 v.3.06 (42) under the Akaike information criterion. Modeltest selected for the 16S rRNA,  
230 *recA* and *nodA* datasets the GTR+I+G model. Four Markov Chains were run simultaneously  
231 for four million generations, sampling every 100 generations. The initial 25% of trees were  
232 regarded as “burnin” and discarded. Convergence of the chains was checked using Tracer  
233 v.1.4 (43). ML analyses were performed using RAxML-VI-HPC v.7.0.4 (44). A total of 100  
234 RAxML searches were conducted, relying on the GTR-GAMMA model of evolution.  
235 Support values were estimated using a multi-parametric bootstrap resampling with 1000  
236 pseudo-replicates.

237

#### 238 *Geographic distances among Fynbos representatives*

239 The genetic variation of rhizobia across spatial scales in the Fynbos was calculated on both  
240 chromosomal (*recA*) and nodulation (*nodA*) data. Genetic distance matrices for both sets of  
241 genes was constructed including our Fynbos isolates and supplemented with previously  
242 described rhizobial strains (see Fig. 1). The *recA* and *nodA* datasets comprised 134 and 128  
243 sequences, respectively, covering genera of the tribes Podalyrieae (*Amphithalea*, *Cyclopia*,

244 *Podalyria*, *Virgilia*), Crotalariaeae (*Aspalathus*, *Crotalaria*, *Rafnia*), Hypocalypsteae  
245 (*Hypocalyptus*), Phaseoleae (*Bolusafra*, *Dipogon*) and Indigofereae (*Indigofera*). Genetic  
246 variation of all pairs of isolates was linked with a geographic distance matrix calculated from  
247 their geographic coordinates using the Geographic Distance Matrix Generator, v.1.2.3 (Ersts,  
248 American Museum of Natural History, Center for Biodiversity and Conservation). Values of  
249 genetic variations were grouped within geographic distance classes (0-200, 201-400, 401-  
250 600, 601-800 km) and plotted as box plots in R v.2.15.3 (45). The correlation between  
251 genetic similarities and geographic distances was investigated using a Mantel test in  
252 Genealex 6.501 (46), and its significance was tested on 9.999 permutations.

253

254 *Trapping experiment*

255 The legume species *Dipogon lignosus* (L.) Verdc. (Phaseoleae), *Indigofera filifolia* Thunb.  
256 (tribe Indigofereae), *Macroptilium atropurpureum* (DC.) Urb. (siratro; tribe Phaseoleae),  
257 *Mimosa pudica* L. (tribe Mimoseae), *Podalyria calyptrata* C.A. Sm. (Podalyrieae) and  
258 *Psoralea pinnata* L. (Psoraleeae) were grown in soil samples from Malawi (Chinyonga,  
259 Blantyre - S15.819431, E35.041753) and South Africa (Table Mountain National Park -  
260 S33.952532, E18.456871). Both sites are part of natural ecosystems with no history of  
261 cultivation or rhizobial inoculation. At each locality, soils were sampled from the top 0-20  
262 cm from at least three field sites and bulked to generate a composite sample for rhizobial  
263 isolation. Soil pH was determined from 4 g samples of sieved (1 mm mesh) mixed in 40 ml  
264 1M KCl.

265 The Cape legume species *I. filifolia* and *P. calyptrata* are endemic to the Western Cape  
266 Province of South Africa. *Psoralea pinnata*, also endemic to the Fynbos, became naturalized  
267 and invasive in South Australia and New Zealand (47). *Dipogon lignosus* and *M. pudica*,  
268 which are native to South Africa and South America respectively, are also considered as

269 invasive (48,49). All legume species of the trapping experiment, except for *Ps. pinnata*,  
270 which is strictly associated with *Mesorhizobium* (alpha-class of Proteobacteria) (26), have  
271 been shown to form associations with *Burkholderia*. (26,49,50). *Siratro* is a widely used  
272 species known to be very promiscuous with regard to symbionts (51) and was proven  
273 previously to be effectively nodulated by *Burkholderia tuberum* (29,52).

274 Nodulation was assessed by growing germinated seedlings (three replicates) in 20 cm  
275 diameter plastic pots filled with acid-washed sterile sand and a layer of 200 g of soil (the  
276 layer of soil was omitted from negative control pots). Seeds were surface-sterilized in 4%  
277 (w/v) sodium hypochlorite for 10 min., rinsed in six changes of sterile water, soaked in  
278 boiled water and pre-germinated at room temperature on 1.5% (w/v) agar plates until root  
279 emergence. Pots were covered with a layer of nylon PA6 beads (Lomold group HQ, South  
280 Africa) and provided with a sterile watering tube to prevent cross-contamination. All plants  
281 were watered with sterile de-ionized water every two days. Nodules were harvested from  
282 seedlings after two months and rhizobia were isolated on YEMA as previously described.

283

#### 284 *Host range study*

285 Seeds of legume species from the tribes Crotalariaeae, Hypocalyptaeae, Indigofereae and  
286 Podalyrieae were used for this study. Seeds were surface-sterilized with concentrated  
287 sulphuric acid for 10 min. followed by 4% sodium hypochlorite for 10 min. Seedlings were  
288 grown in glass tubes with a sterile mixture of Vermiculite/Perlite as a rooting medium and fed  
289 with Jensens N-free plant nutrient medium under aseptic conditions (53). After one week of  
290 plant cultivation, seedlings were inoculated with the wild type strains *B. tuberum* STM678<sup>T</sup>  
291 and *B. phymatum* STM815<sup>T</sup> (54). Plants were harvested after 6 weeks and inspected for  
292 nodule formation and the potential ability to perform symbiotic nitrogen fixation was assessed  
293 by the presence of leghemoglobin (Lb). In addition, nodules were fixed and embedded for

294 light microscopy to assess their internal structure, as this is also a strong indicator of  
295 effectiveness (26,55). Three species of *Podalyria* and one *Virgilia* species, *V. oroboides* (tribe  
296 Podalyrieae), were also inoculated with a GFP-marked *B. tuberum* STM678 variant strain  
297 (29); nodule preparation and morphological observation of the STM678-GFP construct in  
298 nodule sections, using light and fluorescence microscopy, are according to (29). Uninoculated  
299 plants served as controls.

300

301 *Nucleotide sequence accession numbers*

302 The 16S rRNA sequences were deposited in the GenBank database under the accession  
303 numbers KF791602-KF791673 and KF824727-KF824733. The *recA* sequences were  
304 deposited under accession numbers KF791796-KF791864, KF824747-KF824753, KP013139-  
305 PK013158 and KT700208-KT700213. Sequences for the *nodA* sequences were deposited  
306 under the accession numbers KF791743-KF791795, KF824740-KF824746, KP013159-  
307 KP013178 and KT700202-KT700207. Sequences for the *nodC* sequences were deposited  
308 under the accession numbers KP013126-KP013137.

309

310 **Results**311 *16S rRNA gene sequence cut-off levels used for (putative new) species delineation*

312 A large 16S rRNA dataset comprising 75 validly named *Burkholderia* species was  
313 constructed to evaluate the genetic diversity of *Burkholderia* at five different sequence  
314 similarity threshold values ranging from 97% to 99% (Fig. S1). A sequence similarity level  
315 to delineate the true number of sequences at species level was obtained between a cut-off  
316 value of 98.5% and 99%, resulting in 59 and 96 OTUs. Although there is some controversy  
317 about the concept of a species in prokaryotes (56-59) the results of the empirical clustering  
318 analysis, using 16S rRNA data, support 98.5% as a conservative threshold value for species  
319 level definitions within *Burkholderia* and corresponds to the general threshold value of  
320 98.65% estimated to delineate the global prokaryotic diversity (60). A 98.5% threshold value  
321 was used for further diversity calculations of 16S rRNA datasets.

322

323 *Phylogenetic clustering of the Burkholderia richness according to geography and legume*  
324 *subfamily*

325 The diversity of root-nodulating *Burkholderia* was classified according to geography and  
326 their hosts for different DNA regions (16S rRNA, *nifH*, *nodA* and *nodC*). Table 1 shows the  
327 16S rRNA OTUs and clusters of symbiosis genes calculated at a cut-off value of 98.5%,  
328 which are identified from different continents and host associations occurring across  
329 continents and legume subfamilies. From a total of 23 16S rRNA OTUs, eight groups  
330 occurred on more than one continent, including one OTU (number 5) globally distributed  
331 across all four continents assessed and three OTUs (numbers 1, 5 and 12) associated with  
332 both legume subfamilies (Table 1). *Burkholderia tuberum* (OTU number 1) was a highly  
333 recorded species (107 16S rRNA sequences) associated with eight South African genera  
334 (*Amphithalea*, *Aspalathus*, *Cyclopia*, *Hypocalyptus*, *Lebeckia*, *Podalyria*, *Rhynchosia* and

335 *Virgilia*) and from field nodules of the South American genus *Mimosa* (Table 1). Six OTUs  
336 (numbers 4-8 and 15) comprised symbionts of invasive *Mimosa* species, recorded from South  
337 America, as well as their invasive regions in Asia and Australia (Table S1).

