Systematic and Applied Microbiology 38 (2015) 545–554

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

journal homepage: www.elsevier.de/syapm

Burkholderia kirstenboschensis sp. nov. nodulates papilionoid legumes indigenous to South Africa

Emma T. Steenkamp[a], Elritha van Zyl[a], Chrizelle W. Beukes[a], Juanita R. Avontuur[a], Wai Yin Chan[a], Marike Palmer[a], Lunghile S. Mthombeni[a], Francina L. Phalane[b], T. Karabo Sereme[a], Stephanus N. Venter[a,*]

[a] *Department of Microbiology and Plant Pathology, Faculty of Natural and Agricultural Sciences, Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, South Africa*
[b] *Agricultural Research Council (ARC), Plant Protection Research Institute, Pretoria, Gauteng, South Africa*

ARTICLE INFO

Article history:
Received 11 June 2015
Received in revised form 8 September 2015
Accepted 11 September 2015

Keywords:
Hypocalyptus
Virgilia
Beta-rhizobia
DNA–DNA hybridization
Average nucleotide identity (ANI)
Burkholderia

ABSTRACT

Despite the diversity of *Burkholderia* species known to nodulate legumes in introduced and native regions, relatively few taxa have been formally described. For example, the Cape Floristic Region of South Africa is thought to represent one of the major centres of diversity for the rhizobial members of *Burkholderia*, yet only five species have been described from legumes occurring in this region and numerous are still awaiting taxonomic treatment. Here, we investigated the taxonomic status of 12 South African root-nodulating *Burkholderia* isolates from native papilionoid legumes (*Hypocalyptus coluteoides*, *H. oxalidifolius*, *H. sophoroides* and *Virgilia oroboides*). Analysis of four gene regions (16S rRNA, *recA*, *atpD* and *rpoB*) revealed that the isolates represent a genealogically unique and exclusive assemblage within the genus. Its distinctness was supported by all other aspects of the polyphasic approach utilized, including the genome-based criteria DNA–DNA hybridization (≥70.9%) and average nucleotide identities (≥96%). We accordingly propose the name *B. kirstenboschensis* sp. nov. for this taxon with isolate Kb15^T (=LMG 28727^T; =SARC 695^T) as its type strain. Our data showed that intraspecific genome size differences (≥0.81 Mb) and the occurrence of large DNA regions that are apparently unique to single individuals (16–23% of an isolate's genome) can significantly limit the value of data obtained from DNA–DNA hybridization experiments. Substitution of DNA–DNA hybridization with whole genome sequencing as a prerequisite for the description of *Burkholderia* species will undoubtedly speed up the pace at which their diversity are documented, especially in hyperdiverse regions such as the Cape Floristic Region.

1 Introduction

Since the initial discovery that *Burkholderia tuberum* is capable of nodulating several species of the South African legume *Cyclopia* [25], many additional *Burkholderia* species have been demonstrated to efficiently establish the nitrogen-fixing symbiosis with a variety of legumes. To date, formal descriptions have been provided for 15 of these rhizobial or root-nodulating *Burkholderia* species [10,11,13,29,46,64,65,76,79]. Among them, only five species (including *B. tuberum*) have been described as symbionts of South African legumes. The other four species also nodulate legumes in the subfamily Papilionoideae, where both *B. sprentiae* and *B. dilworthii* were isolated from *Lebeckia ambigua* [22,24], *B. rhynchosiae* originated from *Rhynchosia ferulifolia* [23] and *B. aspalathi* was isolated from the root nodules of *Aspalathus abietina* [47].

In comparison to the knowledge available for rhizobial genera in the Alphaproteobacteria, which have been studied for more than a century [80], our understanding of the diversity, distribution and evolution of the legume symbionts in the genus *Burkholderia* (class Betaproteobacteria, family Burkholderiaceae [30]) is limited. Research during the last decade has shown that legumes are often associated with a large diversity of *Burkholderia* symbionts, particularly where the root-nodulating bacteria of indigenous or endemic host species were explored [4,5,7,8,12,53]. Phylogenetic analyses of various protein-coding genes have also shown that the association of these legumes with their *Burkholderia* symbionts likely

