Recombination and horizontal transfer of nodulation and ACC deaminase (*acdS*) genes within *Alpha*- and *Betaproteobacteria* nodulating legumes of the Cape Fynbos biome

Benny Lemaire[*,1,2], **Jannick Van Cauwenberghe**[2,3], **Samson Chimphango**[1], **Charles Stirton**[1], **Olivier Honnay**[2], **Erik Smets**[2,4], **A. Muthama Muasya**[1]

*Corresponding author
Department of Biological Sciences
University of Cape Town, Private Bag X3
Rondebosch 7701
South Africa
E-mail address: benny.lemaire@uct.ac.za

Plant Conservation and Population Biology
KU Leuven, Kasteelpark Arenberg 31, box 02435
3001 Heverlee
Belgium
TEL. + 32 16 32 86 36 FAX + 32 16 32 19 68
E-mail address: benny.lemaire@bio.kuleuven.be

[1]*Department of Biological Sciences, University of Cape Town, Private Bag X3, Rondebosch 7701, Cape Town, South Africa*

[2]*Plant Conservation and Population Biology, KU Leuven, Kasteelpark Arenberg 31, PO Box 2435, 3001 Heverlee, Belgium*

[3]*Centre of Microbial and Plant Genetics, KU Leuven, Kasteelpark Arenberg 20, B-3001 Leuven, Belgium*

[4]*Naturalis Biodiversity Center, Leiden University, 2300 RA Leiden, The Netherlands*

Abstract

The goal of this work is to study the evolution and the degree of horizontal gene transfer (HGT) within rhizobial genera of both *Alpha-* (*Mesorhizobium, Rhizobium*) and *Beta* (*Burkholderia*) *Proteobacteria*, originating from South African Fynbos legumes. By using a phylogenetic approach and comparing multiple chromosomal and symbiosis genes, we revealed conclusive evidence of high degrees of horizontal transfer of nodulation genes among closely related species of both groups of rhizobia, but also among species with distant genetic backgrounds (*Rhizobium* and *Mesorhizobium*), underscoring the importance of lateral transfer of symbiosis traits as an important evolutionary force among rhizobia of the Cape Fynbos biome. The extensive exchange of symbiosis genes in the Fynbos is in contrast with a lack of significant events of HGT among *Burkholderia* symbionts from the South American Cerrado and Caatinga biome. Furthermore, homologous recombination among selected housekeeping genes had a substantial impact on sequence evolution within *Burkholderia* and *Mesorhizobium*. Finally, phylogenetic analyses of the non-symbiosis *acdS* gene in *Mesorhizobium*, a gene often located on symbiosis islands, revealed distinct relationships compared to the chromosomal and symbiosis genes, suggesting a different evolutionary history and independent events of gene transfer. The observed events of HGT and incongruence between different genes necessitate caution in interpreting topologies from individual data types.

Introduction

The large-scale availability and analyses of sequence information from individual genes and complete genomes has demonstrated significant amounts of gene movements or horizontal gene transfer (HGT) in bacterial evolution (Ochman *et al.*, 2000). The acquisition of new genes and metabolic capabilities between a broad spectrum of bacteria in a non-parent-to-offspring manner impacts bacterial diversification (Jain *et al.*, 2003; Vetsigian & Goldenfeld, 2005; Boto, 2010) and ecological adaptation of the recipient cells (Preston *et al.*, 1998; Goldenfeld & Woese, 2007) and has been attributed to several mechanisms such as insertion sequences, transposons, integrons, bacteriophages, genomic islands and plasmids.

In rhizobia, the capacity to establish an effective symbiosis with the host plant and to fix atmospheric nitrogen involves the expression of nodulation (*nod*) and nitrogen fixation (*nif*) genes. These genes are part of the 'accessory' gene pool and are usually located as mobile genetic elements on either transferable plasmids (*Rhizobium* spp. and *Ensifer* spp.) or scattered across the chromosome(s) as genomic islands (*Mesorhizobium* spp. and *Bradyrhizobium* spp.) (Finan, 2002; Masson-Boivin *et al.*, 2009). In *Mesorhizobium* species (i.e. *M. amorphae* and *M huakuii*), symbiosis genes have also been detected on plasmids (Xu & Murooka, 1995; Wang *et al.*, 1999; Zhang *et al.*, 2000). Symbiosis related genes of rhizobia of the beta-subclass of *Proteobacteria*, which so far consist of the genera *Burkholderia* and *Cupriavidus*, are found on plasmids (Chen *et al.*, 2003; Amadou *et al.*, 2008; Gyaneshwar *et al.*, 2011). Extensive evidence for horizontal transmission of symbiosis genes has been revealed by conflicting or incongruent sequence data of plasmid- and chromosomal-located genes among a wide range of rhizobial lineages including both alpha- (Van Berkum *et al.*, 2003; Ormeno-Orrillo *et al.*, 2006, 2013; Barcellos *et al.*, 2007) and beta-rhizobia (Andam *et al.*, 2007; Liu *et al.*, 2012).

In the Fynbos biome, rhizobial studies have recorded *Burkholderia* (*Betaproteobacteria*) symbionts as common root nodulating species associated with the papilionoid legume flora (Kock, 2004; Elliott *et al.*, 2007b; Garau *et al.*, 2009; Gyaneshwar *et al.*, 2011; Beukes *et al.*, 2013; Howieson *et al.*, 2013; Sprent *et al.*, 2013; Lemaire *et al.*, 2015), but the characterization of the symbiosis genes as an important basis for the understanding of gene transfer during rhizobial evolution remains elusive. A recent study of Beukes *et al.* (2013) revealed conflicting relationships between chromosomal (16S rRNA and *recA*) and nodulation (*nodA*)

94 genes among rhizobia mainly from legumes of the tribe Podalyrieae, suggesting HGT
95 as an importance force in the evolution of South African *Burkholderia*. Interestingly,
96 *Burkholderia* species isolated from native legume species from the Brazilian
97 Cerrado/Caatinga biomes, another major biodiversity hotspot for *Burkholderia*
98 dominated by a distinct legume flora (South American Mimosoideae versus South
99 African Papilionoideae), are genetically different in terms of nodulation genes and
100 appear not be subjected to the same levels of HGT (Bontemps *et al.*, 2010; Bournaud
101 *et al.*, 2013).

102 In contrast to *Burkholderia,* only a handful of studies have focussed on the diversity
103 of South African *Alphaproteobacteria*, such as *Mesorhizobium* (Gerding *et al.*, 2012;
104 Hassen *et al.*, 2012; Kanu & Dakora, 2012), which are largely underestimated in
105 terms of diversity (Lemaire *et al.*, 2015), but co-exist as a dominant group with
106 *Burkholderia* in the acidic and oligotrophic Fynbos soils. The study of Lemaire *et al.*
107 (2015) also demonstrated that most isolated Cape mesorhizobia were distantly related
108 to known reference species. Although HGT have been widely described in
109 *Mesorhizobium* (Kaneko *et al.*, 2000; Sullivan *et al.*, 2002; Nandasena *et al.*, 2006,
110 2007), the occurrence and degree of HGT within the evolution of these putatively new
111 *Mesorhizobium* lineages remains to be tested in the Fynbos.

112 Similar to nodulation and nitrogen fixation genes, the ACC deaminase (*acdS*) gene is
113 prevalent in rhizobia, playing an ecological role for plant growth and nodulation via
114 the reduction of deleterious ACC levels (referred to as 1-aminocyclopropane-1-
115 carboxylate and considered as an ethylene precursor) in plants (Ma *et al.*, 2003, 2004;
116 Glick *et al.*, 2007; Conforte *et al.*, 2010; Nascimento *et al.*, 2012a, 2012c). The *acdS*
117 gene is also located on transmittable genetic elements (symbiosis islands) and has
118 been previously shown to evolve via HGT in diverse bacterial groups, including
119 *Mesorhizobium* (Hontzeas *et al.*, 2005; Blaha *et al.*, 2006; Nascimento *et al.*, 2012b).
120 In Fynbos mesorhizobia, however, the occurrence of the *acdS* gene, the degree of
121 HGT, and its potential use for phylogenetic reconstruction has never been
122 investigated.

123 Homologous recombination is another common driving force for prokaryotic
124 evolution (Didelot & Maiden, 2010), diffusing genetic material or creating new allele
125 combination throughout bacterial populations (Fraser *et al.*, 2007). In rhizobia,
126 various degrees of homologous recombination have been demonstrated within species
127 of *Bradyrhizobium* (Vinuesa *et al.*, 2005, 2008; Tang *et al.*, 2012), *Ensifer* (Silva *et*

128 *al.*, 2007), *Mesorhizobium* (Li *et al.*, 2009) and *Rhizobium* (Tian *et al.*, 2010; Van

129 Cauwenberghe *et al.*, 2014).

130 In an attempt to shed some new light on the evolution of Fynbos rhizobia,

131 phylogenies from both housekeeping and symbiosis genes were reconstructed in order

132 to assess incongruent signals as a result of levels of horizontal gene transfer

133 (transmission of symbiosis genes and homologous recombination of chromosomal

134 genes). By using this retrospective approach (sensu Sørensen *et al.*, 2005) the degree

135 of HGT will be investigated in both alpha- (*Mesorhizobium*) and beta- (*Burkholderia*)

136 rhizobia. The objectives of the study were (i) to examine whether HGT of symbiotic

137 plasmids has occurred among rhizobial lineages of *Mesorhizobium* and *Burkholderia*

138 (incongruences between chromosomal and *nod* sequence data) (ii) to investigate

139 recombination rates between homologues of rhizobial strains (incongruences between

140 housekeeping genes) (iii) to evaluate the prevalence of *acdS* and test whether this

141 accessory gene is prone to HGT among *Mesorhizobium* lineages.

