Systematic and Applied Microbiology 35 (2012) 353–358

Contents lists available at SciVerse ScienceDirect

Systematic and Applied Microbiology

journal homepage: www.elsevier.de/syapm

Rhizobium etli taxonomy revised with novel genomic data and analyses

Martha G. López-Guerrero[a], Ernesto Ormeño-Orrillo[a], Encarna Velázquez[b], Marco A. Rogel[a], José Luis Acosta[a,1], Victor Gónzalez[a], Julio Martínez[a], Esperanza Martínez-Romero[a,*]

[a] *Centro de Ciencias Genómicas, UNAM, Cuernavaca, Morelos, Mexico*
[b] *Departamento de Microbiología y Genética, Universidad de Salamanca, Salamanca, Spain*

ARTICLE INFO

Article history:
Received 8 March 2012
Received in revised form 14 May 2012
Accepted 14 June 2012

Keywords:
Nitrogen fixation
Legume nodulation
Phaseolus vulgaris

ABSTRACT

The taxonomic position of *Phaseolus vulgaris* rhizobial strains with available sequenced genomes was examined. Phylogenetic analyses with concatenated conserved genomic fragments accounting for over half of each genome showed that *Rhizobium* strains CIAT 652, Ch24-10 (newly reported genome) and CNPAF 512 constituted a well-supported group independent from *Rhizobium etli* CFN 42^T. DNA–DNA hybridization results indicated that CIAT 652, Ch24-10 and CNPAF 512 could correspond to *R. etli*, although the hybridization values were at the borderline that distinguishes different species. In contrast, experimental hybridization results were higher (over 80%) with *Rhizobium phaseoli* type strain ATCC 14482^T in congruence to phylogenetic and ANIm analyses. The latter criterion allowed the reclassification of *R. etli* strains 8C-3 and Brasil5 as *R. phaseoli*. It was therefore concluded, based on all the evidence, that the CIAT 652, Ch24-10, and CNPAF 512 strains should be reclassified as *R. phaseoli* in spite of several common features linking them to *R. etli*. The *R. phaseoli* and *R. etli* speciation process seems to be a more recent event than the speciation that has occurred among other sister species, such as *R. leguminosarum*–*R. etli* or *R. rhizogenes*–*R. tropici*.

Introduction

The advent of the genomic era has provided both a plethora of molecular markers useful in taxonomy and the possibility to compare whole genomes instead of a few genes. Toward this end, novel algorithms and parameters have been proposed to compare genomes for taxonomic purposes. ANI (average nucleotide identity) has been defined as a very useful parameter to delineate different species and it correlates with DNA–DNA hybridization (DDH) [12,19,32]. Based upon a large set of experimental results from diverse bacteria, thresholds of ANI (94–96%) have been recommended for distinguishing species.

Phaseolus vulgaris (common bean) is the legume grain most consumed for human nutrition and, like other legumes, it forms symbiosis with nitrogen-fixing bacteria. *P. vulgaris* symbiotic bacteria have been widely studied [4,6,13,16,22,24,31,34,37,43] and this legume has become a model for studying nodule-bacterial diversity from plants grown in diverse conditions or geographical regions where bean is native or introduced. In its sites of origin and in some introduced areas, *Rhizobium etli* has been reported as the dominant *P. vulgaris* bean nodule bacterium identified on the basis of 16S rRNA gene sequences [22]. In addition, core gene sequences have been used to characterize nodule isolates, thus providing a better phylogenetic resolution and revealing that other species besides *R. etli* can also represent a significant fraction of the bean nodule occupants [4,13,34,37]. Recently, *R. etli* type strain CFN 42^T was found to have a low recombination with *R. etli* CIAT 652 and other *P. vulgaris* isolates [2]. Richter and Rosselló-Móra [32] calculated the ANI of CFN 42^T and CIAT 652, and reported that CIAT 652 was not a member of *R. etli*. DDH results were not available for comparison and the species affiliation of CIAT 652 was not identified. At the same time, *Rhizobium phaseoli* was re-recognized as a valid and different species from *R. etli* because it presented divergent core genes and experimental DNA–DNA relatedness values significantly lower than 70% with *R. etli* CFN 42^T [30]. *R. etli* and *R. phaseoli* are sympatric species nodulating bean and both have been found recently in Ethiopia [4]. On a phylogenetic basis, using *recA*, *atpD* and *celC* partial gene sequences, CIAT 652 was recognized as belonging to *R. phaseoli* [34] but this taxonomic affiliation was not otherwise confirmed by DDH analysis. Based on the same phylogenetic analysis, the Mim2 strain (a *Mimosa affinis* isolate) was also recognized as *R. phaseoli*, however, DDH and multilocus enzyme electrophoresis (MLEE) placed it within *R. etli* [44]. Clearly, there were conflicting data. Further, it was recommended that "the taxonomic status of the strains currently named *R. etli*

* Corresponding author.
E-mail address: esperanzaeriksson@yahoo.com.mx (E. Martínez-Romero).
[1] Present address: Instituto Nacional de Medicina Genómica, México DF, Mexico.

0723-2020/$ – see front matter © 2012 Elsevier GmbH. All rights reserved.
http://dx.doi.org/10.1016/j.syapm.2012.06.009

Table 1
Rhizobium etli and *R. phaseoli* strains used in this study.