338 In contrast to 16S rRNA, fewer nodulation and nitrogen fixation types were shared among  
339 continents, including only four *nifH* (numbers 2, 4, 5 and 9), five *nodA* (numbers 3, 5, 6, 10  
340 and 16) and three *nodC* (numbers 4, 9 and 10) types. One group of *nodC* (type number 4) and  
341 one group of *nifH* (type number 4) sequences were globally distributed on all the four  
342 continents. A total of five sequence clusters were shared between both subfamilies for *nifH*  
343 (numbers 1 and 4), *nodA* (numbers 3 and 14) and *nodC* (number 4). All *nodA* and *nodC*  
344 sequence clusters associated with both legume subfamilies originate from mimosoids and  
345 from the papilionoid hosts *Macropodium* and *Phaseolus*.

346 Phylogenetic NeighborNet analyses for chromosomal (16S rRNA), nitrogen fixation (*nifH*),  
347 and nodulation (*nodA* and *nodC*) genes revealed the genetic divergence and clustering among  
348 sequence types and their affinities for a geographical locality and legume subfamily (Fig. 2).  
349 The genetic distances, proportional to evolutionary divergences, were more pronounced for  
350 the symbiosis genes (*nifH*, *nodA* and *nodC*) than for the conservative 16S rRNA gene. For  
351 16S rRNA, phylogenetic relationships among OTUs were not structured by geography nor  
352 host (Fig. 2A). Large genetic clusters contained OTUs from different continents and  
353 subfamilies, confirming the previous observation of shared 16S rRNA types across localities  
354 and hosts (Table 1). In contrast to 16S rRNA, NeighborNet analyses of nitrogen fixation  
355 (*nifH*) (Fig. 2B) and nodulation (*nodA* and *nodC*) genes (Fig. 2C-D) identified a strong  
356 pattern according to geography and host. Genetic clusters were identified, separating the  
357 African papilionoids from the South American and Asian mimosoid representatives.

358



359 *Burkholderia diversity, specificity and geographical distribution in legumes of the Fynbos*  
360 *biome*

361 In order to investigate the biodiversity and geographic distribution of *Burkholderia* at a  
362 smaller spatial scale, rhizobia of diverse indigenous Fynbos species were sampled and  
363 analyzed using a combination of culture and non-culture based identification techniques.  
364 Initially, a standard culture method was applied to selected legume lineages covering most  
365 legume groups (Table S2; *Amphithalea*, *Aspalathus*, *Bolusafra*, *Crotalaria*, *Dipogon*,  
366 *Hypocalyptus*, *Indigofera*, *Podalyria*, *Rafnia*, *Rhynchosia* and *Virgilia*) and all rhizobia were  
367 identified as *Burkholderia*, showing only a single colony morphotype in each root nodule.  
368 All cultured strains were authenticated using siratro (Table S2), showing effective nodules  
369 and enhanced plant growth compared with nodule-free controls. Only the strain from *Rafnia*  
370 *acuminata* (Dlodlo 22) failed to form effective nodules on siratro and so was not regarded as  
371 a rhizobial symbiont.

372 In addition, a culture independent approach was performed using direct PCR analyses to  
373 assess the nodule rhizobial diversity and to confirm single strain occupation within a nodule.  
374 PCR amplifications on the total genomic DNA extraction of the intranodular tissue produced  
375 high quality and single-copy sequences for all genetic markers investigated, suggesting one  
376 dominant *Burkholderia* strain as nodule resident. Amplified 16S rRNA products were cloned  
377 for available nodules in selected species within genera of two legume tribes (*Podalyria*:  
378 Muasya, 6490; 6463 and *Indigofera*: Muasya & Stirton, 6502B; 6502C) to test the one-  
379 symbiont one-nodule specificity. For all samples investigated, similar 16S rRNA clones (20  
380 per sample) were obtained showing a single bacterial endosymbiont in each nodule.

381 Sequence data of 16S rRNA, *recA* and *nodA* from rhizobia of 26 Podalyriaceae (13 individuals,  
382 3 genera), 11 Indigoferaceae (8 individuals, 1 genus), 4 Hypocalypteae (2 individuals, 1 genus),  
383 15 Crotalarieae (6 individuals, 3 genera) and 16 Phaseoleae (9 individuals, 3 genera) were

384 analyzed with Maximum Likelihood and Bayesian phylogenetic analyses (Figs. 3-4),  
385 clustering the isolates within diverse reference strains, comprising root-nodulating (*B.*  
386 *dilworthii* WSM3556<sup>T</sup>, *B. dipogonis* LMG19430<sup>T</sup>, *B. kirstenboschensis* Kb15<sup>T</sup>, *B.*  
387 *rhynchosiae* WSM3937<sup>T</sup>, *B. sprentiae* WSM5005<sup>T</sup>, *B. tuberum* STM678<sup>T</sup>) and plant-  
388 beneficial (*B. phytofirmans* PsJN<sup>T</sup>, *B. xenovorans* LB400<sup>T</sup>) lineages. Our isolates from  
389 various host legumes (e.g. *Amphithalea*, *Aspalathus*, *Indigofera*, *Rafnia*, *Rhynchosia*,  
390 *Podalyria*) were closely related to nodulated representatives (*B. kirstenboschensis*, *B.*  
391 *rhynchosiae*, *B. tuberum*) of the current Fynbos record, but the majority of isolates appeared  
392 to be related to bacteria without generally nodulating traits (*B. phytofirmans*, *B. xenovorans*)  
393 or were grouped apart into clusters without known reference species (Fig. 3).

394 Analyses of rhizobial lineages in relation to their geographical provenance showed many  
395 widely distributed 16S rRNA OTU types, suggesting genetic similarity of *Burkholderia* in  
396 Fynbos soils. To evaluate the diversity of Fynbos rhizobia in relation to geography at a  
397 regional scale, we investigated spatial structuring by the common approach of isolation by  
398 distance (61), assuming that geographic distance and population genetic differentiation are  
399 expected to correlate positively because population connectivity occurs more frequently  
400 among adjacent habitats. For close and distantly located populations, genetic variation was  
401 examined among *Burkholderia* strains, showing no effect of geographical distance on the  
402 genetic distance for both *recA* and *nodA* sequence data (Fig. 5). Genetic differentiation was  
403 constant among the different distance classes (0-200; 201-400; 401-600; 601-800 km),  
404 showing mean values of genetic similarities of ca. 94% and 96% for *recA* and *nodA*,  
405 respectively (Fig. 5). A Mantel test examined the associations between pairwise differences  
406 in genetic and geographical distances, rejecting an effect of geographical distance on the  
407 genetic *Burkholderia* variation ( $P > 0.05$ ).

408 Similar to geography, no link was observed between *Burkholderia* strains and host genotype.  
409 For the majority of hosts, different populations of one legume species were associated with a  
410 set of genetically diverse strains of *Burkholderia* for both chromosomal and nodulation data  
411 (Figs. 3-4). Sequence analyses showed that a given *Burkholderia* lineage was associated with  
412 different legume lineages and in turn these host plants accommodated genetically diverse  
413 symbionts.

#### 414 *Nodulation of Cape legumes in African soils and identity of rhizobial groups*

415 Nodulation of the legumes *I. filifolia*, *P. calytrata* and *Ps. pinnata*, which are restricted in  
416 distribution to the Cape Fynbos biome, and the widely distributed species *D. lignosus* and *M.*  
417 *pudica* was assessed in South African (Fynbos region) and Malawian (Savanna grassland)  
418 soils. The pH of the soil from the Fynbos ( $\text{pH} = 4.6 \pm 0.2$ ) was substantially lower than at the  
419 Savanna site ( $\text{pH} = 7.1 \pm 0.3$ ). Distinct symbiotic associations were found among the  
420 legumes with a strong influence of the source of soils on the rhizobia sampled (Figs. 6-7).  
421 *Podalyria calytrata* (Podalyrieae), *I. filifolia* (Indigofereae) and *D. lignosus* (Phaseoleae)  
422 were exclusively nodulated by *Burkholderia* in Fynbos soil, with the exception of one  
423 *Bradyrhizobium* isolate associated with *D. lignosus* that was from Fynbos soil. None of these  
424 legume species nodulated in Malawian soil, except *Ps. pinnata* (Psoraleeae) and siratro  
425 (Phaseoleae) that were able to form nodules in both soils (Table S4), with isolates identified  
426 as *Mesorhizobium* (*Ps. pinnata* – Fynbos), *Burkholderia* (siratro – Fynbos) and  
427 *Bradyrhizobium* (*Ps. pinnata*, siratro – Malawi). *Mimosa pudica* formed no nodules in either  
428 the South African or the Malawian soils.