142

Material and Methods

143

Bacterial strains

144

145   A selection of 22 *Burkholderia* and 24 *Mesorhizobium* isolates from phylogenetically

146   diverse lineages was obtained from previous rhizobial screenings and new collections

147   in the Fynbos region (Lemaire *et al.*, 2015) (Table S1). The *Mesorhizobium* isolates

148   originate from diverse host legumes of the tribes Crotalarieae (*Aspalathus*), Genisteae

149   (*Argyrolobium*), Indigofereae (*Indigofera*) and Psoraleeae (*Psoralea*, *Otholobium*),

150   whereas all *Burkholderia* accessions are from *Podalyria calyptrata* populations (tribe

151   Podalyrieae) collected from 14 sites. One *Rhizobium* isolate (accession Dlodlo 49)

152   from the study of Lemaire *et al.* (2015) with a nodulation gene related to

153   *Mesorhizobium* was also included in this study. The *Rhizobium* and all

154   *Mesorhizobium* strains were successfully authenticated on either the original host,

155   siratro or *Otholobium hirtum*, except for the strain isolated from the host species

156   *Aspalathus spicata* Muasya 5440 (Lemaire *et al.*, 2015). For *Burkholderia*, the ability

157   to nodulate was verified on the original host *Podalyria calyptrata* (data not

158   presented).

159

Amplification and phylogenetic analyses

160

161   PCR reactions were performed in a standard 25 µl reaction mixture (Kapa

162   Biosystems), according to the manufacturer's instructions. All PCR amplifications

163   were generated with primers listed in Table S2, following the PCR conditions as

164   described by the authors. Amplified products were purified using the Exo/Sap enzyme

165   cleaning protocol (Werle *et al.*, 1994) and sent to the Macrogen sequence facility

166   (Macrogen, The Netherlands), using the same PCR primers for sequencing. All

167   GenBank accessions numbers are listed in Table S1.

168   Sequence reads were edited, assembled and aligned in Geneious Pro v.5.1.7

169   (http://www.geneious.com). Alignments were subjected to phylogenetic analyses,

170   using Maximum Likelihood (ML) and Bayesian Inference (BI) criteria, both carried

171   out on the CIPRES web portal (http://www.phylog.org). ML analyses were done with

172   RAxML-VI-HPC v.2.2.3 using GTR-GAMMA as the most complex substitution

173   model available (Stamatakis, 2006). A multi-parametric bootstrap resampling of 1000

174   pseudo-replicates was plotted onto the previously selected best-scored ML tree.

175   Model selection for the Bayesian analyses was conducted with MrModeltest v.3.06

176   (Posada & Crandall, 1998) under the Akaike information criterion. For all datasets,

177   MrModeltest selected the general time reversible (GTR) model of DNA substitutions

178   with gamma-distributed rate variation across invariant sites. This best fitting model of

179   DNA substitution was applied for each separate dataset. In the combined BI analyses,

180   the multiple-gene dataset was partitioned and the same models were assigned to

181   separate unlinked partitions. BI analyses were carried out using MrBayes v.3.1

182   (Ronquist & Huelsenbeck, 2003), running four Markov chains (one cold and three

183   heated) simultaneously for five million generations. Conservatively, 25% of the first

184   trees sampled were regarded as 'burnin' and discarded. Convergence of the chains

185   was checked using Tracer v.1.4 (Rambaut & Drummond, 2007).

186

187   *Testing phylogenetic incongruence between chromosomal and nodulation genes*

188   Parallel evolution between chromosomal and plasmid gene (vertical transmission of

189   accessory genes) trees was evaluated with a topology or co-phylogeny mapping

190   method in Jane v.4 (Conow *et al.*, 2010). The degree of congruence between the

191   topologies was assessed by maximizing the number of co-speciation (vertical gene

192   transfer) and minimizing the possible number of non-codiversification events

193   (horizontal gene transfer) under the default setting of event costs. A permutation

194   procedure tested the null hypothesis that two phylogenies are randomly related or that

195   the observed number of co-speciation events of the initial search was not larger than

196   expected by chance alone. The best scoring ML trees were imported as input trees for

197   the reconciliation analyses, comparing the scores of the optimal and initial

198   reconstruction with those of randomly obtained topologies. Randomization tests were

199   ran with 1000 randomly permuted trees and a population size of 100. The cost

200   distribution of random sample solutions and statistical significance was calculated in a

201   cost histogram in Jane.

202

203   *Analysis of recombination*

204   Sequence alignments of the housekeeping genes *recA*, *atpD*, *gyrB* and *glnA* were

205   subjected to ClonalFrame analyses (Didelot & Falush, 2007) to assess the effect of

206   recombination estimated by the r/m (the ratio of probability that a nucleotide will be

207   altered through recombination and point mutations) (Guttman & Dykhuizen, 1994)

208   and the ρ/θ (the frequency of occurrence of recombination relative to point mutations)

209   (Milkman & Bridges, 1990) statistics. Five independent ClonalFrame runs were

210 performed each consisting of 1,000,000 MCMC iterations (25% burn-in), and trees

211 were sampled every 100 iterations, resulting in a sample size of 7500 trees.

212 Convergence of the MCMC was confirmed by the Gelman & Rubin test (Gelman &

213 Rubin, 1992).

214 **Results**

215 *Phylogenetic analyses of individual chromosomal and symbiosis genes*

216 Four chromosomal (16S rRNA – 1360 bp, *atpD* – 650 bp, *recA* – 620 bp, *gyrB* 650

217 bp) and four symbiosis (*nodA* – 590 bp, *nodB* – 250 bp, *nodC* – 570 bp, *nifH* – 310

218 bp) genes were sequenced and analysed for the *Burkholderia* isolates, including

219 reference sequences of chromosomal and plasmid genes of available genomes of *B.*

220 *graminis*, *B. phytofirmans*, *B. rhynchosiae*, *B. tuberum* and *B. xenovorans*.

221 Phylogenetic reconstruction of the separate datasets produced similar groupings

222 among genes of the core genome (Fig. S1A-D), and among symbiosis related genes

223 (Fig. S1E-H). Some discrepancies were detected, but most conflicts were not

224 statistically supported, except for the placement of two taxa in the 16S rRNA (isolate

225 25I3R1 and 23I2R2) and one *nodC* lineage (29I6R2) relative to the housekeeping and

226 nodulation gene trees, respectively. The *nifH* sequences within *Burkholderia* (Fig.

227 S1H) generated mostly an unresolved topology as a result of similar or identical (12

228 out of 21 isolates) amplicons (pairwise genetic similarity > 99%), comprising only 16

229 variable nucleotides in the dataset.

230 Similarly for *Mesorhizobium*, analyses of sequence data of five chromosomal (16S

231 rRNA – 1329 bp, *atpD* – 516 bp, *recA* – 458 bp, *gyrB* - 637 bp, *glnA* – 953 bp) and

232 four symbiosis (*nodA* – 621 bp, *nodB* – 560 bp, *nodC* – 592 bp, *nifH* – 349 bp) genes

233 generated consistent relationships (Fig. S2) with only a few conflicts observed among

234 different chromosomal loci (Fig. S2 A-E), comprising the isolates of *Psoralea*

235 *rigidula* 5343 (16S rRNA), *Aspalathus aurantiaca* 5397, *Psoralea asarina* 15 (*recA*)

236 and *Argyrolobium lunare* 5369 (*gyrB*). Incongruences among nodulation and nitrogen

237 fixation gene trees (Fig. S2G-J) were detected for the *Mesorhizobium* isolates of

238 *Psoralea asarina* 15 (*nodC, nifH*) and *Otholobium hirtum* 5334 (*nodA*).

239 One *Rhizobium* isolate (accession 49) was included in the analyses, previously

240 demonstrated to harbour *nodA* and *nifH* symbiosis genes from a *Mesorhizobium* strain

241 (Lemaire *et al.*, 2015). Sequencing of the *nodB* and *nodC* genes supports the

242 identification of symbiosis genes related to *Mesorhizobium*, indicating horizontal

243 transfer of symbiosis genes across different rhizobial genera

244 (*Rhizobium*/*Mesorhizobium*).

245 Two strongly supported clades were recovered in the *nodA* and *nodC Mesorhizobium*

246 gene trees, largely separating the isolates of *Otholobium* and *Psoralea* (tribe

247 Psoraleeae) from nodule symbionts of the genera *Aspalathus* (tribe Crotalarieae),

248   *Argyrolobium* (tribe Genisteae) and *Indigofera* (tribe Indigofereae). Only one

249   *Aspalathus* symbiont (*Mesorhizobium sp.* 31) was found in the *Otholobium-Psoralea*

250   clade. The grouping of nodulation genes, according to the host was reflected by high

251   sequence divergence in the *nodA* (72.3 % mean sequence similarity) and *nodC* (80.2

252   % mean sequence similarity) datasets. Moreover, for *nodB*, none of the *Psoralea* or

253   *Otholobium* isolates could be amplified (except for *Mesorhizobium* sp. 5462) (Table

254   S1), suggesting that the used primers are not suitable for rhizobia of the tribe

255   Psoraleeae, which have probably too diverged *nodB* genes.

256   In contrast to the symbiosis genes (Fig. S2G-J), which are correlated with the host

257   range, the housekeeping genes (Fig. S2 A-E) of *Mesorhizobium* species differ from

258   the host phylogeny, showing an intermingled pattern of isolates from *Argyrolobium*,

259   *Aspalathus*, *Otholobium* and *Psoralea* legumes. This result indicates that legume

260   species form symbiosis with *Mesorhizobium* lineages with diverse genetic

261   backgrounds.

