Systematic and Applied Microbiology 37 (2014) 208–215

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

journal homepage: www.elsevier.de/syapm

Phylogeny of the *Rhizobium–Allorhizobium–Agrobacterium* clade supports the delineation of *Neorhizobium* gen. nov.

Seyed Abdollah Mousavi[a,*], Janina Österman[a], Niklas Wahlberg[b], Xavier Nesme[c], Céline Lavire[c], Ludovic Vial[c], Lars Paulin[d], Philippe de Lajudie[e], Kristina Lindström[a,f]

[a] University of Helsinki, Division of Microbiology and Biotechnology, Department of Food and Environmental Sciences, Biocentre 1, Viikinkaari 9, P.O. Box 56, Helsinki FIN-00014, Finland
[b] University of Turku, Laboratory of Genetics, Department of Biology, 20014 Turku, Finland
[c] University of Lyon, Université Lyon1, Ecologie Microbienne, UMR CNRS 5557/USC INRA 1364, 16 rue R. Dubois, F-69622 Villeurbanne Cedex, France
[d] University of Helsinki, Institute of Biotechnology, DNA Sequencing and Genomics Laboratory, Viikinkaari 4, P.O. Box 56, Helsinki FIN-00014, Finland
[e] IRD, Laboratoire des Symbioses Tropicales et Méditerranéennes, Campus de Baillarguet TA A-82/J, 34398 Montpellier Cédex 5, France
[f] University of Helsinki, Department of Environmental Sciences, Viikinkaari 2a, P.O. Box 65, Helsinki FIN-00014, Finland

ARTICLE INFO

Article history:
Received 26 July 2013
Received in revised form
14 December 2013
Accepted 16 December 2013

Keywords:
MLSA
Housekeeping genes
Rhizobium galegae
Agrobacterium
Allorhizobium

ABSTRACT

The genera *Agrobacterium*, *Allorhizobium*, and *Rhizobium* belong to the family *Rhizobiaceae*. However, the placement of a phytopathogenic group of bacteria, the genus *Agrobacterium*, among the nitrogen-fixing bacteria and the unclear position of *Rhizobium galegae* have caused controversy in previous taxonomic studies. To resolve uncertainties in the taxonomy and nomenclature within this family, the phylogenetic relationships of generic members of *Rhizobiaceae* were studied, but with particular emphasis on the taxa included in *Agrobacterium* and the "*R. galegae* complex" (*R. galegae* and related taxa), using multilocus sequence analysis (MLSA) of six protein-coding housekeeping genes among 114 rhizobial and agrobacterial taxa. The results showed that *R. galegae*, *R. vignae*, *R. huautlense*, and *R. alkalisoli* formed a separate clade that clearly represented a new genus, for which the name *Neorhizobium* is proposed. *Agrobacterium* was shown to represent a separate cluster of mainly pathogenic taxa of the family *Rhizobiaceae*. *A. vitis* grouped with *Allorhizobium*, distinct from *Agrobacterium*, and should be reclassified as *Allorhizobium vitis*, whereas *Rhizobium rhizogenes* was considered to be the proper name for former *Agrobacterium rhizogenes*. This phylogenetic study further indicated that the taxonomic status of several taxa could be resolved by the creation of more novel genera.

Introduction

Rhizobia are *Proteobacteria* that can enter nitrogen-fixing symbioses with legumes. Until 1982, all rhizobial species were placed in a single genus, *Rhizobium*, which conformed to this feature. Since then, 16 genera in two subphyla have been distinguished among rhizobia by following standard rules for general bacterial taxonomy. Rhizobial genera include *Agrobacterium*, *Allorhizobium*, *Aminobacter*, *Azorhizobium*, *Bradyrhizobium*, *Devosia*, *Mesorhizobium*, *Methylobacterium*, *Microvirga*, *Ochrobactrum*, *Phyllobacterium*, *Rhizobium*, *Shinella* and *Ensifer* (syn. *Sinorhizobium*) [15,16]. Although *Agrobacterium* members (including *A. larrymoorei*, *A. rhizogenes*, *A. rubi*, *A. tumefaciens*, and *A. vitis*) were identified as plant pathogenic *Proteobacteria* (they induce either tumors or hairy roots on their host plants and do not form symbiotic nodules on the plants), they were placed in the "supercluster" *Rhizobium* (family *Rhizobiaceae*). This caused controversy when naming these pathogenic bacteria. The transfer of the species *Agrobacterium rhizogenes*, *A. rubi*, *A. tumefaciens*, and *A. vitis* to *Rhizobium* was proposed by Young et al. [45], whilst Farrand et al. [6] supported *Agrobacterium* as a proper name for this group of pathogenic rhizobial species. Since then, *R. rhizogenes* has been recognized as a true *Rhizobium*, while other pathogenic taxa were considered as *Agrobacterium*. *Agrobacterium* thus includes the well-defined species (*A. vitis*, *A. larrymoorei*, *A. rubi*, *A. fabrum*) along with the well-delineated genomospecies (e.g. genomovars) that have not yet received a Latin binomial name. Remarkably, the type strains of *A. radiobacter* and *A. tumefaciens* were found to belong to the same genomospecies. As a result, these species names are synonymous and *A. tumefaciens* is no longer a valid species name because it was described after *A. radiobacter* [1,12,17].

