

Paraburkholderia strydomiana* sp. nov. and *Paraburkholderia steynii* sp. nov. – rhizobial symbionts of the fynbos legume *Hypocalyptus sophoroides

Antonie van Leeuwenhoek

CW Beukes¹, ET Steenkamp¹, E van Zyl¹, J Avontuur¹, WY Chan^{1,3}, AI Hassen², M Palmer¹,
LS Mthombeni¹, FL Phalane², Sereme TK¹, SN Venter¹

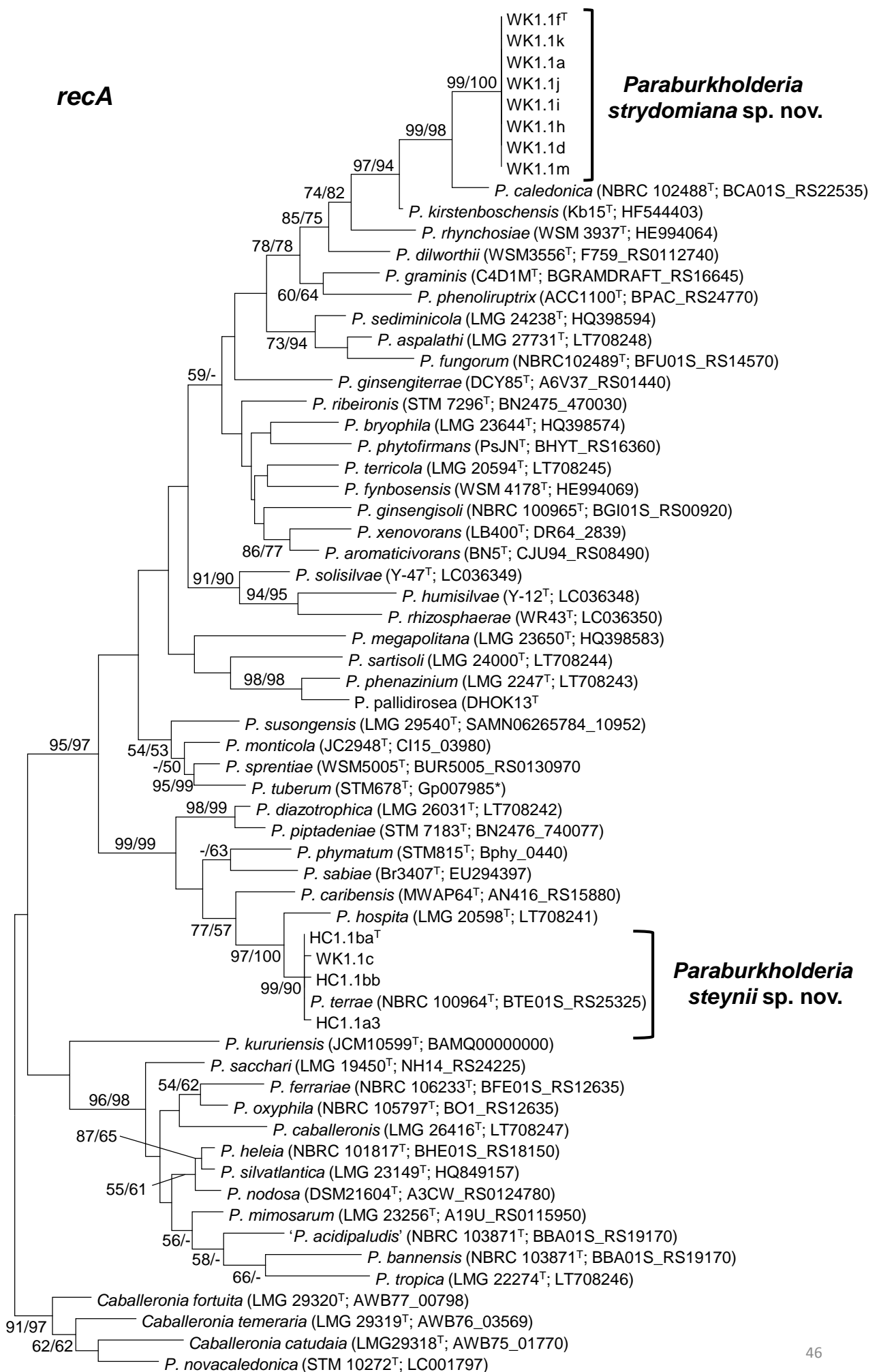
¹Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria 0002, Gauteng, South Africa