262

Concatenated sequence analyses

264   Analyses of the concatenated sequences were generally congruent with those of the

265   individual gene trees. Because short gene fragments appear to lack sufficient

266   phylogenetic information to provide well-resolved trees, combination of single genes

267   with unequal evolutionary rates has been recommended to give a more robust

268   evolutionary tree and to level out conflicting signals of homoplastic character states

269   (Gaunt *et al.*, 2001; Gadagkar *et al.*, 2005; Vinuesa *et al.*, 2005; Rivas *et al.*, 2009;

270   Laranjo *et al.*, 2012). In this study, the concatenated sequences of single gene markers

271   of chromosomal (16S rRNA, *recA*, *atpD*, *gyrB*) and nodulation (*nodA*, *nodB*, *nodC*)

272   genes produced robust phylogenies, resolving relationships among most isolates with

273   high support values for both genera of *Burkholderia* and *Mesorhizobium* (Fig. 1-2).

274   Because the 16S rRNA matrix of the *Mesorhizobium* isolates lacks phylogenetic

275   information (pairwise genetic similarity > 99%), comprising only 5 variable

276   nucleotides within 18 out of 24 sequences, we excluded 16S rRNA from the

277   concatenated chromosomal gene analysis (Fig. 2).

278   A considerable number of *Mesorhizobium* (Fig. S3) and *Burkholderia* (Fig. S4)

279   isolates were not related to 16S rRNA sequences of reference type strains, suggesting

280   novel rhizobial species. For *Burkholderia* (Fig. S4), only seven isolates were highly

281   related (> 99% 16S rRNA sequence similarity) to the common South African

282   rhizobial species *B. tuberum* and *B. dilworthii* (Gyaneshwar *et al.*, 2011; De Meyer *et*

283   *al.*, 2014), and the diazotrophic species *B. xenovorans*, *B. sediminicola,* which were

284   previously isolated from Fynbos legumes (Beukes *et al.*, 2013). One *Burkholderia*

285   isolate was related to *B. sartisoli*, which has never been isolated from root nodules.

286   Similarly for *Mesorhizobium* (Fig. S3) we showed that only one strain was

287   conspecific to *Mesorhizobium chacoense* (symbionts of *Otholobium bracteolatum*

288   42), placing the remaining isolates in distinct phylogenetic clades unrelated to any

289   reference *Mesorhizobium* species. These mesorhizobia isolates are most likely new

290   species as previously suggested by Lemaire *et al.* (2015).

291

Incongruence between chromosomal and symbiosis genes

293   Phylogenetic relationships of the chromosomal genes were largely inconsistent with

294   the nodulation genes for both *Burkholderia* (Fig. 1) and *Mesorhizobium* (Fig. 2).

295   Visual inspection of the combined trees of chromosomal versus nodulation genes

296   revealed congruent relationships for only three *Burkholderia* clades (clade 1:

297   16I4R2/13R2/*B. tuberum*, clade 2: 25I3R6/9I2R2, clade 3:

298   25I1R1/18I8R3/6665CI2R2) and one *Mesorhizobium* grouping (*Aspalathus ericifolia*

299   31/*Otholobium hirtum* 32). To estimate the degree of parallel evolution between

300   chromosomal and nodulation genes, a reconciliation analyses was performed for the

301   *Burkholderia* and *Mesorhizobium* datasets, mapping the nodulation gene tree on the

302   chromosomal gene phylogeny (Fig. S5). The co-divergence approach estimates the

303   maximum number of co-speciation events and visualizes all solutions by introducing

304   a minimum number of non-co-speciation events (duplication, host-switch and sorting

305   events) on the nodulation gene tree. For *Burkholderia*, 8336 equal cost solutions were

306   recovered with co-speciation events ranging between seven and nine, and a total cost

307   of 36 for 18 events of host switches/duplications. A similar degree of non-parallel

308   evolution was also observed for the *Mesorhizobium* analyses, revealing 6432 equal

309   cost solutions with six co-speciation events and 22 lateral transfers/duplications (cost

310   = 44) (Fig. S5). Topological congruence (vertical inheritance/parallel evolution) was

311   further statistically examined with a randomization tests (Fig. S6), providing evidence

312   to reject the null hypothesis of random relationships for both gene phylogenies ($P <$

313   0.01). Despite large-scale symbiosis-gene transfers, the overall chromosomal and

314   symbiosis topology shares a significant number of co-divergence, indicating that

315 events of parallel evolution occurs more frequently than we would expect purely by

316 chance.

317

318 **Recombination in *Mesorhizobium* and *Burkholderia***

319 The frequency and relative impact of recombination on the evolution of housekeeping

320 genes was assessed using the ClonalFrame approach (Didelot & Falush, 2007).

321 Recombination frequencies were estimated for the *Mesorhizobium* and *Burkholderia*

322 datasets, comprising similar levels of genetic divergence with the lowest sequence

323 similarity of 90% for both rhizobial groups. For the *Mesorhizobium* dataset (n = 28

324 strains), the r/m and ρ/θ values were 11.81 (7.52, 16.01; 95% CI) and 2.62 (1.73, 2.91;

325 95% CI), respectively, implying recombination rather than mutations as predominant

326 contribution to the evolution among the tested regions of the *Mesorhizobium* strains.

327 Similarly, substantial levels of recombination were observed among the *Burkholderia*

328 strains (n = 28), with r/m = 7.80 (2.47, 16.42; 95% CI) and ρ/θ = 2.18 (0.68, 3.76;

329 95% CI).

330

331 ***Phylogenetic analysis of the* acdS *gene***

332 The presence of *acdS* in one *Rhizobium* (49), one *Burkholderia* (25I3R1) and all

333 *Mesorhizobium* isolates was confirmed by sequence analyses. Phylogenetic analyses

334 of the *acdS* isolates and closely related reference sequences of *Mesorhizobium*,

335 *Rhizobium* and *Burkholderia*, placed all *Mesorhizobium* isolates within a

336 monophyletic group (100% Bayesian support value - BS, 100% Bootstrap support

337 value - BSS) as a sister group to *Mesorhizobium chacoense* (98% BS, 90 BSS) (Fig.

338 3, S7). The sequence divergence between the *acdS Mesorhizobium* isolates (mean

339 pairwise sequence similarity > 95%) generated well-resolved relationships with high

340 support values for most nodes. The *acdS* gene phylogeny revealed significant

341 incongruent groupings in comparison to both chromosomal and symbiosis-related

342 genes, indicating a different evolutionary history prone to HGT. Concordant

343 relationships between the *acdS* tree (Fig. S2F) and the chromosomal gene trees (Figs.

344 S2A-E) were only detected among a few sister group relationships (e.g.

345 *Mesorhizobium* spp. 31/32, *Mesorhizobium* spp. 5382/5343, *Mesorhizobium* spp.

346 5361/5357, *Mesorhizobium* spp. 5378/5334). One similar relationship

347 (*Mesorhizobium* sp. 31R1-31R2) was observed between the *acdS* and nodulation gene

348 trees (Fig. 2-3).

349 The *acdS* sequence of the *Rhizobium* isolate was placed as a sister group to the

350 *Mesorhizobium* spp., although this relationships is not supported (Fig. 3, S7).

351 Sequence analysis of the *Rhizobium* isolates detected low levels of similarity among

352 available reference strains, showing the highest similarity value with *R. gallicum*

353 (83.6 % sequence similarity).

354 Amplification of the *acdS* region in *Burkholderia* resulted in only one amplicon,

355 suggesting that the primers originally designed for *Mesorhizobium*

356 (*Alphaproteobacteria*) (Nascimento *et al.*, 2012b) are not suitable for *Burkholderia*

357 genus (*Betaproteobacteria*) due to primer mismatches. The *acdS* sequence of

358 *Burkholderia* was placed as a sister group to the reference strain *B. ginsengisoli*

359 NBRC100965 (100% BS, 100 BSS). This species was originally recovered from the

360 rhizosphere of plants from ginseng field (Kim *et al.*, 2006).

361

362 **Discussion**

363 *Horizontal gene transfer of symbiosis genes among rhizobia of the Fynbos biome*

364 Extensive incongruence between phylogenies of nodulation and chromosomal genes

365 of members of the genera *Burkholderia* and *Mesorhizobium* provides evidence for

366 frequent exchange of nodulation genes among rhizobial lineages of the South African

367 Cape Fynbos biome. Within the *Burkholderia* species of the Cape, horizontal gene

368 transfer of symbiosis genes has been previously suggested to explain discordant

369 relationships between the nodulation (*nodA*) gene in comparison to chromosomal

370 genes (Beukes *et al.*, 2013). The observed exchange of nodulation genes located on

371 plasmids and symbiosis islands in *Burkholderia* and *Mesorhizobium*, respectively,

372 indicates that HGT is not restricted to one rhizobial group (i.e. *Burkholderia*), but also

373 occurs among rhizobia of the *Alphaproteobacteria*, suggesting HGT as a common

374 feature in the Fynbos biome. This observation is in contrast with rare events of HGT

375 of symbiosis genes among South American *Burkholderia,* which are associated with

376 *Mimosa* spp. mostly endemic to the Cerrado and Caatinga biomes of Brazil

377 (Bontemps *et al.*, 2010; Mishra *et al.*, 2012). Recently, alpha- and beta-rhizobia of

378 Mexican *Mimosa* spp. were also characterized without an indication of gene exchange

379 of nodulation genes (Bontemps *et al.*, 2015).