Rhizobium galegae was isolated from root nodules of *Galega orientalis*, and is capable of inducing nodules on at least two plant species, *G. orientalis* and *G. officinalis* [14]. In *R. galegae*, the two symbiovars (sv.) orientalis and officinalis were described based on nitrogen fixation host specificity in *Galega* plants [24].

* Corresponding author. Tel.: +358 9191 59279.
E-mail address: seyed.mousavi@helsinki.fi (S.A. Mousavi).

0723-2020/$ – see front matter © 2014 Elsevier GmbH. All rights reserved.
http://dx.doi.org/10.1016/j.syapm.2013.12.007

R. huautlense, *R. alkalisoli* and *R. vignae*, which were isolated from the effective nodules of *Sesbania herbacea*, *Caragana intermedia* and multiple legume species, respectively, were more related to *R. galegae* than the other rhizobia based on 16S rRNA gene phylogeny [18,25,43]. Zakhia et al. [47] reported that several *Rhizobium* isolates from *Astragalus cruciatus*, *Argyrolobium uniflorum*, *Anthyllis henoniana*, *Lotus creticus*, *Medicago marina* and *M. truncatula* were closely related to *R. galegae* and *R. huautlense*, and were therefore candidates for either of these two *Rhizobium* species. *Rhizobium* sp. HAMBI 3429 isolated from *Glycyrrhiza uralensis* is also closely related to *R. galegae* [13]. *R. galegae*, *R. huautlense*, *R. vignae*, *R. alkalisoli* and the candidate species isolated by Zakhia et al. [47] and Li et al. [13] form the "*R. galegae* complex". The phylogenetic position of *R. galegae* has not been stable in single gene analyses, since it was shown to group in the *Agrobacterium* clade in a 16S rRNA gene tree and clustered within *Rhizobium* in a *dnaK* gene tree [4,45]. Moreover, it formed a clade with *A. vitis* in a 23S rRNA gene tree [23], was separate in a *glnA* gene tree and a 16S rRNA gene tree of 160 proteobacterial species [15,37], and clustered with *R. leguminosarum* and *R. etli* in a *glnII* gene tree [37]. These taxonomic issues have suggested the need for more inclusive taxonomic studies in order to clarify the situation of the isolates.

Polyphasic taxonomy (combination of phenotypic, genotypic and phylogenetic data) was considered as a developed bacterial taxonomy 40 years ago. However, the genomes of rhizobia may lose or gain (at high frequency) plasmids or genomic islands bearing genes governing catabolic capacities. Hence, performing phenotypic tests, mostly for utilization of carbon and nitrogen, may not be that informative for rhizobial taxonomy [21]. The 16S rRNA gene is highly conserved, and it has been widely applied in taxonomic and phylogenetic studies of bacteria and especially rhizobia. On the other hand, it is too conserved to allow separation of closely related species [20]. Nevertheless, other housekeeping genes involved in physiological maintenance in prokaryotes are conserved and distributed throughout their genomes. Therefore, multilocus sequence analysis (MLSA) of housekeeping genes is thought to be a more powerful approach for resolving some of the taxonomic issues. The concatenation of at least five protein-encoding housekeeping genes is recommended for MLSA in order to result in a robust phylogeny. Protein-coding accessory genes are not considered as proper markers in taxonomic studies since they are specialized for ecological adaptation and thus might have been acquired independently and could have evolved separately [19,30,39,40].