429 The *Burkholderia* and *Mesorhizobium* symbionts isolated from legumes growing in Fynbos  
430 soils were placed in different clades (Figs. 6-7) and were highly related (99-100% sequence  
431 similarity) to known reference strains previously isolated from various South African  
432 legumes (Table S4). The *recA* and *nodA* sequence data of bradyrhizobia symbionts from the

433 Malawian soils were related (97-99%) to known African, South American and European  
434 isolates (Table S4).

435

436 *Host range of Burkholderia tuberum and B. phymatum among South African legumes*

437 The host range experiment showed that all legumes from the tribes Crotalariaeae,  
438 Hypocalyptaeae, Indigofereae and Podalyrieae were able to nodulate successfully with the  
439 type strain of *B. tuberum*, STM678<sup>T</sup>, except for four *Calpurnia* species, which either did not  
440 produce nodules (*C. aurea* and *C. intrusa*) or showed ineffective nodulation (*C. glabrata*  
441 and *C. sericea*) (Table 2; Fig. S2). All legume species assessed in the host range experiment  
442 are native Fynbos species, except for *Calpurnia*, where only *C. intrusa* is found in the  
443 karroid vegetation near the Fynbos-dominated Swartberg Mountains. The presence of *B.*  
444 *tuberum* in the nodule structure was confirmed in *Podalyria* and *Virgilia* species by  
445 fluorescence microscopy of the GFP transconjugant strain of STM678 (Fig. 8), and in all the  
446 other species by immunogold labelling with a *Burkholderia*-specific antibody (Fig. S2). The  
447 type strain of *Burkholderia phymatum*, STM815<sup>T</sup>, formed functional nodules on four native  
448 Fynbos legume species of the tribe Podalyrieae (*Cyclopia* and *Virgilia*), whereas other  
449 species of the genera *Amphithalea* (tribe Podalyrieae), *Hypocalyptus* (tribe Hypocalyptaeae),  
450 *Aspalathus* and *Lebeckia* (both tribe Crotalariaeae) produced ineffective nodules or remained  
451 nodule-free (Table 2; Fig. S2).

452

453 **Discussion**454 *Spatial distribution of root nodulated Burkholderia at continental scale*

455 The global survey of the *Burkholderia* record revealed various geographical and host-related  
456 patterns within the 16S rRNA and *nifH*, *nodA* and *nodC* datasets at a continental scale.  
457 Chromosomal 16S rRNA types were highly diverse (Fig. 2A, Table 1) and unrelated to the  
458 host subfamily or geographical region, whereas nitrogen fixation and nodulation genes are  
459 clearly structured by a geographical and host factor (Fig. 2B-D) with only a few sequence  
460 groups identified across continents and legume subfamilies (Table 1). The observation of an  
461 association between geography, host legume and nodulation genes, showing two large  
462 clusters of highly diverged nodulation gene types, according to their geographical origin and  
463 host subfamily, corroborates previous *Burkholderia* studies (12,19). All African distributed  
464 rhizobia were clustered in one group, and were highly diverged (<75% gene similarity) from  
465 the remaining mimosoid-related *Burkholderia*.

466 The geographical distribution of the legume host seems to be the key factor, explaining the  
467 nodulation and nitrogen fixation gene phylogenetic structure at a continental scale,  
468 supporting the idea that the rhizobial biogeography largely follows their hosts (20), which  
469 represent two distinct legume floras of South African papilionoids and South American  
470 mimosoids in the Fynbos and Cerrado/Caatinga biomes, respectively (12,62). Evidence is  
471 accumulating that the vast majority of *Mimosa* species native to central Brazil are exclusively  
472 associated with *Burkholderia* (10,55), whereas in Mexico, which is considered as another  
473 large centre of radiation of the genus, most endemic species are not nodulated by beta-  
474 rhizobia (17), but are specifically associated with alpha-proteobacteria and only a few  
475 lineages are able to form interactions with *Burkholderia* (11,63). Distinct nodule occupancies  
476 of beta- and alpha-rhizobia within the native home range of Brazilian and Mexican *Mimosa*  
477 species, respectively, can be largely explained by a combination of geographical separation

478 of the various *Mimosa* clades with distinct symbiont preferences, and their subsequent co-  
479 evolution with rhizobia in contrasting soil types (*e.g.* acid versus neutral/alkaline soils) (11).  
480 Conversely, access and availability of rhizobia, due to varied adaptation to edaphic and  
481 climatic factors, may be a critical factor governing dispersal of legumes outside native areas  
482 and thereby influence legume biogeographic patterns. The latter may be true for South Africa  
483 and Western Australia, which have frequent angiosperm dispersal events in the Cenozoic  
484 (64), associated with similarity of niches (Mediterranean climate, oligotrophic acidic soils),  
485 yet legumes are one of the few (large) families that do not exhibit disjunction between the  
486 two continents. While the endemic Australian tribes Bossiaceae and Mirbelieae are largely  
487 associated with *Bradyrhizobium* lineages (65,66), the tribe Hypocalypteae, which is endemic  
488 to South Africa and resolved as a sister group to the mirbelioids is strictly associated with  
489 *Burkholderia*.

490 The nodulation genes *nodA* and *nodC* are frequently used to understand the symbiotic  
491 specificities and their evolutionary adaptation to a specific host (67). Because nodulation  
492 genes are involved in the synthesis of Nod-factors (*i.e.* rhizobial signaling molecules required  
493 for the earliest host responses) they determine the host specificity (68-70) and have been  
494 frequently shown to align with their *Burkholderia* host (12,17,28). The specificity of the  
495 symbiotic association of *Burkholderia* with mimosoid and papilionoid legumes is clearly  
496 demonstrated in one single species, *B. tuberum*, which has distinct nodulation genes or  
497 symbiotic variants and has been ascribed to symbiovars mimosae and papilionoideae,  
498 respectively (71,72). However, a link between *nodA* types and the legume subfamily is not  
499 strictly predictable for all species. *Macroptilium atropurpureum* (siratro, Papilionoideae) for  
500 example, known as a valuable plant for trapping a broad range of alpha- and beta-rhizobia  
501 (52), is able to nodulate with both *B. tuberum* sv. papilionoideae (*e.g.* STM678<sup>T</sup>) (29) and sv.  
502 mimosae strains (*e.g.* STM4801) (71). Similarly, the mimosoid symbiont *B. phymatum*

503 STM815<sup>T</sup> has been isolated from nodules of the papilionoid *P. vulgaris*, which is known for  
504 its wide range of symbiotic partners (31). Apart from the records involving promiscuous host  
505 legumes (siratro, *P. vulgaris*), *Burkholderia* species and their nodulation genes appear to  
506 group and evolve in close concert with their mimosoid and papilionoid hosts. However,  
507 evidence is accumulating that, although rhizobial species (e.g. *B. tuberum* sv. papilionoideae)  
508 associated with the subfamily of Papilionoideae appear incapable of nodulating mimosoid  
509 hosts (29), the opposite is not the case (12). In addition to common bean (73), diverse  
510 papilionoids such as the Fynbos species *Dipogon lignosus* (49) and legumes of the genera  
511 *Cyclopia* and *Virgilia* (Table 2, Fig. S2) have been demonstrated to form effective nodules  
512 with the mimosoid-nodulating *B. phymatum*-type symbiont (17, 74), confirming its broad  
513 host range and ability to associate with legumes of the mimosoid and papilionoid subfamily.  
514 While symbiosis genes are largely structured according to legume subfamily, 16S rRNA  
515 clusters are more diverse (Fig. 2A), affiliated with various hosts from different parts of the  
516 world (Table 1). A widespread occurrence of *Burkholderia* strains, especially for 16S rRNA  
517 types (Table 1), indicates an inter-continental and global distribution pattern for different  
518 strains of burkholderias (e.g. *B. diazotrophica*, *B. mimosarum*, *B. phymatum*, *B. sabiae* and  
519 *B. tuberum*). The occurrence and vast diversity of *Burkholderia* outside Africa and South  
520 America are mostly recorded from pan(sub)tropically distributed *Mimosa* species (*M.*  
521 *pudica*, *M. pigra*, *M. diplotricha*). *Burkholderia* symbionts of these widespread invasive  
522 plant species are included in the clustering analyses and close relationships of nodulation  
523 genes with their native distributed relatives support previous observations that rhizobia are  
524 co-transported with the seeds or plants from their native to new invasive habitats. Following  
525 the co-introduction hypothesis (75), symbionts that have been co-introduced with their hosts  
526 or which have hitchhiked on introduced material over long-distances, bridging geographical  
527 barriers between continents, has been evidenced in many studies (15,49,76-78). For



528 *Burkholderia*, a plausible long-distance migration event from South Africa to New Zealand,  
529 possibly dispersed across the Australian continent, has been reported in the South African  
530 papilionoid *Dipogon lignosus* (tribe Phaseoleae) (79), which is invasive in New Zealand and  
531 Australia (49, 79) as revealed by high sequence similarities of the symbiosis genes (*nodA*  
532 sequence clusters 6, 10; *nodC* sequence clusters 9, 17) between invasive populations of  
533 *Dipogon* and native South African relatives from the genera *Bolusafrsa*, *Crotalaria*,  
534 *Cyclopia*, *Hypocalyptus*, *Indigofera*, *Podalyria* and *Rhynchosia*.

535

#### 536 *Geographical distribution and specificity of Fynbos Burkholderia*

537 While the global *Burkholderia* diversity was structured for the nodulation genes at legume  
538 subfamily level, an interaction between rhizobia, host legumes and geographical distribution  
539 was not shown at regional scale, showing widely spread and locally diverse *Burkholderia*  
540 populations in the Fynbos. Our results corroborate a previous study, demonstrating the  
541 widespread occurrence of *Burkholderia* and the absence of a site sampling effect on the  
542 rhizobial diversity of selected Hypocalypteae and Podalyrieae species (24,30). Using  
543 geographical distances as a proxy for population connectivity, genetic variation is expected  
544 to correlate positively with the sampling site distances. Our study does not show any  
545 correlation between genetic variation and geographical distance, suggesting the absence of  
546 genetic isolation through high rates of rhizobial dispersal of both chromosomal and  
547 symbiosis traits.