380 The symbiotic nodulation genes, which are involved in host recognition by the

381 synthesis of signalling molecules (Nod factors), are expected to evolve under

382 constraints imposed by the interaction with the host plant (Perret *et al.*, 2000; Spaink,

383 2000). Hence the evolutionary history of nodulation genes is usually related to the

384 host plant (Haukka *et al.*, 1998; Laguerre *et al.*, 2001; Suominen *et al.*, 2001; Lu *et

385 al.*, 2009). In Fynbos mesorhizobia, two distinct symbiosis clades (*nodA-nodB-nodC-

386 nifH*) were recovered, largely grouped by the host tribes Psoraleeae and

387 Crotalarieae/Genisteae (Fig. 2); all symbionts of the legumes of *Otholobium* and

388 *Psoralea* (tribe Psoraleeae) were clustered within a monophyletic group, while all

389 *Aspalathus* (tribe Crotalarieae) and *Argyrolobium* (tribe Genisteae) symbionts were

390 placed in a clade with distinct nodulation and nitrogen fixation genes, except for one

391 *Aspalathus* symbiont (accession 31). While the *nod* gene phylogenies of mesorhizobia

392 are strongly aligned with the host, at least at tribal (but not generic) level, it does not

393 explain the complex evolutionary history of nodulation genes for *Burkholderia*, which

394 were all originally isolated from the same host *Podalyria calyptrata*. In previous

395 studies, the association between *Burkholderia* isolates and the host was not strong,

showing one *Burkholderia* species nodulating diverse host species from different legume tribes and genera (Beukes *et al.*, 2013). Host range studies confirmed the aspecificity of the *Burkholderia*-legume interaction, showing one rhizobial strain able to form effective nodules in various legume species (Gyaneshwar *et al.*, 2011; Liu *et al.*, 2012; Angus *et al.*, 2013; Sprent *et al.*, 2013). This observation may indicate that South African legumes do not have stringent requirements for a particular *Burkholderia* genotype and allow relaxed co-evolution between the symbiotic partners.

While *Burkholderia* seems to have a broad host-range with local papilionoid species they appear incapable of nodulating South American mimosoid hosts (Gyaneshwar *et al.*, 2011). Interestingly, mimosoid-nodulating *Burkholderia* from the South Americas exhibit a broader host range, which are able to form interactions with papilionoid species (Martínez-Romero, 2009; Talbi *et al.*, 2010; Gyaneshwar *et al.*, 2011; Liu *et al.*, 2014; Moulin *et al.*, 2014). The naturally broader host range of these *Burkholderia* species (e.g. *B. phymatum*) and consequently the low pressure of the bacterial symbiont to adapt to a legume host by the exchange of symbiosis-specific genes (Segovia *et al.*, 1991) might explain the relative lack of HGT observed in South American *Mimosa* symbionts (Bontemps *et al.*, 2010, 2015; Mishra *et al.*, 2012).

In the Fynbos biome, the lateral transfer of nodulation genes in *Burkholderia* might also be the result of other factors, such as the flexibility and adaptability of the genome to highly diverse ecological environments (Miché *et al.*, 2002). Rhizobial populations seem to interact reciprocally by exchanging symbiotic elements, comprising genes related to nodulation, nitrogen fixation, auxin synthesis, hydrogenase components and ACC deaminase activity (de Oliveira Cunha *et al.*, 2012; Zuleta *et al.*, 2014), in order to respond to highly diverse and changeable environments, and extend their capacity to colonize new habitats, which allow the host plants to associate with the most adapted rhizobia to the environment (Suominen *et al.*, 2001; Moulin *et al.*, 2004; Vinuesa *et al.*, 2005; Zhao *et al.*, 2008).

Role of recombination in Fynbos rhizobia

The rates of recombination relative to those of mutation showed similar results of recombination for *Mesorhizobium* and *Burkholderia* strains, indicating a high impact of homologous recombination or low mutation rates. The ratio of the probabilities that a given nucleotide is changed by recombination or mutation (r/m) is roughly eleven

and seven for the *Mesorhizobium* and *Burkholderia* isolates, respectively. Although similar high values of recombination relative to mutation (r/m = 2-10, *sensu* Vos & Didelot, 2009) have been recorded in many rhizobial groups (e.g. Tian *et al.*, 2010, 2012; Van Cauwenberghe *et al.*, 2014), the observation of recombination among distinct *Mesorhizobium* (*M. ciceri*, *M. loti*, *M. huakuii*) and *Burkholderia* species (including at least five species) is remarkable, because the success rate of exchange of homologous genetic material decreases exponentially with the genetic distance of interacting species (Majewski, 2001). Consequently, high rates of recombination occur more frequently between close relatives than among divergent organisms (Didelot & Maiden, 2010; Popa *et al.*, 2011). Nevertheless, events of recombination across bacterial divisions and domains have been reported (Garcia-Vallve *et al.*, 2000; Rest & Mindell, 2003).

Why these *Mesorhizobium* and *Burkholderia* are shuffling around alleles by homologous recombination is still an open question. Although *Mesorhizobium* and *Burkholderia* are phylogenetically distinct (alpha- and beta-subclass of *Proteobacteria*), similar recombination rates in both rhizobial groups could imply that events of recombination are more related with comparable ecologies rather than to genetic background (Wiedenbeck & Cohan, 2011). Although speculative, the considerable level of recombination (and gene movement of symbiosis genes; see above) in *Burkholderia,* and linked to its renowned genomic plasticity (Miché *et al.*, 2002; Vial *et al.*, 2007), is not evolutionary constrained to beta-rhizobia, but is a common feature in the Fynbos, occurring among mesorhizobia adapted to the same ecological environment.

*Evolution, occurrence and ecological significance of the ACC deaminase (*acdS*) gene*

The location of the *acdS* gene varies in different species but is often located on transferable elements such as plasmids in *Rhizobium* and *Ensifer* (Ma *et al.*, 2003; Young *et al.*, 2006; Kuhn *et al.*, 2008) and symbiosis islands in *Mesorhizobium* (Sullivan *et al.*, 2002; Nascimento *et al.*, 2012b). In *Burkholderia*, analyses of genome data identified *acdS* on the chromosome, except for *B. phymatum* STM815[T] and *B. phenoliruptrix* BR3459a, which have two copies of the *acdS* gene, one on the chromosome and the other on the plasmid (Nascimento *et al.*, 2014). Despite the variation of the position on transmittable elements, the *acdS* gene is expected to

464 evolve mainly through HGT, at least for species having *acdS* on plasmids and
465 symbiosis islands, as previously demonstrated in phylogenetic studies of both *Alpha-*
466 and *Betaproteobacteria* (Hontzeas *et al.*, 2005; Blaha *et al.*, 2006; Glick *et al.*, 2007;
467 Nandasena *et al.*, 2007; Nascimento *et al.*, 2012b).

468 In *Mesorhizobium*, *acdS* has been reported in many species, which are shown to be
469 prone to HGT, most likely through symbiotic island exchange (Nascimento *et al.*,
470 2012b; Laranjo *et al.*, 2014). In the study of Nascimento *et al.* (2012b), the *acdS* tree
471 revealed similar relationships in comparison to the symbiosis gene trees and correlates
472 well with the host range, rather than the 16S rRNA phylogeny. In the current study,
473 few congruent relationships were observed between the *acdS*, housekeeping and
474 nodulation gene trees, indicating that ACC deaminase genes of these South African
475 *Proteobacteria* are extensively subjected to HGT with genes on transmittable
476 elements (i.e. plasmids and symbiotic islands) being prone to such different
477 evolutionary histories. Future genome studies are needed to investigate the genome
478 characteristics and the exact location of the *acdS* gene within the multipartite genome;
479 a genome arrangement prevalent among plant-associated symbionts (Harrison *et al.*,
480 2010; Landeta *et al.*, 2011). It is also important to note that the genes located on
481 accessory replicons or smaller chromosomes may evolve at a higher substitution rate
482 compared to genes present within the larger primary chromosomes (Cooper *et al.*,
483 2010; MacLean *et al.*, 2014) and hence may consequently affect the inference of
484 phylogenetic relationships between different sets of genes (i.e. housekeeping,
485 nodulation, *acdS*).

486 The ecological significance of the microbial ACC deaminase activity to stimulate
487 plant growth (Glick *et al.*, 2007) and the observed prevalence of the ACC deaminase
488 gene throughout all *Mesorhizobium* spp. indicate that this enzyme is playing an
489 important role in the nodulation process of these strains by increasing their ecological
490 competitiveness and symbiotic performance (Ma *et al.*, 2003, 2004; Uchiumi *et al.*,
491 2004; Nascimento *et al.*, 2012b, 2012c; Brígido *et al.*, 2013). The presence of *acdS*
492 genes in all *Mesorhizobium* and one *Rhizobium* and *Burkholderia* strain, originating
493 from different geographical locations and diverse legume groups of the tribes
494 Crotalarieae, Genisteae, Podalyrieae and Psoraleeae, indicates that ACC deaminase is
495 a common and important plant-beneficial property among Fynbos rhizobia,
496 particularly for lineages of the genus *Mesorhizobium*.

497

498 *Conclusion*

499 In this multilocus sequence analysis, we provided phylogenetic evidence for

500 horizontal transfer of plasmid located genes within species of *Burkholderia* and

501 *Mesorhizobium,* and extensive exchange of housekeeping genes through homologous

502 recombination. No evidence of HGT between alpha- and beta-rhizobia was observed.

503 The dynamic nature of gene transfer and acquisition observed in selected 'core' and

504 'accessory' genes among *Burkholderia* and *Mesorhizobium* in the Fynbos biome is

505 most likely only the tip of the iceberg, and future genomic work is necessary to reveal

506 the true extent of the migratory lifestyle of (accessory) genes among rhizobia of the

507 Fynbos biome.

508

References

Amadou C, Pascal G, Mangenot S, *et al.* (2008) Genome sequence of the beta-rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia. *Genome Res.* **18**: 1472-1483.