* Corresponding author. Tel.: +27 12 420 4100; fax: +27 12 420 3266.
E-mail addresses: Emma.steenkamp@up.ac.za (E.T. Steenkamp), Elritha.vanzyl@fabi.up.ac.za (E. van Zyl), Chrizelle.beukes@fabi.up.ac.za (C.W. Beukes), Juanita.avontuur@fabi.up.ac.za (J.R. Avontuur), Annie.chan@fabi.up.ac.za (W.Y. Chan), Marike.duplessis@fabi.up.ac.za (M. Palmer), MthombeniL@dwa.gov.za (L.S. Mthombeni), PhalaneF@arc.agric.za (F.L. Phalane), Theodorah.Sereme@ZA.nestle.com (T.K. Sereme), Fanus.venter@up.ac.za (S.N. Venter).

http://dx.doi.org/10.1016/j.syapm.2015.09.003
0723-2020/

represents an ancient partnership with a long and complex evolutionary history [5,7]. For example, the South African Cape Floristic Region (CFR) with its indigenous legumes, some of which trace their origins to the early Oligocene [62], represent a unique centre of diversity for the rhizobial species of this genus [5,42].

This study investigated the taxonomic status of 12 *Burkholderia* isolates that were originally isolated from the root nodules of the papilionoid hosts *Hypocalyptus coluteoides*, *H. oxalidifolius*, *H. sophoroides* and *Virgilia oroboides* growing in various locations in the CFR of South Africa [5]. Based on sequence analyses of the 16S ribosomal RNA (rRNA) and recombinase A (*recA*) genes, these isolates were suggested to represent a distinct group nested within the so-called "environmental clade" of *Burkholderia* [5]. This clade was recently suggested to represent a new genus (i.e. "*Caballeronia*" as proposed by Gyaneshwar et al. [36] or "*Paraburkholderia*" as proposed by Sawana et al. [61]; neither proposal have been formally implemented and a type species for this new taxon has not been identified). This proposed generic taxon includes most of the described rhizobia, as well as the plant-associated species, but excludes all clinically important *Burkholderia* species. The specific aims of this study were to (i) evaluate the conspecificity of the 12 South African isolates, (ii) evaluate their uniqueness within the overall genus and (iii) to provide a formal description of the taxon thus delineated. To achieve these aims, we utilized a polyphasic approach, where a range of phenotypic characters, phylogenetic information using Multi-Locus Sequence Analysis (MLSA) and genome-based criteria such as DNA–DNA hybridization and average nucleotide identities (ANI) [33,39,59] were used. To accomplish the latter, we also determined the whole genome sequences for two isolates of our newly recognized taxon.

2 Materials and Methods

2.1 Bacterial strains and growth conditions

The 12 *Burkholderia* isolates all originated from locations in the Western and Eastern Cape Provinces of South Africa. Three isolates (HC1.1be, HC1.1bc and HC1.1a2) associated with *H. sophoroides* were recovered from soil collected in Old du Toit's Kloof Pass, two isolates (RAU2b and RAU2d2) were recovered from *H. coluteoides* growing at the Storms River Bridge, and two isolates originated from *H. oxalidifolius* (RAU6.4d and RAU6.4f) occurring in the Fernkloof Nature Reserve. Five isolates (Kb2, Kb13, Kb14, Kb16 and Kb15^T) were recovered from nodules of *V. oroboides* occurring in the Kirstenbosch National Botanical Gardens. The capacity of these isolates to nodulate their original hosts and/or other hosts such as *Vigna unguiculata* (cowpea) or *Macroptilium atropurpureum* (siratro) has been demonstrated previously [5]. Isolate Kb15^T is also available from the Belgian Coordinated Collections of Microorganisms (Universiteit Gent, Belgium; LMG 28727^T) or from the South African Rhizobium Collection (Pretoria, South Africa; SARC 695^T). For comparative purposes, *B. caledonica* strain LMG 19076^T, *B. fungorum* strain LMG 16225^T, *B. megapolitana* strain LMG 23650^T and *B. dilworthii* strain LMG 27173^T were also included in the study.