548 In the Fynbos region, local environmental variables, rather than spatial dispersal factors, are  
549 most likely the major ecological drivers for rhizobial distributions. In a recent study, Lemaire  
550 and associates (26) showed that genetic variation of Fynbos *Burkholderia* was correlated  
551 with differences in site elevation, a feature also observed in symbionts of South American



552 *Mimosa* species (10); hence the indirect effects of temperature and rainfall may play a  
553 significant role in the rhizobial community structure.

554 Symbiotic associations of Fynbos legumes for *Burkholderia* have been described in many  
555 lineages with various degrees of specificity. In the tribe Podalyriaceae, a strong preference for  
556 *Burkholderia* is observed, showing all legume species and genera (except for *Calpurnia*  
557 which is not endemic to the Fynbos – Table 2, Figure S2) strictly nodulated with  
558 *Burkholderia* (12,24,26). Other common plant groups such as the tribes Crotalariaeae and  
559 Indigoferaceae also contain *Burkholderia*-philous species, although (closely related) legume  
560 lineages within the same tribes and co-occurring in the similar habitats have been recorded  
561 with classical alpha-rhizobial lineages ((26) and references therein).

562 In this study, the *Burkholderia*-legume interaction was further investigated at a finer  
563 taxonomic scale. Diverse phylogenetic clusters of *Burkholderia* strains were observed within  
564 native legume genera of the tribes Crotalariaeae (*Aspalathus*, *Crotalaria*, *Lebeckia*, *Rafnia*),  
565 Indigoferaceae (*Indigofera*), Phaseoleae (*Bolusafra*, *Dipogon*, *Rhynchosia*), Podalyriaceae  
566 (*Amphithalea*, *Podalyria*, *Virgilia*), but without a host specific pattern (Figs. 3-4). For both  
567 chromosomal and nodulation genes, the latter symbiotic genes determining host specificity  
568 (68), a relaxed association among genetically similar rhizobia and different legume species,  
569 genera and tribes was demonstrated. The variation of host-*Burkholderia* interactions  
570 corroborates a previous rhizobial screening in selected legume genera of the tribes  
571 Hypocalypteae (*Hypocalyptus*) and Podalyriaceae (*Cyclopia*, *Podalyria*, *Virgilia*) (24,30). In  
572 South America, a similar relaxed host specific interaction has been described for  
573 *Burkholderia* and their mimosoid hosts (10,62). The predominance or prevalence of  
574 *Burkholderia* strains in both papilionoid and mimosoid legumes, but without a host specific  
575 pattern, indicates that the host genotype has not been a major factor on the local *Burkholderia*  
576 distribution. This observation is in line with the current host range study, showing selected

577 South African papilionoid species able to form effective nodules with the strains  
578 *Burkholderia tuberum* STM678<sup>T</sup> and *Burkholderia phymatum* STM815<sup>T</sup>. Strains of *B.*  
579 *phymatum*, which is found as a common symbiont of *Mimosa* in French Guiana, Papua New  
580 Guinea, India and China (12,16,17,71), has not been isolated from field nodules collected in  
581 the Fynbos, yet they are able to nodulate selected papilionoids (*Dipogon*, *Cyclopia*, *Virgilia*).  
582 The promiscuous character of the papilionoid-*Burkholderia* symbiosis has previously been  
583 demonstrated in other species of Podalyrieae (12) and Phaseoleae (29,49,52).

584 Although Fynbos legumes were generally associated with diverse *Burkholderia* species,  
585 individual root nodules consistently accommodated a single strain. The observation of a  
586 single *Burkholderia* strain per nodule may suggest high selective constraints of the host  
587 towards their symbiont. In order to retain a stable and mutualistic interaction, legumes  
588 generally hinder the emergence of opportunistic rhizobial strains and select cooperative (i.e.  
589 effectively nitrogen-fixing rhizobia) ones over non-beneficial symbionts (referred to as  
590 partner choice) (80,81) by providing only one beneficial symbiont with ample carbon  
591 resources while an uncooperative nodule occupant is disfavored with host resources (referred  
592 to as host sanctions) (82,83). However, the general observation of a relaxed interaction or  
593 accommodation of diverse rhizobial strains per host individual may indicate that the one-  
594 nodule one-strain interaction is a result of high competitiveness for nodulation among  
595 rhizobial strains, rather than to selection by the host plant.

596

#### 597 *Nodulation of Fynbos legumes outside their distribution range*

598 A legume growing in non-native soil can only form nodules when naturalized populations of  
599 compatible rhizobia are available in the soil. In our inoculation experiment, *siratro* and *Ps.*  
600 *pinnata* nodulated in soils collected from South Africa and Malawi, whereas *P. calypttrata*, *I.*  
601 *filifolia* and *D. lignosus* were nodule-free in the Malawian soil. The inability to form nodules

602 in Malawian soil suggests that these legumes, known to exhibit a strong host preference for  
603 *Burkholderia* (24,26,84), did not find their specific *Burkholderia* symbionts in the Malawian  
604 (Savanna) soil, which was substantially higher in pH compared to the Cape soil. The  
605 occurrence and success of *Burkholderia* in South African (Fynbos) soils, but also in the  
606 South American Cerrado/Caatinga biomes, can be linked with the general ecological  
607 adaptation of these symbionts to acidic soil conditions, which may play a prominent role as  
608 ecological driver on the rhizobial diversity (19,27,28,32). In Malawi, legume nodulation by  
609 *Burkholderia* has never been reported as far as we know, and further *Burkholderia* surveys in  
610 other African soils are needed to provide evidence for a more limited distribution pattern on  
611 the African continent with the Fynbos biome reported as a major center of diversity.

612 The inability of legumes to form a symbiosis with *Burkholderia* in Malawian soils does not  
613 necessarily indicate the absence of *Burkholderia* in other regions of Africa (e.g. see report of  
614 *Burkholderia* nodulating the non-native common bean in Moroccan soil (31)), but may also  
615 result from incompatible types of symbiosis genes within local *Burkholderia* communities. In  
616 this context, the observation that *Mimosa pudica* is unable to nodulate within the  
617 *Burkholderia*-rich Fynbos soils, strongly suggests that the necessary mimosoid type  
618 nodulation genes (which are genetically distinct from the papilionoid type nodulation genes)  
619 are not naturally occurring in these soils. The absence of effective rhizobia and their  
620 compatible symbiosis genes is a potential barrier to the colonization of novel habitats by the  
621 host legumes. For exotic legumes such as *Mimosa pudica*, it appears that the host needs to  
622 bring its own native symbionts into the new environment for an optimal and successful  
623 colonization and distribution (15,16).

624 In contrast to legumes with a specific preference for *Burkholderia*, *Ps. pinnata* was nodulated  
625 by *Mesorhizobium* in Fynbos soils and by *Bradyrhizobium* in Malawian soil, indicating a  
626 more relaxed interaction, albeit one that does not involve beta-rhizobia. Although field

627 nodules of this genus have been consistently associated with *Mesorhizobium* in the Fynbos  
628 (26), *Bradyrhizobium* was also able to nodulate *Psoralea* effectively, probably in the absence  
629 of their preferred *Mesorhizobium* symbionts in these Savanna soils. The genus *Psoralea* has a  
630 centre of diversity in the Fynbos but several species occur in montane grasslands in North-  
631 Eastern South Africa, Mozambique and Swaziland, and two species are naturalized in  
632 Australia (47). The current *Mesorhizobium* diversity from Fynbos *Psoralea* has been placed  
633 in a separate cluster unrelated to known 16S rRNA or *nodA* gene types from other African  
634 localities, suggesting rhizobial strains restricted to the Cape region. The *Bradyrhizobium*  
635 isolates from the Malawian soils, however, were closely related to *B. elkanii*, and are  
636 geographically widespread and able to nodulate a broad range of legumes from different  
637 continents (65,85-88). In a recent study by Parker (89), a phylogenetic analysis on a broad  
638 sampling of *Bradyrhizobium* strains from diverse plant groups provided evidence for a broad  
639 host range of most bradyrhizobia lineages, including *B. elkanii*, that are associated with  
640 diverse legume tribes.

641

#### 642 *Concluding remarks*

643 *Burkholderia* populations, like many free-living microbes and other (classical) rhizobial  
644 groups, are widespread and occur on different continents (except Antarctica and Europe), a  
645 phenomenon which can be explained by their capacity for long-distance dispersal. By  
646 investigating nodulation genes of publicly available sequence data, rather than taxonomic  
647 identities (16S rRNA types), we observed a strong biogeographic relationship, which  
648 corresponds largely to two main groups of *Burkholderia* with distinct host related affinities.  
649 Indeed, various phylogenetic studies have described taxonomically diverse papilionoid- and  
650 mimosoid-associated rhizobia with a geographical structure preserved in the nodulation  
651 genes (*nodA* and *nodC*), supporting the hypothesis that traits (i.e. nodulation genes) rather

652 than taxon names (i.e. chromosomal genes) are the fundamental units of biogeography (90).  
653 In contrast to the global investigation of *Burkholderia*, regionally distributed species in the  
654 Fynbos did not show any geographical distribution pattern. Within the Cape region, genetic  
655 variation for both chromosomal and nodulation genes was unrelated to geographical or host  
656 factors, suggesting that nodulating *Burkholderia* are omnipresent in the Fynbos biome and do  
657 not constrain the distribution of their native host legumes in terms of compatible symbionts.  
658

659 **Funding Information**

660 This work was supported by the National Research Foundation (NRF) project grant  
661 Biology of Cape Legumes. BL owe special gratitude to the Research Foundation  
662 Flanders (FWO, 1273513N), the Claude Leon Foundation and the Smuts Memorial  
663 Botanical Fellowship. The funders had no role in study design, data collection and  
664 interpretation, or the decision to submit the work for publication.