²Agricultura Research Council (ARC), Plant Health and Protection Institute, Pretoria 0002, Gauteng, South Africa

³Agricultural Research Council (ARC), Biotechnology Platform, Onderstepoort, 0110, Gauteng, South Africa

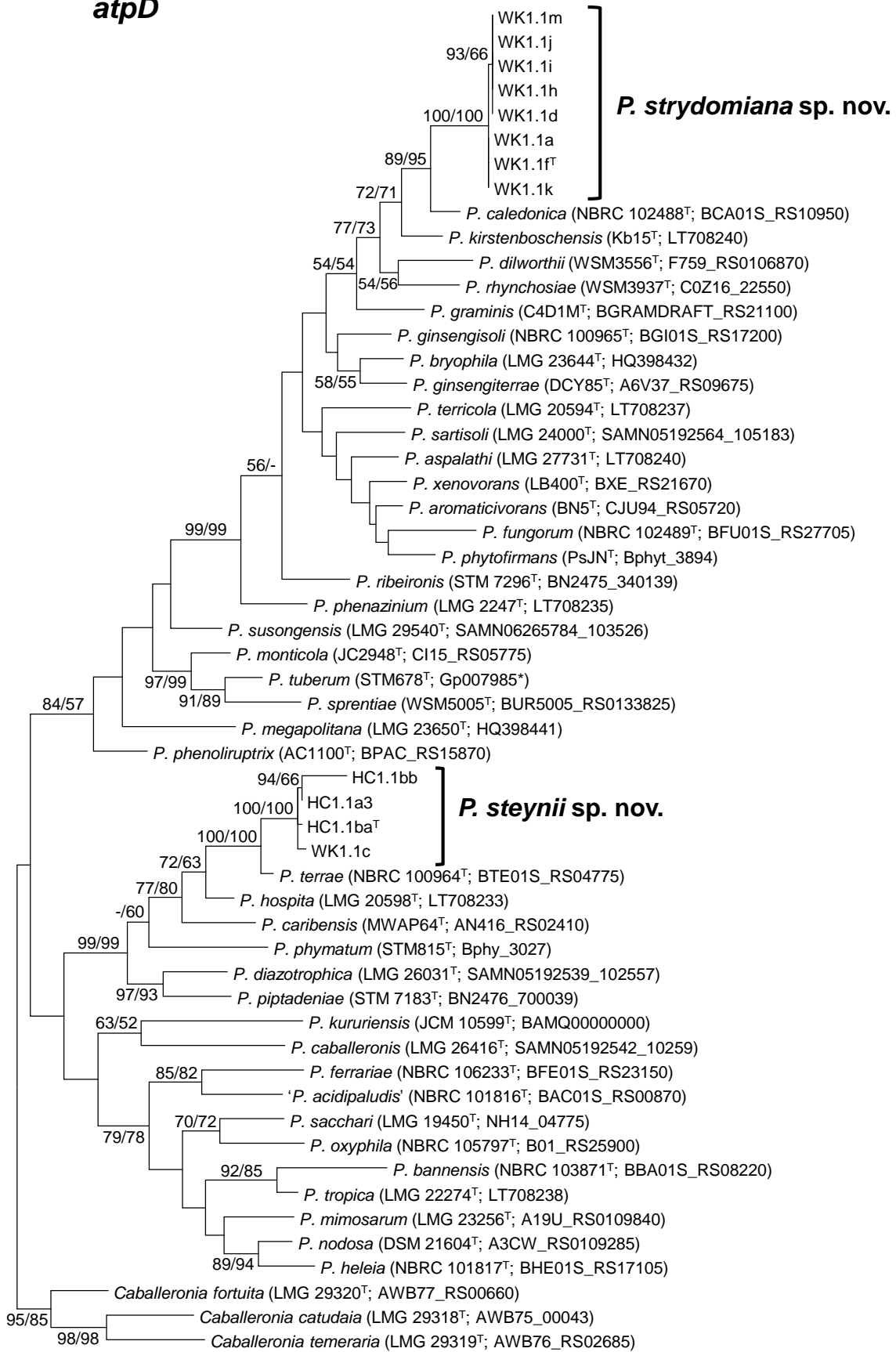
Supplementary Figure S1 One of nine most parsimonious tree inferred from the analysis of *recA* sequences of a group of 12 *Paraburkholderia* isolates associated with *Hypocalyptus sophoroides*. The length of the tree is 739 steps, the consistency index is 0.30, the retention index is 0.69 and the homoplasy index is 0.70. Statistical support for the groupings are shown in the order MP/ML and only support $\geq 50\%$ are listed. Information regarding the type strain, GenBank accession number/locus tag for each species is indicated in brackets, while the number accompanied by * is a GOLD Project ID. Species names which appear in inverted commas ('...') are combinations which have not yet been validly published. The scale bar corresponds to 10 character state changes.