All isolates were routinely grown at 28 °C on Tryptone Yeast Extract Agar (TYA) enriched with $CaCl_2.2H_2O$ (0.088 g/l) or on Yeast Mannitol Agar (YMA). Isolates were preserved at -70 °C in 20% glycerol as part of the University of Pretoria's Rhizobium Culture Collection.

2.2 16S rRNA gene analysis and MLSA

This study utilized the previously determined sequences for the 16S rRNA and *recA* genes [5], together with those for the genes *atpD* and *rpoB* that respectively encode the beta subunits for ATP synthase and RNA polymerase. The *atpD* and *rpoB* gene fragments were amplified with primers *atpDF* (5′ GAT CGT ACA GTG CAT CGG 3′) and *atpDR* (5′ ATC GTG CCG ACC ATG TAG 3′) [3] and primers *RpoB*-1394F (5′ TGG CGG AAA ACC AGT TCC GCG 3′) and *RpoB*-2430R (5′ AGC CGT TCC ACG GCA TGA ACG 3′), respectively. The latter primer set was designed based on publicly available *rpoB* sequences by making use of BioEdit v7.0.5.3 [37] and Primer3 [60]. These two primer sets targeted, respectively, 1200 base pair (bp) and 1000 bp regions of the *atpD* and *rpoB* genes (see Table S1 for details regarding PCR and cycling conditions).

PCR products were cleaned using polyethylene glycol precipitation [69] and sequenced in both directions using the original PCR primers for *rpoB* and specific sequencing primers for *atpD* [3]. For this purpose, the ABI PRISM Big Dye Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems, CA, USA) and an ABI 377 Automated Capillary DNA sequencer (Applied Biosystems) were utilized. The *atpD* and *rpoB* sequences were then manually curated using ChromasLite v2.01 (Technelysium, Queensland, Australia) and BioEdit.

For 10 of the *Burkholderia* isolates examined in this study, single gene nucleotide datasets were compiled for the *atpD* and *rpoB* sequences generated here and the 16S rRNA and *recA* sequences determined previously (see Beukes et al. [5] for accession numbers). For isolates Kb15^T and RAU2d2, the relevant gene sequences used in the nucleotide datasets were obtained from the genome sequences of these bacteria (see below). These datasets also included the sequences for the recognized type strains of described *Burkholderia* species as they appear in the List of Prokaryotic Names with Standing in Nomenclature (LPSN) [26,54] (www.bacterio.net). The 16S rRNA dataset was aligned with the online version of MAFFT (Multiple Alignment using Fast Fourier Transformation; http://mafft.cbrc.jp/alignment/server/) using the Q-INS-I strategy that takes secondary structure into account [38]. The nucleotide datasets for the protein coding genes were manually aligned according to the inferred amino acid sequences. A concatenated dataset was generated with SequenceMatrix [74].

The 16S rRNA dataset was analyzed using CLC bio Main Workbench v7.03 (QIAGEN, Århus, Denmark) to determine the percentage sequence similarity among the various taxa included. The four single gene datasets and the concatenated dataset were subjected to maximum parsimony and maximum-likelihood phylogenetic analyses. For the maximum parsimony analyses, MEGA6 [72] was utilized to determine the most parsimonious tree(s) by making use of the heuristic tree search algorithm with 1000 random addition-sequence replicates and Tree Bisection-Reconnection (TBR) branch swapping. Branch lengths were calculated by means of the average pathway method [50]. For the maximum-likelihood analyses, PhyML v3.1 [35] or MEGA6 [72] were utilized with the best-fit substitution model parameters as indicated by MEGA5 or jModelTest v0.1.01 [28,35,56,71]. The 16S rRNA data used the model of Tamura and Nei [70], the *atpD* data used the TIM2 "transitional" model [56] while the *rpoB*, *recA* and concatenated datasets all utilized the General Time Reversible model (GTR) [73]; in all cases these models incorporated gamma (G) correction of among site rate variation and a proportion of invariable sites (I). In these analyses, the best of the Nearest-Neighbour Interchange (NNI) and Subtree-Pruning-Regrafting (SPR) search algorithms were used for tree searches, which were initiated using the best of five random starting trees. For both the maximum parsimony and maximum-likelihood analyses, branch support was estimated using non-parametric bootstrap analysis [27] based on 1000 pseudoreplicates.