Strains	Host	Origin	Reference
R. phaseoli			
Ch24-10	*Zea mays* and *Phaseolus vulgaris*	Puebla, Mexico	[35]
CIAT 652	*P. vulgaris*	Buitrera, Colombia	[42]
CNPAF 512	*P. vulgaris*	Brazil	[9,23]
ATCC 14482^T	*P. vulgaris*	Beltsville, Maryland	ATCC
R. etli			
CFN 42^T	*P. vulgaris*	Guanajuato, Mexico	[29]

should be revised" [34]. Consequently, our aim in this study was to revise the taxonomic status of different *R. etli* strains with available sequenced genomes.

Materials and methods

Strains, growth and DNA extraction

Strains used in the DDH studies are shown in Table 1. They were grown in 5 mL liquid PY medium for DNA extraction. Ch24-10 was grown on PY plates and fresh cultures were grown in 50 mL liquid PY. DNA was extracted by the DNA Isolation Kit for Cells and Tissues (Roche, USA). *Rhizobium* strains were maintained in YM with glycerol at −70 °C.

Phylogenetic analysis

rpoB primers and PCR conditions were as described previously [21,25]. PCR sequences were compared to sequences obtained from whole genomes. Sequence alignments were generated and edited with BioEdit 7 [14]. Percentage identity between sequences was obtained after removing all columns with gaps from the alignments. Best-fit models of sequence evolution were selected for each gene with JModelTest 0.1.1 using the Akaike information criterion [28]. Maximum likelihood (ML) and neighbor-joining (NJ) phylogenies were constructed with Mega 5 [39]. Support for tree nodes was evaluated by bootstrap analysis with 100 or 1000 pseudoreplicates for ML and NJ, respectively.

Genomic sequencing

The genomic sequence from strain Ch24-10 was obtained using two platforms: the Roche 454 pyrosequencing system (350 bp from 3K long-tag paired end sequencing protocol in the Genome Sequencer FLX) using a commercial service, and the Illumina technology (*Genome Analyzer GAIIx*, paired-end protocol, 200 base pairs-inserts library, reads with 36 nucleotides in length) at the Unidad Universitaria de Secuenciación Masiva de DNA (USMDNA) of the Universidad Nacional Autónoma de México (UNAM).

Genome assembly

Roche reads were *de novo* assembled using GSassembler, Newbler version 2.5.3 with default parameters. Illumina reads were assembled using TAIPAN [36] and then SSAKE [45] separately, both with default parameters. The contigs generated by these programs were reassembled using minimus2 [http://sourceforge.net/apps/mediawiki/amos/index.php?title=Minimus2].

Both assemblies, the one generated with Newbler and the one generated with minimus2 (TAIPAN + SSAKE), were reassembled again to generate a hybrid assembly with minimus2.

Prediction and annotation of the Ch24-10 genome

The contigs generated with the minimus2 assembler were assembled in scaffolds with ABACAS [5] using the complete genomes of CFN 42^T and CIAT 652 as anchors. Next, the pseudochromosome and pseudoplasmids were constructed for each replicon armed with ABACAS, by adding the following sequence "NNNNCATTCCATTCATTAATTAATTAATGAATGAATGNNNNN" (containing the six open reading frames) at the 5′ end of each contig [40]. Prediction and annotation of genes for each pseudoreplicon were undertaken with the CG-Pipeline program [17]. The draft genome sequence of *Rhizobium* Ch24-10 obtained by hybrid assembly using sequences derived by Illumina and 454 Roche technologies was then used. An 80× genome coverage was obtained. The results of this Whole Genome Shotgun project have been deposited at DDBJ/EMBL/GenBank under the accession AHJU00000000. The version described in this paper is the first version, AHJU01000000.

Genome tree construction

Regions shared among *Rhizobium* strains CIAT 652, CFN 42^T, Ch24-10, and CNPAF 512, *Rhizobium leguminosarum* sv. viciae 3841 and *Sinorhizobium meliloti* 1021 were identified using Mugsy [3] with the following parameters: minlength = 30, distance = 100, duplications 1, and fullsearch refine. Each orthologous region was extracted by an *ad hoc* Perl script (homemade) and filtered for long gaps with trimAl [8] with the parameter -automated1. All regions were concatenated to build a genome tree by the ML and neighbor-net network methods. ML was performed using RaxML [38] with the GTR + I + G nucleotide substitution model, 1000 distinct randomized maximum parsimony trees and the parameters: p 12345, e 0.0000001, c 8, j STRICT, and k 1000. Finally, the neighbor-net was created with the Splits Tree 4 program [15]. Other genome comparisons were performed as described [20].

ANIm

ANIm values were calculated using the JSpecies package [32].

DNA–DNA hybridization (DDH)

The procedure was as described in [41,44]. DNA was quantified with NanoDrop 2000 (Thermo Scientific) and in gels. DNA was digested with the *Eco*R1 restriction enzyme and electrophoresis was performed in 1% agarose gels. Only lanes with homogeneous DNA quantities were hybridized in Southern blot experiments to total DNA from reference strains with probes labeled with [α-32P] dCTP using RediPrimeTM II (GE Healthcare). Rapid-hyb buffer was used for hybridization and washings were carried out using 2× to 1× SSC with 0.1% SDS at 65 °C.

Filters were cut and individual lanes were counted in scintillation liquid in a LS6500 multi-purpose scintillation counter (Beckman Coulter). Three independent experiments were performed with similar results.

Results

Phylogenetic and similarity gene analysis

R. etli and *R. phaseoli* strains had highly similar 16S rRNA genes (over 99.3% identical) but were slightly separated in a NJ phylogenetic tree (Fig. 1). Ch24-10, CNPAF 512 and CIAT 652 strains reported as *R. etli* were found to group with *R. phaseoli* in the phylogenetic analysis of housekeeping genes *recA*, *atpD* and *rpoB* (Fig. 2). Furthermore, the genomic-based phylogenetic analyses