recA



Supplementary Figure S2 One out of two most parsimonious trees inferred from the analysis of *atpD* sequences of a group of 12 *Paraburkholderia* isolates associated with *Hypocalyptus sophoroides*. The length of the tree is 1076 steps, the consistency index is 0.32, the retention index is 0.70 and the homoplasy index is 0.68. Statistical support for the groupings is shown in the order MP/ML and only support $\geq 50\%$ is listed. Information regarding the type strain, GenBank accession number/locus tag for each species is indicated in brackets, while the number accompanied by * is a GOLD Project ID. Species names which appear in inverted commas ('...') are combinations which have not yet been validly published. The scale bar corresponds to 20 character state changes.

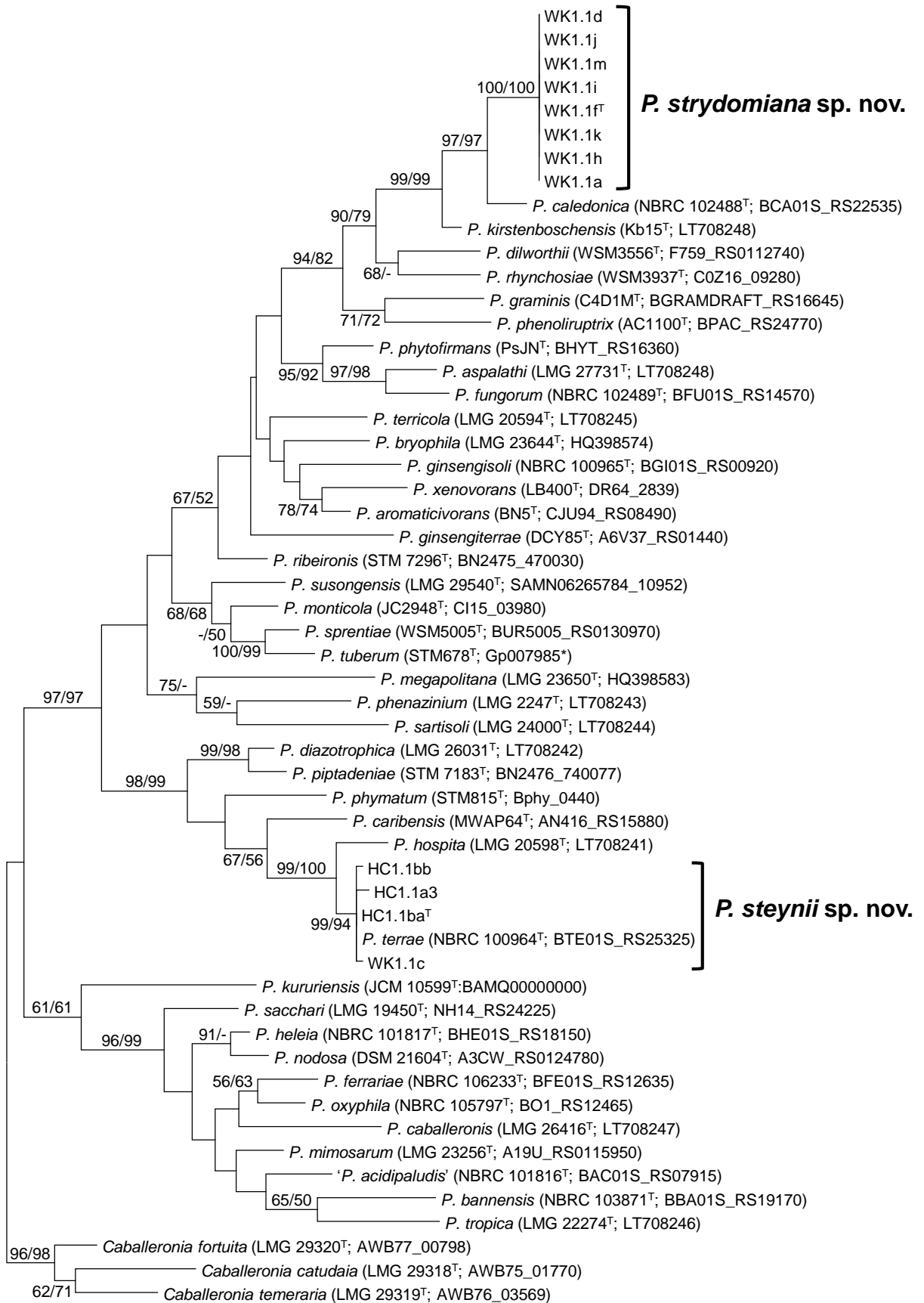
atpD



20

Supplementary Figure S3 One of 12 most parsimonious trees inferred after analysing the *recA* sequences of the 12 *Paraburkholderia* isolates associated with *Hypocalyptus sophoroides*. The tree has a length of 793 steps, a consistency index of 0.31, a retention index of 0.69 and a homoplasy index of 0.69. Bootstrap support is indicated only when $\geq 50\%$ and appear in the order MP/ML. Listed in brackets is the type strain and GenBank accession number/locus tag for each species, while the number accompanied with * is a GOLD Project ID. Species appearing in inverted commas ('...') are combinations which have not yet been validly published. The scale bar corresponds to 10 character state changes.