In this study, the aim was (1) to investigate the phylogenetic relatedness of the "*R. galegae* complex" and *Agrobacterium* members, as well as other rhizobia, (2) to resolve the unascertained phylogenetic status of the "*R. galegae* complex", and (3) to clarify nomenclatural aspects of *Agrobacterium* and the "*R. galegae* complex". Therefore, MLSA was performed on 114 rhizobial taxa, and six protein-coding housekeeping genes were selected based on the criteria of their universal distribution, their location on the chromosome, their high conservation and unique occurrence in the genome: *atpD* (ATP synthase F1, beta subunit), *glnA* (glutamine synthetase type I), *glnII* (glutamine synthetase type II), *recA* (recombinase A), *rpoB* (RNA polymerase, beta subunit), and *thrC* (threonine synthase).

Materials and methods

Bacterial strains and DNA preparation

Bacterial strains representing the genera *Agrobacterium* and *Rhizobium* were obtained from different culture collections, as detailed in Table S1. The strains were cultured in 5 mL tryptone-yeast extract (TY) broth at 28 °C for two days, and were then grown on yeast mannitol agar (YMA) at 28 °C for 48 h [29]. Single colonies of the bacteria were cultured in 5 mL TY broth for subsequent preservation in 20% (v/v) glycerol at −80 °C. The UltraClean Microbial DNA Isolation Kit (MO BIO Laboratories, Inc.) was used for DNA extraction from the samples, and DNA preparations were stored at −20 °C.

Amplification and sequencing of the genes

The sequences (>1300 bp) of 16S rRNA genes from 51 strains of the genera *Agrobacterium*, *Allorhizobium*, *Azorhizobium*, *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium* and *Ensifer* (syn. *Sinorhizobium*) were obtained from GenBank (http://www.ncbi.nlm.nih.gov/genbank; Table S1). The primers used for amplification of the *atpD*, *glnA*, *glnII*, *recA*, *rpoB*, and *thrC* genes in the present study are listed in Table 1. The protein-coding housekeeping genes were amplified by means of a DNA Engine Thermocycler (Bio-Rad Laboratories, Inc.) following the instructions provided by Finnzymes (Thermo Fisher Scientific Inc.) in a 50 μL reaction mixture containing 1 μL of DNA template (prepared as above), 37.0 μL MilliQ water, 0.25 μL (25 pmol) of each primer, 0.5 μL (1 U) Phusion DNA polymerase, 10 μL HF Buffer, and 1 μL (10 mM) dNTPs. The quality and size of PCR products were checked by electrophoresis of 5 μL amplicons on 1.5% agarose gels stained with ethidium bromide. Some of the PCR products were extracted by the Qiaquick Gel Extraction Kit (Qiagen Inc.). The amplicons were stored at −20 °C. The obtained PCR products were purified with AMPure XP (Beckman Coulter, Inc.), sequenced using BigDye Terminator Chemistry v.3.1 and analyzed on an ABI 3130×l Sequencer (Life Technologies) at the Sequencing and Genomics Lab., Institute of Biotechnology, University of Helsinki.

The sequences were edited in GAP4 [31], and were blasted in NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastHome). Furthermore, the sequences of the six housekeeping genes from 19 strains of agrobacteria were obtained from the Laboratoire d'Ecologie Microbienne, Lyon University (France). The sequences of the corresponding loci of 14 whole-genome sequenced rhizobia were also retrieved from NCBI.

Phylogenetic analyses of the genes

All the sequences of the housekeeping genes were aligned by ClustalW [11], as executed by BioEdit version 7.0.5.3 [10], and MUSCLE [5] software at EML-EBI [9] was used for the 16S rRNA gene. Readseq-biosequence was used to convert the format of the sequences from Fasta to Nexus (http://www.ebi.ac.uk/Tools/sfc/readseq/). The best-fit models of nucleotide substitution were selected by Akaike information criterion applied in MEGA5, jModelTest 0.1.1 and FindModel [22,34; http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html]. The 16S rRNA sequences were analyzed using the neighbor-joining method in the MEGA5 program. The gene trees of six housekeeping genes and the phylogenetic hypothesis of the combined genes of 114 taxa were constructed using Bayesian inferences. Bayesian analyses of each single locus and the combined loci were run twice applying the Metropolis coupled Markov chain Monte Carlo (MCMCMC) algorithm for 10 million generations for the single genes, and 20 million generations for the combined genes with MrBayes 3.2 [28]. The most appropriate model of evolution for each locus was used to construct the combined gene tree. The results of the runs were analyzed by Tracer v1.5, and the phylogenetic trees were visualized and annotated with FigTree v1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/).