665

666 **Acknowledgments**

667 We would like to acknowledge CapeNature and SanParks Table Mountain and  
668 Eastern Cape Parks Board for access within the nature reserves.

669

670

671 **References**

- 672 1. **Bissett A, Richardson AE, Baker G, Wakelin S, Thrall PH.** 2010. Life history  
673 determines biogeographical patterns of soil bacterial communities over multiple  
674 spatial scales. *Mol Ecol* **19**:4315-4327.
- 675 2. **Finlay BJ.** 2002. Global dispersal of free-living microbial eukaryote species.  
676 *Science* **296**:1061-1063.
- 677 3. **Cho J-C, Tiedje JM.** 2000. Biogeography and degree of endemism of fluorescent  
678 *Pseudomonas* strains in soil. *Appl Environ Microbiol* **66**:5448-5456.
- 679 4. **Horner-Devine MC, Lage M, Hughes JB, Bohannan BJM.** 2004. A taxa-area  
680 relationship for bacteria. *Nature* **432**:750-753.
- 681 5. **Reche I, Pulido-Villena E, Morales-Baquero R, Casamayor EO.** 2005. Does  
682 ecosystem size determine aquatic bacterial richness? *Ecology* **86**:1715-1722.
- 683 6. **Bala A, Murphy P, Giller KE.** 2003. Distribution and diversity of rhizobia  
684 nodulating agroforestry legumes in soils from three continents in the tropics. *Mol*  
685 *Ecol* **12**:917-929.
- 686 7. **Rahi P, Kapoor R, Young JPW, Gulati A.** 2012. A genetic discontinuity in root-  
687 nodulating bacteria of cultivated pea in the Indian trans-Himalayas. *Mol Ecol* **21**:145-  
688 159.
- 689 8. **Silva C, Vinuesa P, Eguiarte LE, Souza V, Martinez-Romero E.** 2005.  
690 Evolutionary genetics and biogeographic structure of *Rhizobium gallicum* sensu lato,  
691 a widely distributed bacterial symbiont of diverse legumes. *Mol Ecol* **14**:4033-4050.
- 692 9. **Van Cauwenberghe J, Verstraete B, Lemaire B, Lievens B, Michiels J,**  
693 **Honnay O.** 2014. Population structure of root nodulating *Rhizobium leguminosarum*  
694 in *Vicia cracca* populations at local to regional geographic scales. *Syst Appl*  
695 *Microbiol* **37**:613-621.
- 696 10. **Bontemps C, Elliott GN, Simon MF, Dos Reis FB, Gross E, Lawton RC, Neto**  
697 **NE, de Fatima Loureiro M, de Faria SM, Sprent JI, James EK, Young JPW.**  
698 2010. *Burkholderia* species are ancient symbionts of legumes. *Mol Ecol* **19**:44-52.
- 699 11. **Bontemps C, Rogel MA, Wiechmann A, Mussabekova A, Moody S, Simon**  
700 **MF, Moulin L, Elliott GN, Lacerat-Didier L, Dasilva C, Grether R, Camargo-**  
701 **Ricalde SL, Chen W, Sprent JI, Martínez-Romero E, Young JPW, James EK.**  
702 2016. Endemic *Mimosa* species from Mexico prefer alphaproteobacterial rhizobial  
703 symbionts. *New Phytol* **209**:319-333.
- 704 12. **Gyaneshwar P, Hirsch AM, Moulin L, Chen W-M, Elliott GN, Bontemps C,**  
705 **Estrada-de los Santos P, Gross E, dos Reis Jr. FB, Sprent JI, Young JPW, James**  
706 **EK.** 2011. Legume-nodulating betaproteobacteria: diversity, host range, and future  
707 prospects. *Mol Plant Microbe Interact* **24**:1276-1288.
- 708 13. **Chen W-M, James EK, Chou J-H, Sheu S-Y, Yang S-Z, Sprent JI.** 2005b.  $\beta$ -  
709 Rhizobia from *Mimosa pigra*, a newly discovered invasive plant in Taiwan. *New*  
710 *Phytol* **168**:661-675.
- 711 14. **Gehlot HS, Tak N, Kaushik M, Mitra S, Chen W-M, Poweleit N, Panwar D,**  
712 **Poonar N, Parihar R, Tak A, Sankhla IS, Ojha A, Rao SR, Simon MF, dos Reis**  
713 **Junior FB, Perigolo N, Tripathi AK, Sprent JI, Young JPW, James EK,**  
714 **Gyaneshwar P.** 2013. An invasive *Mimosa* in India does not adopt the symbionts of  
715 its native relatives. *Ann Bot* **112**:179-196.
- 716 15. **Parker MA, Wurtz AK, Paynter Q.** 2007. Nodule symbiosis of invasive  
717 *Mimosa pigra* in Australia and in ancestral habitats: a comparative analysis. *Biol*  
718 *Invasions* **9**:127-138.



- 719 16. Liu X, Wei S, Wang F, James EK, Guo X, Zagar C, Xia LG, Dong X, Wang  
720 YP. 2012. *Burkholderia* and *Cupriavidus* spp. are the preferred symbionts of *Mimosa*  
721 spp. in southern China. *FEMS Microbiol Ecol* **80**:417-426.
- 722 17. Elliott GN, Chen W-M, Chou J-H, Wang H-C, Sheu S-Y, Perin L, Reis VM,  
723 Moulin L, Simon MF, Bontemps C, Sutherland JM, Bessi R, de Faria SM,  
724 Trinick MJ, Prescott AR, Sprent JI, James EK. 2007b. *Burkholderia phymatum* is  
725 a highly effective nitrogen-fixing symbiont of *Mimosa* spp. and fixes nitrogen *ex*  
726 *planta*. *New Phytol* **173**:168-180.
- 727 18. Andrus AD, Andam C, Parker MA. 2012. American origin of *Cupriavidus*  
728 bacteria associated with invasive *Mimosa* legumes in the Philippines. *FEMS*  
729 *Microbiol Ecol* **80**:747-750.
- 730 19. Suárez-Moreno ZR, Caballero-Mellado J, Coutinho BG, Mendonça-Previato  
731 L, James EK, Venturi V. 2012. Common features of environmental and potentially  
732 beneficial plant-associated *Burkholderia*. *Microb Ecol* **63**:249-266.
- 733 20. Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green  
734 JL, Horner-Devine MC, Kane M, Krumins JA, Kuske CR. 2006. Microbial  
735 biogeography: putting microorganisms on the map. *Nat Rev Microbiol* **4**:102-112.
- 736 21. De Meyer SE, Cnockaert M, Ardley JK, van Wyk B-E, Vandamme PA,  
737 Howieson JG. 2014. *Burkholderia dilworthii* sp. nov. isolated from *Lebeckia*  
738 *ambigua* root nodules from South Africa. *Int J Syst Evol Microbiol* **64**:1090-1095.
- 739 22. De Meyer SE, Cnockaert M, Ardley JK, Maker G, Yates R, Howieson JG,  
740 Vandamme P. 2013a. *Burkholderia sprentiae* sp. nov. isolated from *Lebeckia*  
741 *ambigua* root nodules from South Africa. *Int J Syst Evol Microbiol* **63**:3950-3957.
- 742 23. Howieson JG, De Meyer SE, Vivas-Marfisi A, Ratnayake S, Ardley JK, Yates  
743 RJ. 2013. Novel *Burkholderia* bacteria isolated from *Lebeckia ambigua*, a perennial  
744 suffrutescent legume of the fynbos. *Soil Biol Biochem* **60**:55-64.
- 745 24. Beukes CW, Venter SN, Law IJ, Phalane FL, Steenkamp ET. 2013. South  
746 African papilionoid legumes are nodulated by diverse *Burkholderia* with unique  
747 nodulation and nitrogen-fixation loci. *PLoS One* **8**:e68406.
- 748 25. Steenkamp ET, van Zyl E, Beukes CW, Avontuur JR, Chan WY, Palmer M,  
749 Mthombeni LS, Phalane FL, Sereme TK, Venter SN. 2015. *Burkholderia*  
750 *kirstenboschensis* sp. nov. nodulates papilionoid legumes indigenous to South Africa.  
751 *Syst Appl Microbiol* **38**:545-554.
- 752 26. Lemaire B, Dlodlo O, Chimphango S, Stirton CH, Schrire BD, Boatwright  
753 JS, Honnay O, Smets E, Sprent JI, Euan KJ, Muasya AM. 2015. Symbiotic  
754 diversity, specificity and distribution of rhizobia in native legumes of the core Cape  
755 Subregion (South Africa). *FEMS Microbiol Ecol* **91**:1-17.
- 756 27. De Meyer SE, Cnockaert M, Ardley JK, Trengove RD, Garau G, Howieson  
757 JG, Vandamme P. 2013b. *Burkholderia rhynchosiae* sp. nov. isolated from  
758 *Rhynchosia ferulifolia* root nodules from South Africa. *Int J Syst Evol Microbiol*  
759 **63**:3944-2949.
- 760 28. Garau G, Yates RJ, Deiana P, Howieson JG. 2009. Novel strains of nodulating  
761 *Burkholderia* have a role in nitrogen fixation with papilionoid herbaceous legumes  
762 adapted to acid, infertile soils. *Soil Biol Biochem* **41**:125-134.
- 763 29. Elliott GN, Chen WM, Bontemps C, Chou JH, Young JPW, Sprent JI, James  
764 EK. 2007. Nodulation of *Cyclopia* spp. (Leguminosae, Papilionoideae) by  
765 *Burkholderia tuberum*. *Ann Bot* **100**:1403-1411.
- 766 30. Lemaire B, Van Cauwenbergh J, Verstraete B, Chimphango S, Stirton C,  
767 Honnay O, Smets E, Sprent J, James EK, Muasya AM. 2016. Characterization of  
768 the papilionoid-*Burkholderia* interaction in the Fynbos biome: The diversity and