Andam CP, Mondo SJ & Parker MA (2007) Monophyly of *nodA* and *nifH* genes across Texan and Costa Rican populations of *Cupriavidus* nodule symbionts. *Appl. Environ. Microbiol.* **73**: 4686-4690.

Angus AA, Lee A, Lum MR, Shehayeb M, Hessabi R, Fujishige NA, Yerrapragada S, Kano S, Song N & Yang P (2013) Nodulation and effective nitrogen fixation of *Macroptilium atropurpureum* (siratro) by *Burkholderia tuberum*, a nodulating and plant growth promoting beta-proteobacterium, are influenced by environmental factors. *Plant Soil* **369**: 543-562.

Barcellos FG, Menna P, da Silva Batista JS & Hungria M (2007) Evidence of horizontal transfer of symbiotic genes from a *Bradyrhizobium japonicum* inoculant strain to indigenous diazotrophs *Sinorhizobium* (*Ensifer*) *fredii* and *Bradyrhizobium elkanii* in a Brazilian Savannah soil. *Appl. Environ. Microbiol.* **73**: 2635-2643.

Beukes CW, Venter SN, Law IJ, Phalane FL & Steenkamp ET (2013) South African papilionoid legumes are nodulated by diverse *Burkholderia* with unique nodulation and nitrogen-fixation loci. *PLoS One* **8**: e68406.

Blaha D, Prigent-Combaret C, Mirza MS & Moënne-Loccoz Y (2006) Phylogeny of the 1-aminocyclopropane-1-carboxylic acid deaminase-encoding gene *acdS* in phytobeneficial and pathogenic Proteobacteria and relation with strain biogeography. *FEMS Microbiol. Ecol.* **56**: 455-470.

Bontemps C, Elliott GN, Simon MF, *et al.* (2010) *Burkholderia* species are ancient symbionts of legumes. *Mol. Ecol.* **19**: 44-52.

Bontemps C, Rogel MA, Wiechmann A, *et al.* (2015) Endemic *Mimosa* species from Mexico prefer alphaproteobacterial rhizobial symbionts
. *New Phytol.*

Boto L (2010) Horizontal gene transfer in evolution: facts and challenges. *Proc. R. Soc. Lond. B. Biol. Sci.* **277**: 819-827.

Bournaud C, de Faria SM, dos Santos JMF, Tisseyre P, Silva M, Chaintreuil C, Gross E, James EK, Prin Y & Moulin L (2013) *Burkholderia* species are the most common and preferred nodulating symbionts of the *Piptadenia* group (tribe Mimoseae). *PLoS One* **8**: e63478.

Brígido C, Nascimento FX, Duan J, Glick BR & Oliveira S (2013) Expression of an exogenous 1-aminocyclopropane-1-carboxylate deaminase gene in *Mesorhizobium* spp. reduces the negative effects of salt stress in chickpea. *FEMS Microbiol. Lett.* **349**: 46-53.

553   Chen W-M, Moulin L, Bontemps C, Vandamme P, Béna G & Boivin-Masson
554 C (2003) Legume symbiotic nitrogen fixation by beta-Proteobacteria is widespread in
555 nature. *J. Bacteriol.* **185**: 7266-7272.

556   Conforte VP, Echeverria M, Sánchez C, Ugalde RA, Menéndez AB & Lepek
557 VC (2010) Engineered ACC deaminase-expressing free-living cells of *Mesorhizobium*
558 *loti* show increased nodulation efficiency and competitiveness on *Lotus* spp. *J. Gen.*
559 *Appl. Microbiol.* **56**: 331-338.

560   Conow C, Fielder D, Ovadia Y & Libeskind-Hadas R (2010) Jane: a new tool
561 for the cophylogeny reconstruction problem. *Algorithms Mol. Biol.* **5**: 16.

562   Cooper VS, Vohr SH, Wrocklage SC & Hatcher PJ (2010) Why genes evolve
563 faster on secondary chromosomes in bacteria.

564   De Meyer SE, Cnockaert M, Ardley JK, van Wyk B-E, Vandamme PA &
565 Howieson JG (2014) *Burkholderia dilworthii* sp. nov. isolated from *Lebeckia*
566 *ambigua* root nodules from South Africa. *Int. J. Syst. Evol. Microbiol.* **64**: 1090-1095.

567   de Oliveira Cunha C, Zuleta LFG, de Almeida LGP, Ciapina LP, Borges WL,
568 Pitard RM, Baldani JI, Straliotto R, de Faria SM & Hungria M (2012) Complete
569 genome sequence of *Burkholderia phenoliruptrix* BR3459a (CLA1), a heat-tolerant,
570 nitrogen-fixing symbiont of *Mimosa flocculosa*. *J. Bacteriol.* **194**: 6675-6676.

571   Didelot X & Falush D (2007) Inference of bacterial microevolution using
572 multilocus sequence data. *Genetics* **175**: 1251-1266.

573   Didelot X & Maiden MCJ (2010) Impact of recombination on bacterial
574 evolution. *Trends Microbiol.* **18**: 315-322.

575   Elliott GN, Chen W-M, Chou J-H, *et al.* (2007b) *Burkholderia phymatum* is a
576 highly effective nitrogen-fixing symbiont of *Mimosa* spp. and fixes nitrogen *ex*
577 *planta*. *New Phytol.* **173**: 168-180.

578   Finan TM (2002) Evolving insights: symbiosis islands and horizontal gene
579 transfer. *J. Bacteriol.* **184**: 2855-2856.

580   Fraser C, Hanage WP & Spratt BG (2007) Recombination and the nature of
581 bacterial speciation. *Science* **315**: 476-480.

582   Gadagkar SR, Rosenberg MS & Kumar S (2005) Inferring species
583 phylogenies from multiple genes: concatenated sequence tree versus consensus gene
584 tree. *J. Exp. Zool. B Mol. Dev. Evol.* **304**: 64-74.

585   Garau G, Yates RJ, Deiana P & Howieson JG (2009) Novel strains of
586 nodulating *Burkholderia* have a role in nitrogen fixation with papilionoid herbaceous
587 legumes adapted to acid, infertile soils. *Soil Biol. Biochem.* **41**: 125-134.

588   Garcia-Vallve S, Romeu A & Palau J (2000) Horizontal gene transfer of
589 glycosyl hydrolases of the rumen fungi. *Mol. Biol. Evol.* **17**: 352-361.

590   Gaunt MW, Turner SL, Rigottier-Gois L, Lloyd-Macgilp SA & Young JP
591   (2001) Phylogenies of *atpD* and *recA* support the small subunit rRNA-based
592   classification of rhizobia. *Int. J. Syst. Evol. Microbiol.* **51**: 2037-2048.

593   Gelman A & Rubin DB (1992) Inference from iterative simulation using
594   multiple sequences. *Statistical Science* **7**: 457-472.

595   Gerding M, O'Hara GW, Bräu L, Nandasena K & Howieson JG (2012)
596   Diverse *Mesorhizobium* spp. with unique *nodA* nodulating the South African legume
597   species of the genus *Lessertia. Plant Soil* **358**: 385-401.

598   Glick BR, Todorovic B, Czarny J, Cheng Z, Duan J & McConkey B (2007)
599   Promotion of plant growth by bacterial ACC deaminase. *Crit. Rev. Plant Sci.* **26**: 227-
600   242.

601   Goldenfeld N & Woese C (2007) Biology's next revolution. *Nature* **445**: 369-
602   369.

603   Guttman DS & Dykhuizen DE (1994) Clonal divergence in *Escherichia coli*
604   as a result of recombination, not mutation. *Science* **266**: 1380-1383.

605   Gyaneshwar P, Hirsch AM, Moulin L*, et al.* (2011) Legume-nodulating
606   betaproteobacteria: diversity, host range, and future prospects. *Mol. Plant. Microbe*
607   *Interact.* **24**: 1276-1288.

608   Harrison PW, Lower RPJ, Kim NKD & Young JPW (2010) Introducing the
609   bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiol.* **18**: 141-
610   148.

611   Hassen AI, Bopape FL, Habig J & Lamprecht SC (2012) Nodulation of
612   rooibos (*Aspalathus linearis* Burm. f.), an indigenous South African legume, by
613   members of both the α-Proteobacteria and β-Proteobacteria. *Biol. Fertil. Soils* **48**:
614   295-303.

615   Haukka K, Lindström K & Young JPW (1998) Three phylogenetic groups of
616   *nodA* and *nifH* genes in *Sinorhizobium* and *Mesorhizobium* isolates from leguminous
617   trees growing in Africa and Latin America. *Appl. Environ. Microbiol.* **64**: 419-426.

618   Hontzeas N, Richardson AO, Belimov A, Safronova V, Abu-Omar MM &
619   Glick BR (2005) Evidence for horizontal transfer of 1-aminocyclopropane-1-
620   carboxylate deaminase genes. *Appl. Environ. Microbiol.* **71**: 7556-7558.

621   Howieson JG, De Meyer SE, Vivas-Marfisi A, Ratnayake S, Ardley JK &
622   Yates RJ (2013) Novel *Burkholderia* bacteria isolated from *Lebeckia ambigua*, a
623   perennial suffrutescent legume of the fynbos. *Soil Biol. Biochem.* **60**: 55-64.

624   Jain R, Rivera MC, Moore JE & Lake JA (2003) Horizontal gene transfer
625   accelerates genome innovation and evolution. *Mol. Biol. Evol.* **20**: 1598-1602.

626   Kaneko T, Nakamura Y, Sato S, Asamizu E, Kato T, Sasamoto S, Watanabe
627   A, Idesawa K, Ishikawa A & Kawashima K (2000) Complete genome structure of the
628   nitrogen-fixing symbiotic bacterium *Mesorhizobium loti. DNA Res.* **7**: 331-338.