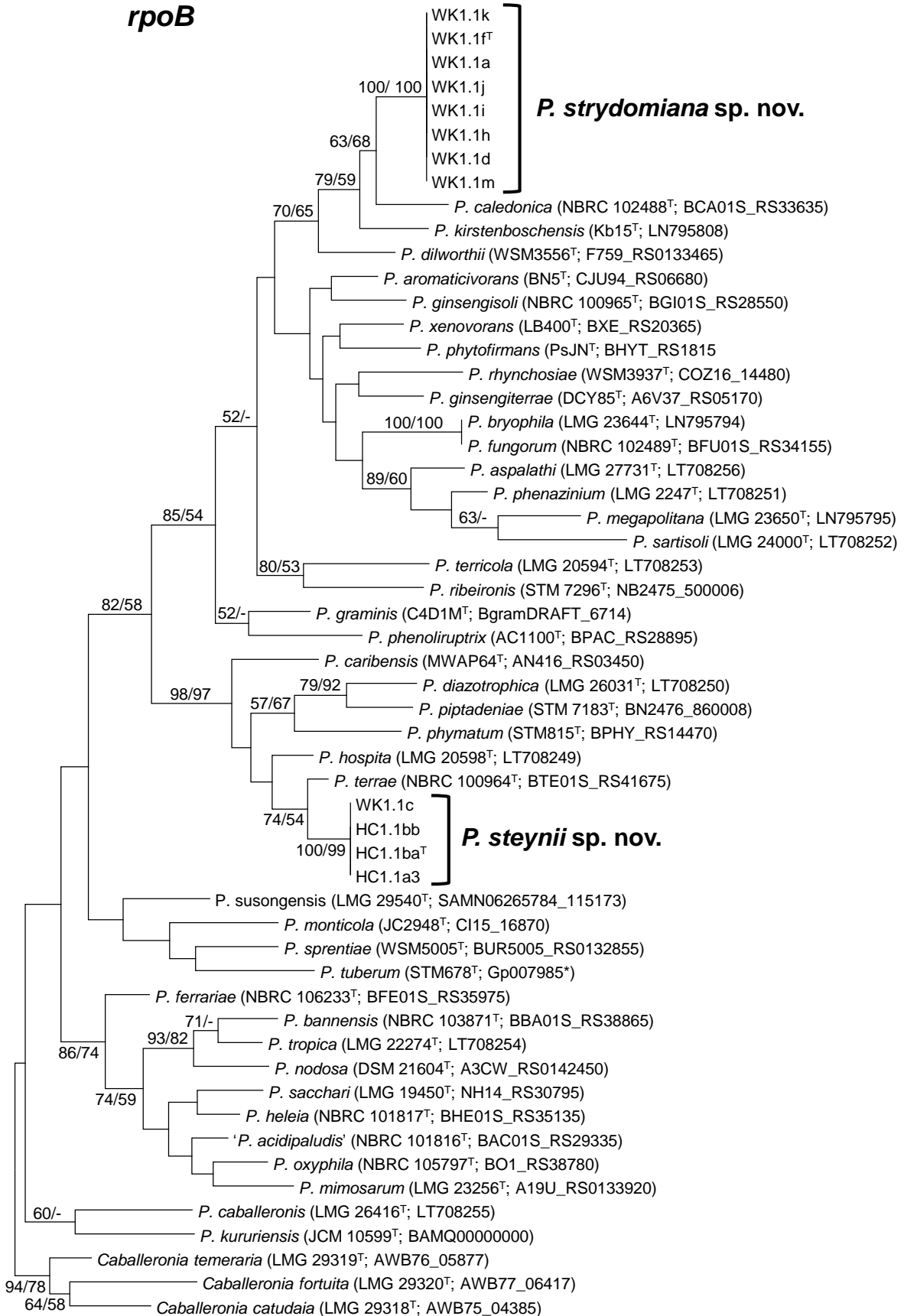
recA



10

Supplementary Figure S4 One of two most parsimonious trees resulting from the analysis of *rpoB* sequences from 12 *Paraburkholderia* isolates associated with *Hypocalyptus sophoroides*. The tree length is 830 steps, it has a consistency index of 0.28, a retention index of 0.65 and a homoplasy index 0.72. Bootstrap support is indicated in the order MP/ML and only support of $\geq 50\%$ is shown. Information regarding the type strain, GenBank accession number/locus tag of each species is listed brackets, while * denotes a GOLD Project ID. Species names in inverted commas ('...') are combinations which have not yet been validly published. The scale bar corresponds to 10 character state changes.