- 769 distribution of beta-rhizobia nodulating *Podalyria calyptata* (Fabaceae, Podalyrieae).  
 770 Syst Appl Microbiol **39**:41-48.
- 771 31. **Talbi C, Delgado MJ, Girard L, Ramírez-Trujillo A, Caballero-Mellado J,**  
 772 **Bedmar EJ.** 2010. *Burkholderia phymatum* strains capable of nodulating *Phaseolus*  
 773 *vulgaris* are present in Moroccan soils. Appl Environ Microbiol **76**:4587-4591.
- 774 32. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and  
 775 high throughput. Nucleic Acids Res **32**:1792-1797.
- 776 33. **Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB,**  
 777 **Lesniewski RA, Oakley BB, Parks DH, Robinson CJ.** 2009. Introducing mothur:  
 778 open-source, platform-independent, community-supported software for describing  
 779 and comparing microbial communities. Appl Environ Microbiol **75**:7537-7541.
- 780 34. **Huson DH, Bryant D.** 2006. Application of phylogenetic networks in  
 781 evolutionary studies. Mol Biol Evol **23**:254-267.
- 782 35. **Bryant D, Moulton V.** 2004. Neighbor-Net: An agglomerative method for the  
 783 construction of phylogenetic networks. Mol Biol Evol **21**:255-265.
- 784 36. **Vincent JM.** 1970. A manual for the practical study of the root-nodule bacteria. A  
 785 manual for the practical study of the root-nodule bacteria.
- 786 37. **Weisburg WG, Barns SM, Pelletier DA, Lane DJ.** 1991. 16S ribosomal DNA  
 787 amplification for phylogenetic study. J Bacteriol **173**:697-703.
- 788 38. **Gaunt MW, Turner SL, Rigottier-Gois L, Lloyd-Macgilp SA, Young JP.**  
 789 2001. Phylogenies of *atpD* and *recA* support the small subunit rRNA-based  
 790 classification of rhizobia. Int J Syst Evol Microbiol **51**:2037-2048.
- 791 39. **Haukka K, Lindström K, Young JPW.** 1998. Three phylogenetic groups of  
 792 *nodA* and *nifH* genes in *Sinorhizobium* and *Mesorhizobium* isolates from leguminous  
 793 trees growing in Africa and Latin America. Appl Environ Microbiol **64**:419-426.
- 794 40. **Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman J, Smith JA, Struhl**  
 795 **K.** 1992. Short protocols in molecular biology. 5th ed. John Wiley & Sons, New  
 796 York, NY.
- 797 41. **Ronquist F, Huelsenbeck JP.** 2003. MrBayes 3: Bayesian phylogenetic inference  
 798 under mixed models. Bioinformatics **19**:1572-1572.
- 799 42. **Posada D, Crandall KA.** 1998. Modeltest: testing the model of DNA  
 800 substitution. Bioinformatics **14**:817-817.
- 801 43. **Rambaut A, Drummond AJ.** 2007. Tracer v1. 4, Available from  
 802 <http://beast.bio.ed.ac.uk/Tracer>.
- 803 44. **Stamatakis A.** 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic  
 804 analyses with thousands of taxa and mixed models. Bioinformatics **22**:2688-2690.
- 805 45. **R Core Team** (2005) R: A language and environment for statistical computing. R  
 806 Foundation for Statistical Computing. Vienna, Austria, 2013. [http://www.R-](http://www.R-project.org)  
 807 [project.org](http://www.R-project.org).
- 808 46. **Peakall R, Smouse PE.** 2006. GENALEX 6: genetic analysis in Excel.  
 809 Population genetic software for teaching and research. Mol Ecol Notes **6**:288-295.
- 810 47. **Stirton CH, Stajsic V, Bello A.** 2015. Naturalised species of *Psoralea* (Fabaceae:  
 811 Psoraleeae) in Australia. Muelleria **33**:97-107.
- 812 48. **Lewis GP, Schrire B, Mackinder B, Lock M.** 2005. Legumes of the World.  
 813 Royal Botanic Gardens Kew.
- 814 49. **Liu WYY, Ridgway HJ, James TK, James EK, Chen W-M, Sprent JI, Young**  
 815 **JPW, Andrews M.** 2014. *Burkholderia* sp. induces functional nodules on the South  
 816 African invasive legume *Dipogon lignosus* (Phaseoleae) in New Zealand soils.  
 817 Microb Ecol **68**:542-555.

- 818 50. **Chen W-M, de Faria SM, Stralioetto R, Pitard RM, Simoes-Araujo JL, Chou**  
819 **J-H, Chou Y-J, Barrios E, Prescott AR, Elliott GN.** 2005a. Proof that *Burkholderia*  
820 strains form effective symbioses with legumes: a study of novel *Mimosa*-nodulating  
821 strains from South America. *Appl Environ Microbiol* **71**:7461-7471.
- 822 51. **Trinick MJ, Miller C, Hadobas PA.** 1991. Formation and structure of root  
823 nodules induced on *Macropodium atropurpureum* inoculated with various species of  
824 *Rhizobium*. *Can J Bot* **69**:1520-1532.
- 825 52. **Angus AA, Lee A, Lum MR, Shehayeb M, Hessabi R, Fujishige NA,**  
826 **Yerrapragada S, Kano S, Song N, Yang P.** 2013. Nodulation and effective nitrogen  
827 fixation of *Macropodium atropurpureum* (siratiro) by *Burkholderia tuberum*, a  
828 nodulating and plant growth promoting beta-proteobacterium, are influenced by  
829 environmental factors. *Plant Soil* **369**:543-562.
- 830 53. **Somasegaran P, Hoben HJ.** 1994. Handbook for rhizobia: methods in legume-  
831 rhizobium technology. Springer-Verlag, New York.
- 832 54. **Vandamme P, Goris J, Chen W-M, De Vos P, Willems A.** 2002. *Burkholderia*  
833 *tuberum* sp. nov. and *Burkholderia phymatum* sp. nov., nodulate the roots of tropical  
834 legumes. *Syst Appl Microbiol* **25**:507-512.
- 835 55. **dos Reis Junior FB, Simon MF, Gross E, Boddey RM, Elliott GN, Neto NE,**  
836 **de Fatima Loureiro M, de Queiroz LP, Scotti MR, Chen W-M, Rubio MC, de**  
837 **Faria SM, Bontemps C, Goi SR, Young JPW, Sprent JI, James EK.** 2010.  
838 Nodulation and nitrogen fixation by *Mimosa* spp. in the Cerrado and Caatinga biomes  
839 of Brazil. *New Phytol* **186**:934-946.
- 840 56. **Cohan FM.** 2002. What are bacterial species? *Annu Rev Microbiol* **56**:457-487.
- 841 57. **Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ,**  
842 **Stackebrandt E, Peer YVD, Vandamme P, Thompson FL, Swings J.** 2005. Re-  
843 evaluating prokaryotic species. *Nat Rev Microbiol* **3**:733-739.
- 844 58. **Stackebrandt E, Frederiksen W, Garrity GM, Grimont PAD, Kämpfer P,**  
845 **Maiden MCJ, Nesme X, Rosselló-Mora R, Swings J, Trüper HG, Vauterin L,**  
846 **Ward AC, Whitman WB.** 2002. Report of the ad hoc committee for the re-  
847 evaluation of the species definition in bacteriology. *Int J Syst Evol Microbiol*  
848 **52**:1043-1047.
- 849 59. **Doolittle WF, Papke RT.** 2006. Genomics and the bacterial species problem.  
850 *Genome Biol* **7**:116.
- 851 60. **Kim M, Oh H-S, Park S-C, Chun J.** 2014. Towards a taxonomic coherence  
852 between average nucleotide identity and 16S rRNA gene sequence similarity for  
853 species demarcation of prokaryotes *Int J Syst Evol Microbiol* **64**:1825-1825.
- 854 61. **Wright S.** 1943. Isolation by distance. *Genetics* **28**:114-138.
- 855 62. **Bournaud C, de Faria SM, dos Santos JMF, Tisseyre P, Silva M, Chaintreuil**  
856 **C, Gross E, James EK, Prin Y, Moulin L.** 2013. *Burkholderia* species are the most  
857 common and preferred nodulating symbionts of the *Piptadenia* group (tribe  
858 Mimoseae). *PLoS One* **8**:e63478.
- 859 63. **Ormeño-Orrillo E, Rogel MA, Chueire LMO, Tiedje JM, Martínez-Romero**  
860 **E, Hungria M.** 2012. Genome sequences of *Burkholderia* sp. strains CCGE1002 and  
861 H160, isolated from legume nodules in Mexico and Brazil. *J Bacteriol* **194**:6927-  
862 6927.
- 863 64. **Crisp MD, Arroyo MTK, Cook LG, Gandolfo MA, Jordan GJ, McGlone MS,**  
864 **Weston PH, Westoby M, Wilf P, Linder HP.** 2009. Phylogenetic biome  
865 conservatism on a global scale. *Nature* **458**:754-756.