629   Kanu SA & Dakora FD (2012) Symbiotic nitrogen contribution and
630 biodiversity of root-nodule bacteria nodulating *Psoralea* species in the Cape Fynbos,
631 South Africa. *Soil Biol. Biochem.* **54**: 68-76.

632   Kim H-B, Park M-J, Yang H-C, An D-S, Jin H-Z & Yang D-C (2006)
633 *Burkholderia ginsengisoli* sp. nov., a β-glucosidase-producing bacterium isolated
634 from soil of a ginseng field. *Int. J. Syst. Evol. Microbiol.* **56**: 2529-2533.

635   Kock MM (2004) Diversity of root nodulating bacteria associated with
636 *Cyclopia* species. Thesis, PhD Thesis. University of Pretoria., Pretoria.

637   Kuhn S, Stiens M, Pühler A & Schlüter A (2008) Prevalence of pSmeSM11a-
638 like plasmids in indigenous *Sinorhizobium meliloti* strains isolated in the course of a
639 field release experiment with genetically modified *S. meliloti* strains. *FEMS*
640 *Microbiol. Ecol.* **63**: 118-131.

641   Laguerre G, Nour SM, Macheret V, Sanjuan J, Drouin P & Amarger N (2001)
642 Classification of rhizobia based on *nodC* and *nifH* gene analysis reveals a close
643 phylogenetic relationship among *Phaseolus vulgaris* symbionts. *Microbiology* **147**:
644 981-993.

645   Landeta C, Dávalos A, Cevallos MÁ, Geiger O, Brom S & Romero D (2011)
646 Plasmids with a chromosome-like role in rhizobia. *J. Bacteriol.* **193**: 1317-1326.

647   Laranjo M, Young JPW & Oliveira S (2012) Multilocus sequence analysis
648 reveals multiple symbiovars within *Mesorhizobium* species. *Syst. Appl. Microbiol.* **35**:
649 359-367.

650   Laranjo M, Alexandre A & Oliveira S (2014) Legume growth-promoting
651 rhizobia: an overview on the *Mesorhizobium* genus. *Microbiol. Res.* **169**: 2-17.

652   Lemaire B, Dlodlo O, Chimphango S, *et al.* (2015) Symbiotic diversity,
653 specificity and distribution of rhizobia in native legumes of the core Cape Subregion
654 (South Africa). *FEMS Microbiol. Ecol.* **91**: 1-17.

655   Li QF, Zhang XP, Zou L, Chen Q, Fewer DP & Lindstrom K (2009)
656 Horizontal gene transfer and recombination shape mesorhizobial populations in the
657 gene center of the host plants *Astragalus luteolus* and *Astragalus ernestii* in Sichuan,
658 China. *FEMS Microbiol. Ecol.* **70**: 227-235.

659   Liu WYY, Ridgway HJ, James TK, James EK, Chen W-M, Sprent JI, Young
660 JPW & Andrews M (2014) *Burkholderia* sp. induces functional nodules on the South
661 African invasive legume *Dipogon lignosus* (Phaseoleae) in New Zealand soils.
662 *Microb. Ecol.* **68**: 542-555.

663   Liu X, Wei S, Wang F, James EK, Guo X, Zagar C, Xia LG, Dong X & Wang
664 YP (2012) *Burkholderia* and *Cupriavidus* spp. are the preferred symbionts of *Mimosa*
665 spp. in Southern China. *FEMS Microbiol. Ecol.* **80**: 417-426.

666   Lu YL, Chen WF, Wang ET, Guan SH, Yan XR & Chen WX (2009) Genetic
667 diversity and biogeography of rhizobia associated with *Caragana* species in three
668 ecological regions of China. *Syst. Appl. Microbiol.* **32**: 351-361.

669   Ma W, Guinel FC & Glick BR (2003) *Rhizobium leguminosarum* biovar
670 *viciae* 1-aminocyclopropane-1-carboxylate deaminase promotes nodulation of pea
671 plants. *Appl. Environ. Microbiol.* **69**: 4396-4402.

672   Ma W, Charles TC & Glick BR (2004) Expression of an exogenous 1-
673 aminocyclopropane-1-carboxylate deaminase gene in *Sinorhizobium meliloti*
674 increases its ability to nodulate alfalfa. *Appl. Environ. Microbiol.* **70**: 5891-5897.

675   MacLean AM, Milunovic B, Golding GB & Finan TM (2014) Examination of
676 prokaryotic multipartite genome evolution through experimental genome reduction.

677   Majewski J (2001) Sexual isolation in bacteria. *FEMS Microbiol. Lett.* **199**:
678 161-169.

679   Martínez-Romero E (2009) Coevolution in *Rhizobium*-legume symbiosis?
680 *DNA Cell Biol.* **28**: 361-370.

681   Masson-Boivin C, Giraud E, Perret X & Batut J (2009) Establishing nitrogen-
682 fixing symbiosis with legumes: how many rhizobium recipes? *Trends Microbiol.* **17**:
683 458-466.

684   Miché L, Faure D, Blot M, Cabanne-Giuli E & Balandreau J (2002) Detection
685 and activity of insertion sequences in environmental strains of *Burkholderia*. *Environ.*
686 *Microbiol.* **3**: 766-773.

687   Milkman R & Bridges MM (1990) Molecular evolution of the *Escherichia*
688 *coli* chromosome. III. Clonal frames. *Genetics* **126**: 505-517.

689   Mishra RPN, Tisseyre P, Melkonian R, Chaintreuil C, Miche L, Klonowska A,
690 Gonzalez S, Bena G, Laguerre G & Moulin L (2012) Genetic diversity of *Mimosa*
691 *pudica* rhizobial symbionts in soils of French Guiana: investigating the origin and
692 diversity of *Burkholderia phymatum* and other beta-rhizobia. *FEMS Microbiol. Ecol.*
693 **79**: 487-503.

694   Moulin L, Béna G, Boivin-Masson C & Stępkowski T (2004) Phylogenetic
695 analyses of symbiotic nodulation genes support vertical and lateral gene co-transfer
696 within the *Bradyrhizobium* genus. *Mol. Phylogenet. Evol.* **30**: 720-732.

697   Moulin L, Klonowska A, Caroline B, Booth K, Vriezen JAC, Melkonian R,
698 James EK, Young JPW, Bena G & Hauser L (2014) Complete Genome sequence of
699 *Burkholderia phymatum* STM815T, a broad host range and efficient nitrogen-fixing
700 symbiont of *Mimosa* species. *Standards in Genomic Sciences* **9**: 763.

701   Nandasena KG, O'Hara GW, Tiwari RP & Howieson JG (2006) Rapid in situ
702 evolution of nodulating strains for *Biserrula pelecinus* L. through lateral transfer of a
703 symbiosis island from the original mesorhizobial inoculant. *Appl. Environ. Microbiol.*
704 **72**: 7365-7367.

705   Nandasena KG, O'Hara GW, Tiwari RP, Sezmiş E & Howieson JG (2007) In
706 situ lateral transfer of symbiosis islands results in rapid evolution of diverse
707 competitive strains of mesorhizobia suboptimal in symbiotic nitrogen fixation on the
708 pasture legume *Biserrula pelecinus* L. *Environ. Microbiol.* **9**: 2496-2511.

709   Nascimento F, Brígido C, Alho L, Glick BR & Oliveira S (2012a) Enhanced
710 chickpea growth-promotion ability of a *Mesorhizobium* strain expressing an
711 exogenous ACC deaminase gene. *Plant Soil* **353**: 221-230.

712   Nascimento FX, Brígido C, Glick BR & Oliveira S (2012b) ACC deaminase
713 genes are conserved among *Mesorhizobium* species able to nodulate the same host
714 plant. *FEMS Microbiol. Lett.* **336**: 26-37.

715   Nascimento FX, Brígido C, Glick BR, Oliveira S & Alho L (2012c)
716 *Mesorhizobium ciceri* LMS-1 expressing an exogenous ACC deaminase increases its
717 nodulation abilities and chickpea plant resistance to soil constraints. *Lett. Appl.*
718 *Microbiol.* **55**: 15-21.

719   Nascimento FX, Rossi MJ, Soares CRFS, McConkey BJ & Glick BR (2014)
720 New insights into 1-aminocyclopropane-1-carboxylate (ACC) deaminase phylogeny,
721 evolution and ecological significance. *PLoS One* **9**: e99168.

722   Ochman H, Lawrence JG & Groisman EA (2000) Lateral gene transfer and the
723 nature of bacterial innovation. *Nature* **405**: 299-304.

724   Ormeno-Orrillo E, Vinuesa P, Zuniga-Davila D & Martínez-Romero E (2006)
725 Molecular diversity of native bradyrhizobia isolated from Lima bean (*Phaseolus*
726 *lunatus* L.) in Peru. *Syst. Appl. Microbiol.* **29**: 253-262.

727   Ormeño-Orrillo E, Servín-Garcidueñas LE, Imperial J, Rey L, Ruiz-Argueso T
728 & Martinez-Romero E (2013) Phylogenetic evidence of the transfer of *nodZ* and *nolL*
729 genes from *Bradyrhizobium* to other rhizobia. *Mol. Phylogenet. Evol.* **67**: 626-630.

730   Perret X, Staehelin C & Broughton WJ (2000) Molecular basis of symbiotic
731 promiscuity. *Microbiol. Mol. Biol. Rev.* **64**: 180-201.

732   Popa O, Hazkani-Covo E, Landan G, Martin W & Dagan T (2011) Directed
733 networks reveal genomic barriers and DNA repair bypasses to lateral gene transfer
734 among prokaryotes. *Genome Res.* **21**: 599-609.