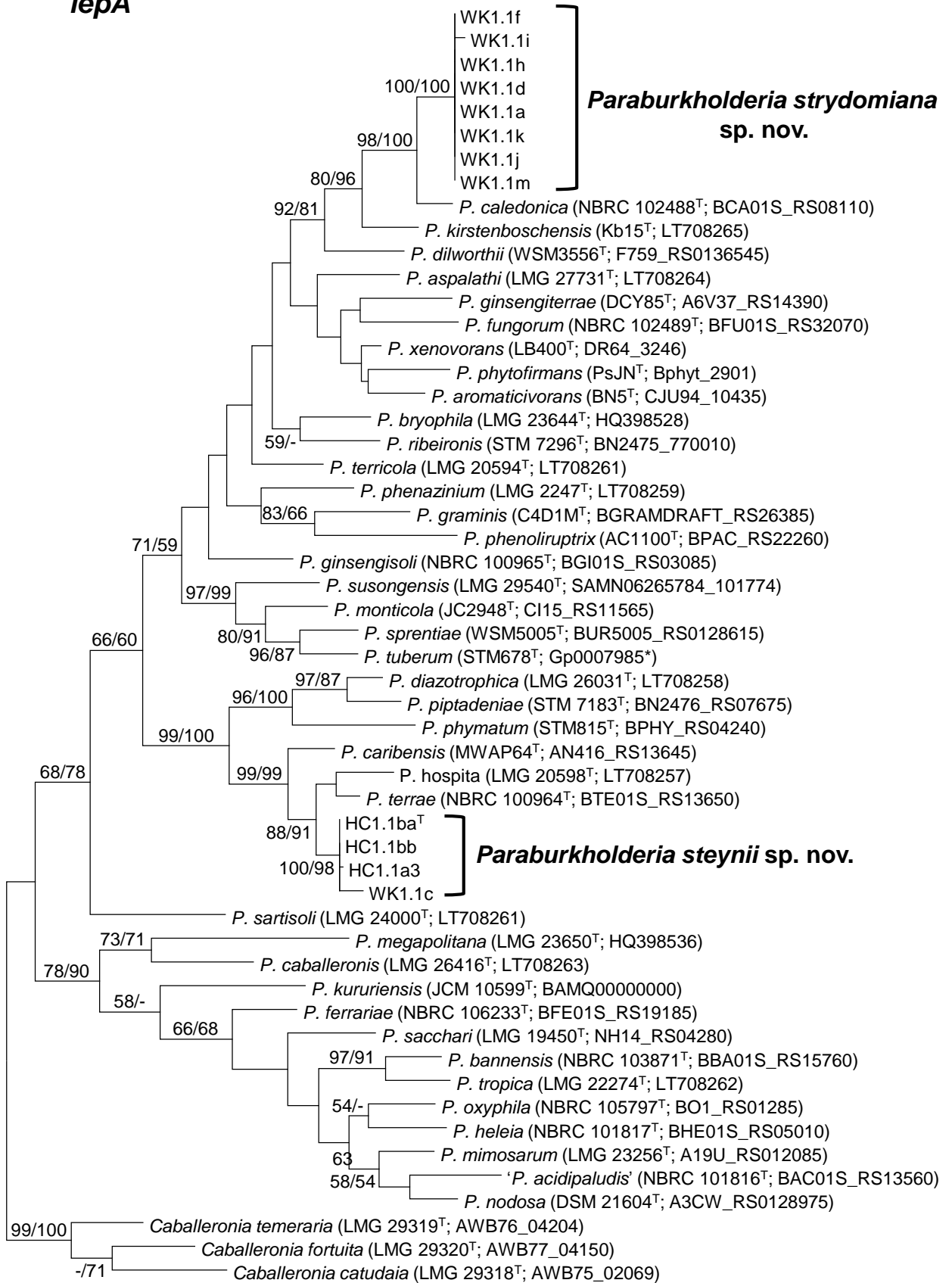
rpoB



10

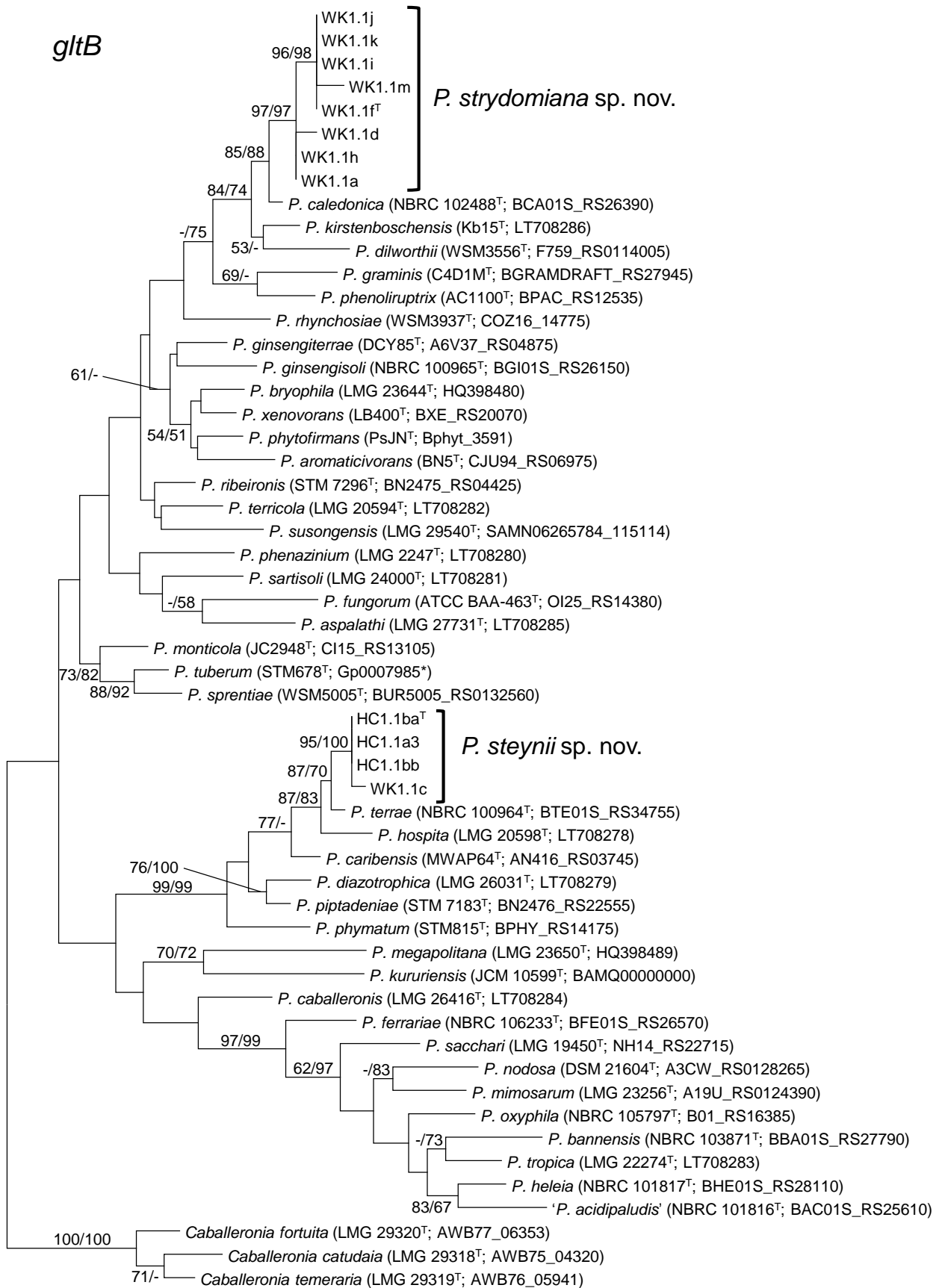
Supplementary Figure S5 One of three most parsimonious trees resulting from the analysis of *lepA* sequences for the 12 *Paraburkholderia* isolates associated with *Hypocalyptus sophoroides* in this study. This tree is 1502 steps long, has a consistency index of 0.30, a retention index of 0.64, and a homoplasy index of 0.70. Bootstrap support is indicated in the order MP/ML and only support of $\geq 50\%$ are shown. Type strain and GenBank accession number/locus tag for each species is listed in brackets, while * denotes a Gold Project ID. Species in inverted commas ('...') are combinations awaiting valid publication. The scale bar corresponds to 20 character state changes.