- 866 65. **Stepkowski T, Watkin E, McInnes A, Gurda D, Gracz J, Steenkamp ET.**  
867 2012. Distinct *Bradyrhizobium* communities nodulate legumes native to temperate and  
868 tropical monsoon Australia. *Mol Phylogenet Evol* **63**:265-277.
- 869 66. **Lafay B, Burdon JJ.** 1998. Molecular diversity of rhizobia occurring on native  
870 shrubby legumes in southeastern Australia. *Appl Environ Microbiol* **64**:3989-3997.
- 871 67. **Roche P, Maillet F, Plazanet C, Debelle F, Ferro M, Truchet G, Promé J-C,**  
872 **Dénarié J.** 1996. The common *nodABC* genes of *Rhizobium meliloti* are host-range  
873 determinants. *Proc Natl Acad Sci U S A* **93**:15305-15310.
- 874 68. **Spaink HP.** 2000. Root nodulation and infection factors produced by rhizobial  
875 bacteria. *Annu Rev Microbiol* **54**:257-288.
- 876 69. **Perret X, Staehelin C, Broughton WJ.** 2000. Molecular basis of symbiotic  
877 promiscuity. *Microbiol Mol Biol Rev* **64**:180-201.
- 878 70. **Kobayashi H, Broughton WJ.** 2008. Fine-tuning of symbiotic genes in rhizobia:  
879 Flavonoid signal transduction cascade, vol 7. Springer Netherlands.
- 880 71. **Mishra RPN, Tisseyre P, Melkonian R, Chaintreuil C, Miche L, Klonowska**  
881 **A, Gonzalez S, Bena G, Laguerre G, Moulin L.** 2012. Genetic diversity of *Mimosa*  
882 *pudica* rhizobial symbionts in soils of French Guiana: investigating the origin and  
883 diversity of *Burkholderia phymatum* and other beta-rhizobia. *FEMS Microbiol Ecol*  
884 **79**:487-503.
- 885 72. **Rogel MA, Ormeno-Orrillo E, Romero EM.** 2011. Symbiobars in rhizobia  
886 reflect bacterial adaptation to legumes. *Syst Appl Microbiol* **34**:96-104.
- 887 73. **Martínez-Romero E.** 2009. Coevolution in *Rhizobium*-legume symbiosis? *DNA*  
888 *Cell Biol* **28**:361-370.
- 889 74. **Moulin L, Klonowska A, Caroline B, Booth K, Vriezen JAC, Melkonian R,**  
890 **James EK, Young JPW, Bena G, Hauser L.** 2014. Complete Genome sequence of  
891 *Burkholderia phymatum* STM815T, a broad host range and efficient nitrogen-fixing  
892 symbiont of *Mimosa* species. *Standards in Genomic Sciences* **9**:763.
- 893 75. **Porter SS, Stanton ML, Rice KJ.** 2011. Mutualism and adaptive divergence: co-  
894 invasion of a heterogeneous grassland by an exotic legume-rhizobium symbiosis.  
895 *PLoS One* **6**:e27935.
- 896 76. **Crisóstomo JA, Rodríguez-Echeverría S, Freitas H.** 2013. Co-introduction of  
897 exotic rhizobia to the rhizosphere of the invasive legume *Acacia saligna*, an  
898 intercontinental study. *Appl Soil Ecol* **64**:118-126.
- 899 77. **Rodríguez- Echeverría S.** 2010. Rhizobial hitchhikers from Down Under:  
900 invasional meltdown in a plant-bacteria mutualism? *J Biogeogr* **37**:1611-1622.
- 901 78. **Ndlovu J, Richardson DM, Wilson JRU, Le Roux JJ.** 2013. Co-invasion of  
902 South African ecosystems by an Australian legume and its rhizobial symbionts. *J*  
903 *Biogeogr* **40**:1240-1251.
- 904 79. **Sheu SY, Chen MH, Liu WYY, Andrews M, James EK, Ardley JK, De**  
905 **Meyer SE, James TK, Howieson JG, Coutinho BG, Chen W-M.** 2015.  
906 *Burkholderia dipogonis* sp. nov., isolated from root nodules of *Dipogon lignosus* in  
907 New Zealand and Western Australia. *Int J Syst Evol*  
908 *Microbiol*:10.1099/ijsem.1090.000639.
- 909 80. **Sachs JL, Mueller UG, Wilcox TP, Bull JJ.** 2004. The evolution of cooperation.  
910 *Q Rev Biol* **79**:135-160.
- 911 81. **Simms EL, Taylor DL, Povich J, Shefferson RP, Sachs JL, Urbina M,**  
912 **Tausczik Y.** 2006. An empirical test of partner choice mechanisms in a wild legume-  
913 rhizobium interaction. *Proc R Soc Lond B Biol Sci* **273**:77-81.
- 914 82. **Kiers ET, Denison RF.** 2008. Sanctions, cooperation, and the stability of plant-  
915 rhizosphere mutualisms. *Annu Rev Ecol Evol Syst* **39**:215-236.

- 916 83. **Kiers ET, Rousseau RA, West SA, Denison RF.** 2003. Host sanctions and the  
917 legume-rhizobium mutualism. *Nature* **425**:78-81.
- 918 84. **Sprent JI, Ardley JK, James EK.** 2013. From North to South: A latitudinal look  
919 at legume nodulation processes. *S Afr J Bot* **89**:31-41.
- 920 85. **Aserse AA, Räsänen LA, Aseffa F, Hailemariam A, Lindström K.** 2012.  
921 Phylogenetically diverse groups of *Bradyrhizobium* isolated from nodules of  
922 *Crotalaria* spp., *Indigofera* spp., *Erythrina brucei* and *Glycine max* growing in  
923 Ethiopia. *Mol Phylogenet Evol* **65**:595-609.
- 924 86. **Steenkamp ET, Stępkowski T, Przymusiak A, Botha WJ, Law IJ.** 2008.  
925 Cowpea and peanut in southern Africa are nodulated by diverse *Bradyrhizobium*  
926 strains harboring nodulation genes that belong to the large pantropical clade common  
927 in Africa. *Mol Phylogenet Evol* **48**:1131-1144.
- 928 87. **Stępkowski T, Hughes CE, Law IJ, Markiewicz A, Gurda D, Chlebicka A,**  
929 **Moulin L.** 2007. Diversification of lupine *Bradyrhizobium* strains: evidence from  
930 nodulation gene trees. *Appl Environ Microbiol* **73**:3254-3264.
- 931 88. **Stępkowski T, Moulin L, Krzyżńska A, McInnes A, Law IJ, Howieson J.**  
932 2005. European origin of *Bradyrhizobium* populations infecting lupins and serradella  
933 in soils of Western Australia and South Africa. *Appl Environ Microbiol* **71**:7041-  
934 7052.
- 935 89. **Parker MA.** 2015. The spread of *Bradyrhizobium* lineages across host legume  
936 clades: from *Abarema* to *Zygia*. *Microb Ecol* **69**:630-640.
- 937 90. **Green JL, Bohannan BJ, Whitaker RJ.** 2008. Microbial biogeography: from  
938 taxonomy to traits. *Science* **320**:1039-1043.
- 939

**Figure Legends****Table 1.**

Occurrence of 16S rRNA OTUs and sequence clusters of symbiosis genes (*nifH*, *nodA* and *nodC*) shared among different continents (South America - SAM, Africa - AFR, Australasia - AUS and Asia - ASI) and host subfamilies (Mimosoideae - MIM and Papilionoideae - PAP). The host genera and reference strains of *Burkholderia* are listed per group (98.5% sequence similarity threshold value). - = not present

<sup>1</sup> *Burkholderia phymatum* STM815<sup>T</sup> was allegedly isolated from the papilionoid *Machaerium lunatum* in French Guiana but has never been proven to renodulate its original host (12) or an alternative *Machaerium* species (*M. brasilense*, (17)).