735   Posada D & Crandall KA (1998) Modeltest: testing the model of DNA
736 substitution. *Bioinformatics* **14**: 817-817.

737   Preston GM, Haubold B & Rainey PB (1998) Bacterial genomics and
738 adaptation to life on plants: implications for the evolution of pathogenicity and
739 symbiosis. *Curr. Opin. Microbiol.* **1**: 589-597.

740   Rambaut A & Drummond AJ (2007) *Tracer v1. 4, Available from*
741 http://beast.bio.ed.ac.uk/Tracer.

742   Rest JS & Mindell DP (2003) Retroids in archaea: phylogeny and lateral
743 origins. *Mol. Biol. Evol.* **20**: 1134-1142.

744   Rivas R, Martens M, De Lajudie P & Willems A (2009) Multilocus sequence
745 analysis of the genus *Bradyrhizobium*. *Syst. Appl. Microbiol.* **32**: 101-110.

746   Ronquist F & Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic
747 inference under mixed models. *Bioinformatics* **19**: 1572-1572.

748   Segovia L, Pinero D, Palacios R & Martinez-Romero E (1991) Genetic
749 structure of a soil population of nonsymbiotic *Rhizobium leguminosarum. Appl.*
750 *Environ. Microbiol.* **57**: 426-433.

751   Silva C, Kan FL & Martínez-Romero E (2007) Population genetic structure of
752 *Sinorhizobium meliloti* and *S. medicae* isolated from nodules of *Medicago* spp. in
753 Mexico. *FEMS Microbiol. Ecol.* **60**: 477-489.

754   Sørensen SJ, Bailey M, Hansen LH, Kroer N & Wuertz S (2005) Studying
755 plasmid horizontal transfer in situ: a critical review. *Nat. Rev. Microbiol.* **3**: 700-710.

756   Spaink HP (2000) Root nodulation and infection factors produced by rhizobial
757 bacteria. *Annu. Rev. Microbiol.* **54**: 257-288.

758   Sprent JI, Ardley JK & James EK (2013) From North to South: A latitudinal
759 look at legume nodulation processes. *S. Afr. J. Bot.* **89**: 31-41.

760   Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based
761 phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**:
762 2688-2690.

763   Sullivan JT, Trzebiatowski JR, Cruickshank RW, Gouzy J, Brown SD, Elliot
764 RM, Fleetwood DJ, McCallum NG, Rossbach U & Stuart GS (2002) Comparative
765 sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. *J.*
766 *Bacteriol.* **184**: 3086-3095.

767   Suominen L, Roos C, Lortet G, Paulin L & Lindström K (2001) Identification
768 and structure of the *Rhizobium galegae* common nodulation genes: evidence for
769 horizontal gene transfer. *Mol. Biol. Evol.* **18**: 907-916.

770   Talbi C, Delgado MJ, Girard L, Ramírez-Trujillo A, Caballero-Mellado J &
771 Bedmar EJ (2010) *Burkholderia phymatum* strains capable of nodulating *Phaseolus*
772 *vulgaris* are present in Moroccan soils. *Appl. Environ. Microbiol.* **76**: 4587-4591.

773   Tang J, Bromfield ESP, Rodrigue N, Cloutier S & Tambong JT (2012)
774 Microevolution of symbiotic *Bradyrhizobium* populations associated with soybeans in
775 east North America. *Ecol. Evol.* **2**: 2943-2961.

776   Tian CF, Young JP, Wang ET, Tamimi SM & Chen WX (2010) Population
777 mixing of *Rhizobium leguminosarum* bv. *viciae* nodulating *Vicia faba*: the role of
778 recombination and lateral gene transfer. *FEMS Microbiol. Ecol.* **73**: 563-576.

779   Tian CF, Zhou YJ, Zhang YM, Li QQ, Zhang YZ, Li DF, Wang S, Wang J,
780 Gilbert LB & Li YR (2012) Comparative genomics of rhizobia nodulating soybean
781 suggests extensive recruitment of lineage-specific genes in adaptations. *Proc. Natl.*
782 *Acad. Sci. U. S. A*. **109**: 8629-8634.

783   Uchiumi T, Ohwada T, Itakura M, Mitsui H, Nukui N, Dawadi P, Kaneko T,
784 Tabata S, Yokoyama T & Tejima K (2004) Expression islands clustered on the
785 symbiosis island of the *Mesorhizobium loti* genome. *J. Bacteriol.* **186**: 2439-2448.

786   Van Berkum P, Terefework Z, Paulin L, Suomalainen S, Lindström K &
787 Eardly BD (2003) Discordant phylogenies within the rrn loci of rhizobia. *J. Bacteriol.*
788 **185**: 2988-2998.

789   Van Cauwenberghe J, Verstraete B, Lemaire B, Lievens B, Michiels J &
790 Honnay O (2014) Population structure of root nodulating *Rhizobium leguminosarum*
791 in *Vicia cracca* populations at local to regional geographic scales. *Syst. Appl.*
792 *Microbiol.* **37**: 613-621.

793   Vetsigian K & Goldenfeld N (2005) Global divergence of microbial genome
794 sequences mediated by propagating fronts. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 7332-
795 7337.

796   Vial L, Groleau M, Dekimpe V & Deziel E (2007) *Burkholderia* diversity and
797 versatility: an inventory of the extracellular products. *J. Microbiol. Biotechnol.* **17**:
798 1407-1429.

799   Vinuesa P, Silva C, Werner D & Martínez-Romero E (2005) Population
800 genetics and phylogenetic inference in bacterial molecular systematics: the roles of
801 migration and recombination in *Bradyrhizobium* species cohesion and delineation.
802 *Mol. Phylogenet. Evol.* **34**: 29-54.

803   Vinuesa P, Silva C, Lorite MJ, Izaguirre-Mayoral ML, Bedmar EJ &
804 Martínez-Romero E (2005) Molecular systematics of rhizobia based on maximum
805 likelihood and Bayesian phylogenies inferred from *rrs*, *atpD*, *recA* and *nifH*
806 sequences, and their use in the classification of *Sesbania* microsymbionts from
807 Venezuelan wetlands. *Syst. Appl. Microbiol.* **28**: 702-716.

808   Vinuesa P, Rojas-Jiménez K, Contreras-Moreira B, Mahna SK, Prasad BN,
809 Moe H, Selvaraju SB, Thierfelder H & Werner D (2008) Multilocus sequence
810 analysis for assessment of the biogeography and evolutionary genetics of four
811 *Bradyrhizobium* species that nodulate soybeans on the Asiatic continent. *Appl.*
812 *Environ. Microbiol.* **74**: 6987-6996.

813   Vos M & Didelot X (2009) A comparison of homologous recombination rates
814 in bacteria and archaea. *ISME J.* **3**: 199-208.

815   Wang ET, Van Berkum P, Sui XH, Beyene D, Chen WX & Martínez-Romero
816 E (1999) Diversity of rhizobia associated with Amorpha fruticosa isolated from
817 Chinese soils and description of Mesorhizobium amorphae sp. nov. *Int. J. Syst.*
818 *Bacteriol.* **49**: 51-65.

819   Werle E, Schneider C, Renner M, Volker M & Fiehn W (1994) Convenient
820 single step, one tube purification of PCR products for direct sequencing. *Nucleic*
821 *Acids Res.* **22**: 4354-4355.

822   Wiedenbeck J & Cohan FM (2011) Origins of bacterial diversity through
823 horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiol.*
824 *Rev.* **35**: 957-976.

825   Xu Y & Murooka Y (1995) A large plasmid isolated from *Rhizobium huakuii*
826 bv. *renge* that includes genes for both nodulation of *Astragalus sinicus* cv. Japan and
827 nitrogen fixation. *J. Ferment. Bioeng.* **80**: 276-279.

828   Young JPW, Crossman LC, Johnston AWB, Thomson NR, Ghazoui ZF, Hull
829 KH, Wexler M, Curson ARJ, Todd JD & Poole PS (2006) The genome of *Rhizobium*
830 *leguminosarum* has recognizable core and accessory components. *Genome Biol.* **7**:
831 R34.

832   Zhang X-X, Turner SL, Guo X-W, Yang H-J, Debellé F, Yang G-P, Dénarié J,
833 Young JPW & Li F-D (2000) The common nodulation genes of *Astragalus sinicus*
834 rhizobia are conserved despite chromosomal diversity. *Appl. Environ. Microbiol.* **66**:
835 2988-2995.

836   Zhao CT, Wang ET, Chen WF & Chen WX (2008) Diverse genomic species
837 and evidences of symbiotic gene lateral transfer detected among the rhizobia
838 associated with *Astragalus* species grown in the temperate regions of China. *FEMS*
839 *Microbiol. Lett.* **286**: 263-273.

840   Zuleta LF, de Cunha C, De Carvalho FM, Ciapina LP, Souza RC, Mercante
841 FM, De Faria SM, Baldani JI, Straliotto R & Hungria M (2014) The complete genome
842 of *Burkholderia phenoliruptrix* strain BR3459a, a symbiont of *Mimosa flocculosa*:
843 highlighting the coexistence of symbiotic and pathogenic genes. *BMC Genomics* **15**:
844 535.