lepA



20

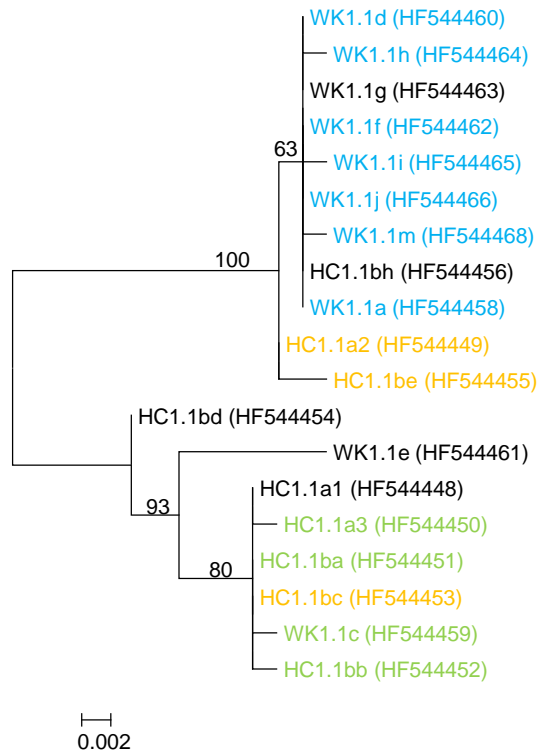
Supplementary Figure S6 The most parsimonious tree inferred from the analysis of the *gltB* sequences of 12 *Paraburkholderia* isolates associated with *Hypocalyptus sophoroides*. The tree has a length of 648 steps, a consistency index is of 0.33, a retention index of 0.69 and a homoplasy index of 0.67. Statistical support for the groupings are indicated in the order MP/ML and only support of $\geq 50\%$ are shown. Listed in brackets is the type strain and GenBank accession number/locus tag for each species, while * indicated a GOLD Project ID. Species names in inverted commas ('...') are combinations which are still awaiting valid publication. The scale bar corresponds to 10 character state changes.



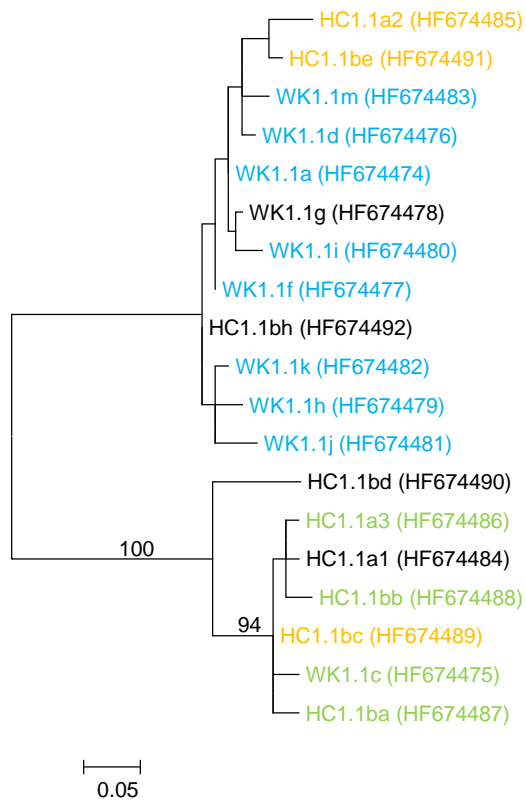
Supplementary Figure S7 One of of four most parsimonious trees inferred from analysing a five gene concatenated dataset (*atpD* + *recA* + *rpoB* + *gltB* + *lepA*) including the 12 focal *Paraburkholderia* species. The tree has a length of 5059 steps, a consistency index of 0.29, a retention index of 0.64 and a homoplasy index of 0.71. Statistical support of $\geq 50\%$ is indicated in the order MP/ML. Listed in brackets is the type strain number for each species. Names in inverted commas ('...') are combinations which are still awaiting valid publication. The scale bar corresponds to 50 character state changes.

Supplementary Figure S8 A. Maximum likelihood phylogeny of a segment of the *nifH* locus for all available *H. sophoroides* associated rhizobial isolates. The tree is rooted at the midpoint, while the key indicates to which *Paraburkholderia* species the isolate belongs. Listed in brackets is the GenBank accession number for that isolate. The scale bar indicates the number of nucleotide changes per site. **B.** Maximum likelihood phylogeny of a fragment of the *nodA* locus for all available *H. sophoroides* rhizobial isolates. Listed in brackets are the GenBank accession number for the isolate. The tree is rooted at the midpoint and follows the same key as for part A. The scale bar indicates the number of nucleotide changes per site.

A.



B.



P. steynii

P. strydomiana

P. kirstenboschensis

As yet unnamed isolates