16S rRNA OTU	Geographic Distribution	Host Subfamily	Reference strain	Host Genera of Mimosoideae	Host genera of Papilionoideae
1	SAM-AFR	MIM-PAP	<i>B. tuberum</i>	<i>Mimosa</i>	<i>Amphithalea</i> , <i>Aspalathus</i> , <i>Cyclopia</i> , <i>Hypocalyptus</i> , <i>Lebeckia</i> , <i>Macroptilium</i> , <i>Podalyria</i> , <i>Rhynchosia</i> , <i>Virgilia</i>
3	AFR-AUS	PAP	<i>B. dipogonis</i>	-	<i>Bolusafra</i> , <i>Crotalaria</i> , <i>Cyclopia</i> , <i>Dipogon</i> , <i>Hypocalyptus</i> , <i>Podalyria</i> , <i>Rafnia</i> , <i>Virgilia</i>
4	SAM-ASI-AUS	MIM	<i>B. mimosarum</i>	<i>Mimosa</i>	-
5	SAM-AUS-ASI-AFR	MIM-PAP	<i>B. phymatum</i>	<i>Mimosa</i> , <i>Parapiptadenia</i> , <i>Piptadenia</i>	<i>Machaerium</i> <sup>1</sup> , <i>Phaseolus</i>
6	SAM-ASI	MIM	<i>B. sabiae</i>	<i>Abarema</i> , <i>Mimosa</i> , <i>Parapiptadenia</i>	-
7	SAM-AUS	MIM	<i>B. diazotrophica</i>	<i>Mimosa</i> , <i>Piptadenia</i> , <i>Anadenanthera</i>	-
8	SAM-AUS	MIM	-	<i>Mimosa</i>	-
12	AFR-SAM	MIM-PAP	-	<i>Mimosa</i>	<i>Hypocalyptus</i>

Sequence cluster <i>nifH</i> gene					
1	SAM	MIM-PAP	<i>B. tuberum</i>	<i>Mimosa</i>	<i>Macroptilium</i>
2	AFR-AUS	PAP	<i>B. tuberum</i>	-	<i>Amphithalea, Aspalathus, Crotalaria, Cyclopia, Dipogon, Hypocalyptus, Indigofera, Lebeckia, Podalyria, Rafnia, Virgilia</i>
4	SAM-AUS-ASI-AFR	MIM-PAP	<i>B. phymatum/B. diazotrophica</i>	<i>Abarema, Anadenanthera, Mimosa</i>	<i>Machaerium</i> <sup>1</sup> , <i>Phaseolus</i>
5	SAM-ASI-AUS	MIM	<i>B. mimosarum</i>	<i>Mimosa</i>	-
9	AFR-AUS	PAP	<i>B. rhynchosiae</i>	-	<i>Dipogon, Rhynchosia</i>
Sequence cluster <i>nodA</i> gene					
3	SAM-ASI-AUS	MIM-PAP <sup>1</sup>	<i>B. phymatum</i>	<i>Mimosa</i>	<i>Machaerium</i> <sup>1</sup>
5	SAM-ASI-AUS	MIM	<i>B. mimosarum</i>	<i>Mimosa</i>	-
6	AFR-AUS	PAP	<i>B. dipogonis</i>	-	<i>Crotalaria, Cyclopia, Dipogon, Hypocalyptus, Virgilia, Podalyria</i>
10	AFR-AUS	PAP	<i>B. rhynchosiae</i>	-	<i>Bolusafr, Dipogon, Indigofera, Rhynchosia</i>
14	SAM	MIM-PAP	<i>B. tuberum</i>	<i>Mimosa</i>	<i>Macroptilium</i>
16	SAM-ASI	MIM	<i>B. sabiae</i>	<i>Mimosa</i>	-
Sequence cluster <i>nodC</i> gene					
4	SAM-AUS-ASI-AFR	MIM-PAP	<i>B. phymatum/B. diazotrophica</i>	<i>Anadenanthera, Mimosa</i>	<i>Phaseolus</i>
9	AFR-AUS	PAP	<i>B. dipogonis</i>	-	<i>Crotalaria, Dipogon</i>
10	SAM-ASI	MIM	<i>B. mimosarum</i>	<i>Mimosa</i>	-

**Table 2.**

Nodulation of selected Fynbos species after inoculation with *Burkholderia tuberum* STM678<sup>T</sup> or *B. phymatum* STM815<sup>T</sup>. E = effective nodulation; I = ineffective nodulation, considered if inoculated plants are not greener than uninoculated controls and only few and white nodules are visible; - = not tested. New reports of nodulation are indicated in bold.

<sup>1</sup>Nodules tested with both *Burkholderia tuberum* STM678<sup>T</sup> and STM678GFP.

\*Data from Elliott *et al.* (29)

Tribe	Legume species tested	<i>Burkholderia tuberum</i> STM678 <sup>T</sup>	<i>Burkholderia phymatum</i> STM815 <sup>T</sup>
Crotalariaeae	<i>Aspalathus carnosa</i> Bergius	E	no nodules
	<i>Lebeckia ambigua</i> E.Mey.	E	no nodules
Hypocalypteae	<i>Hypocalyptus coluteoides</i> (Lam.) R.Dahlgren	E	-
	<i>Hypocalyptus sophoroides</i> (P.J.Bergius) Baill.	E	I
Indigoferaeae	<i>Indigofera filifolia</i> Thunb.	E	-
Podalyrieae	<i>Amphithalea ericifolia</i> (L.) Eckl. & Zeyh	E	I
	<i>Calpurnia aurea</i> (Aiton) Benth.	no nodules	-
	<b><i>Calpurnia glabrata</i> Brummitt</b>	I	-
	<i>Calpurnia intrusa</i> (W.T.Aiton) E.Mey.	no nodules	-
	<b><i>Calpurnia sericea</i> Harv.</b>	I	-



<i>Cyclopia subternata</i> Vogel	E	E
<i>Cyclopia genistoides</i> (L.) Vent.	E*	E
<i>Cyclopia intermedia</i> E.Mey.	E*	E
<b><i>Liparia laevigata</i> Thunb.</b>	E	-
<i>Liparia splendens</i> (Burm.f.) Bos & de Wit	E	-
<i>Podalyria burchellii</i> DC.	E	-
<i>Podalyria calyptrata</i> (Retz.) Willd.	E <sup>1</sup>	-
<i>Podalyria canescens</i> E.Mey.	E <sup>1</sup>	-
<b><i>Podalyria leipoldtii</i> L.Bolus</b>	E	-
<i>Podalyria myrtillifolia</i> Willd.	E <sup>1</sup>	-
<b><i>Podalyria rotundifolia</i> (P.J.Bergius) A.L.Schutte</b>	E	-
<i>Podalyria sericea</i> R.Br	E	-
<i>Stirtonanthus taylorianus</i> (L.Bolus) B.-E.van Wyk & A.L.Schutte	E	-
<i>Virgilia oroboides</i> (P.J.Bergius) T.M.Salter	E <sup>1</sup>	E
<b><i>Xiphotheca fruticosa</i> (L.) A.L.Schutte &amp; B.-E.van Wyk</b>	E	-



**Figure 1**

Map of South Africa showing the geographical distribution of sampling sites within the Western and Eastern Cape Provinces. Records of our isolates are indicated with white squares, whereas samples from other studies are shown with black dots.

**Figure 2**

NeighborNet networks of (A) 16S rRNA, (B) *nifH*, (C) *nodA* and (D) *nodC* sequence types. Sequence types exclusively recorded from one continent are shown by colored circles (Africa – green circles, South America – red circles, Asia – blue circles, Australasia – yellow circles). Numbers of sequence clusters sharing isolates from different continents and/or legume subfamily are shown in grey squares as listed in Table 1. Bootstrap support values below and above 50% are shown with grey and black branches, respectively. Scale bar represents substitutions per site.

**Figure 3**

Phylogenetic tree of rhizobial isolates of the Fynbos biome based on 16S rRNA and *recA* data. Support values for the Bayesian and Maximum Likelihood analyses are given at the nodes (Bayesian posterior probabilities – bootstrap support values for the Maximum Likelihood analysis). Reference strains are shown in bold.

**Figure 4**

Phylogenetic tree of rhizobial endosymbionts based on *nodA* data. Support values for the Bayesian and Maximum Likelihood analyses are given at the nodes (Bayesian posterior probabilities – bootstrap support values for the Maximum Likelihood analysis). Reference strains are shown in bold.

**Figure 5**

Box plots of pairwise genetic distances for (A) *recA* and (B) *nodA* sequence data grouped within four spatial distance classes (0-200; 201-400; 401-600-601-800 km). Box plots represent observations within 95% confidence intervals and the whiskers

extend from the box to the highest and lowest values, excluding outliers, which are shown as circles. The line across the box indicates the median.

**Figure 6**

Phylogenetic tree based on *recA* sequences of rhizobial isolates sampled from the trapping experiments. The closest reference strains obtained from BLASTN searches (see Table S4) are included in the analyses. Bayesian support values are given at the nodes. Geographic distribution of the isolates and reference strains are shown for each taxon. Number of substitutions per site is shown on the phylogram.

**Figure 7**

Phylogenetic tree based on *nodA* sequences of rhizobial isolates sampled from the trapping experiments. The closest reference strains obtained from BLASTN searches (see Table S4) are included in the analyses. Bayesian support values are given at the nodes. Geographic distribution of the isolates and reference strains are shown for each taxon. Number of substitutions per site is shown on the phylogram.

**Figure 8**

Fluorescence (A,C,E,F) and normal transmitted light (B, D) microscopy of sections (50  $\mu\text{m}$ ) from nodules of *Podalyria calyprata* (A-B), *P. canescens* (C, D), *P. myrtillifolia* (E) and *Virgilia oroboides* (F) showing infected cells containing symbiotic bacteroids (\*) as either green fluorescent (A, C, E, F) or dense opaque (B, D) regions in the nodule center. The green-yellow colour in the nodule cortex (A, C, E, F) results from autofluorescence of lignin and suberin. Bars = 100  $\mu\text{m}$ .



