Figure legends
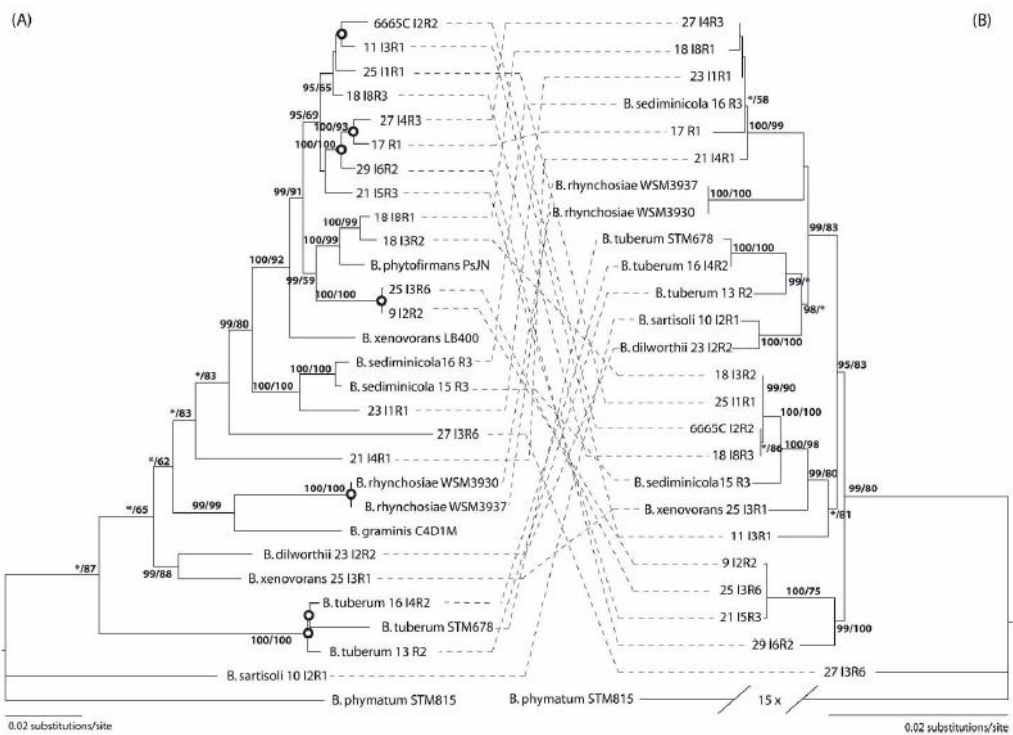
Fig. 1. Phylogenetic incongruences between chromosomal and *nod* sequence data of *Burkholderia* isolates. Comparison of the best Maximum Likelihood trees based on **(A)** chromosomal (16S rRNA, *recA*, *atpD*, *gyrB*) and **(B)** nodulation genes (*nodA*, *nodB*, *nodC*) (right tree). Support values for the Bayesian and Maximum Likelihood analyses are shown at the nodes. Dashed lines indicate the species association between the chromosomal and nodulation gene trees. Nodes highlighted by a circle represent events of co-speciation as revealed by the reconciliation analysis.
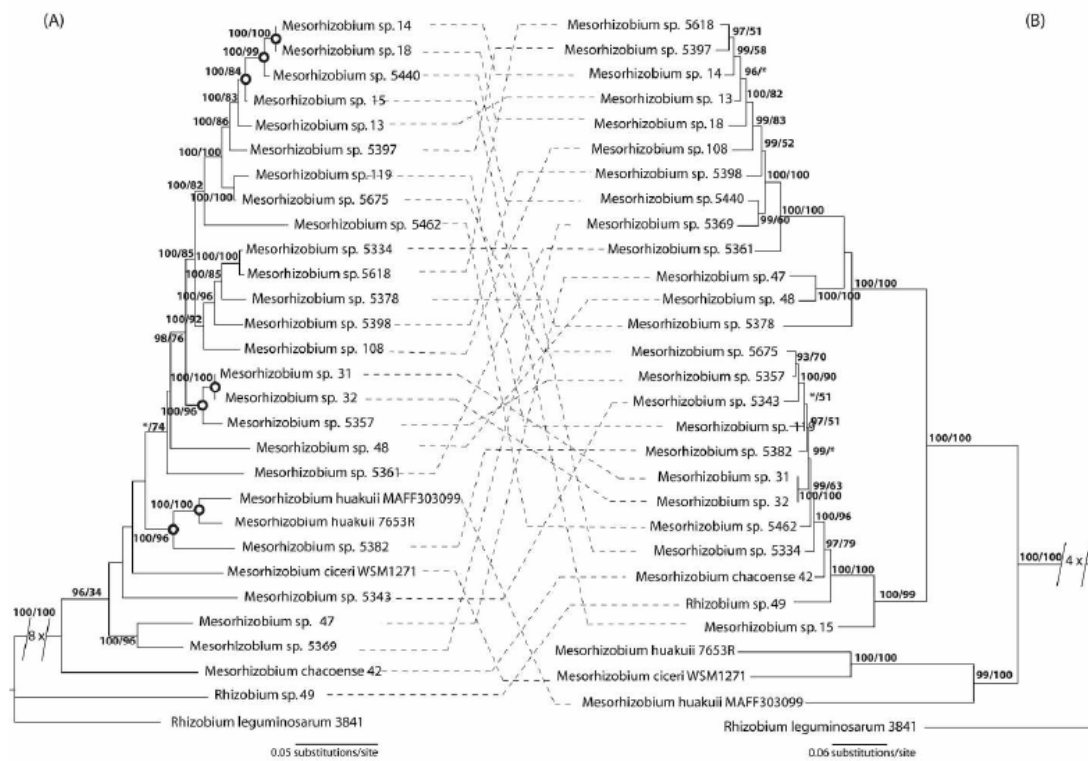
Fig. 2. Phylogenetic incongruences between chromosomal and *nod* sequence data of *Mesorhizobium* isolates. Comparison of best Maximum Likelihood trees based on **(A)** chromosomal (*recA*, *atpD*, *gyrB, glnA*) and **(B)** nodulation genes (*nodA*, *nodB*, *nodC*) (right tree). Support values for the Bayesian and Maximum Likelihood analyses are shown at the nodes. Dashed lines indicate the species association between the chromosomal and nodulation gene trees. Nodes highlighted by a circle represent events of co-speciation as revealed by the reconciliation analysis.
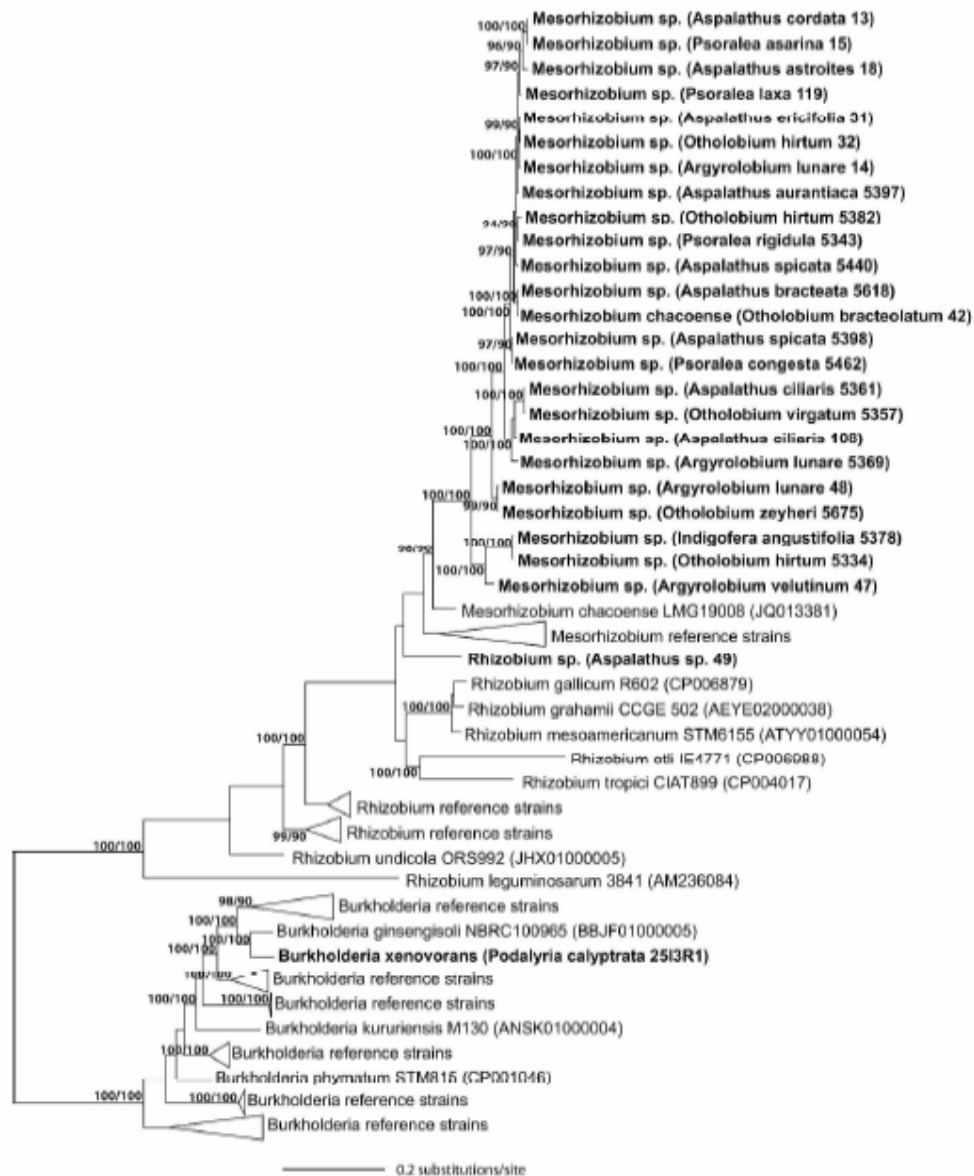
Fig. 3. Phylogenetic relationships based on *acdS* sequences of *Mesorhizobium*, *Burkholderia* and *Rhizobium* isolates. Major lineages are schematically represented by triangles. Support values for the Bayesian and Maximum Likelihood analyses are shown at the nodes. A full phylogram is presented in Fig. S7.