ELSEVIER

System. Appl. Microbiol. 27, 703–706 (2004)
http://www.elsevier.de/syapm

SYSTEMATIC AND APPLIED MICROBIOLOGY

Phylogenetic Relationships of Rhizobia Based on Citrate Synthase Gene Sequences

Ismael Hernández-Lucas[1], Marco Antonio Rogel-Hernández[1], Lorenzo Segovia[2], Keilor Rojas-Jiménez[3], and Esperanza Martínez-Romero[1]

[1] Centro de Investigación sobre Fijación de Nitrógeno, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México
[2] Departamento de Ingeniería Celular y Biocatálisis, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México
[3] FB Biologie der Philipps-Universität, Marburg, Germany

Received: May 4, 2004

Summary

Partial nucleotide sequences of the citrate synthase (*gltA*) gene from different rhizobia genera were determined. Tree topologies based on this housekeeping gene were similar to that obtained using 16S rRNA sequences. However *gltA* appeared to be more reliable at determining phylogenetic relationships of closely related taxa. We propose *gltA* sequences as an additional tool to be used in molecular phylogenetic studies.

Key words: Phylogeny – *gltA* – 16S rRNA

Introduction

The members of the genus *Rhizobium*, *Bradyrhizobium*, *Mesorhizobium*, *Azorhizobium*, and *Sinorhizobium*, collectively named rhizobia, have been classified on the basis of polyphasic taxonomy [17]. This approach uses phenotypic and genotypic characteristics and includes 16S rRNA gene sequence determination. Hitherto, the analysis of 16S rRNA genes represented the most powerful method for investigating the phylogenetic relationships among microorganisms. However, this gene has limitations as a taxonomic tool such as slow evolution, genetic crossing-over and, because it is a highly conserved molecule, may not be useful for discriminating closely related bacterial species. Furthermore, the presence of a 16S rRNA gene as part of a complete ribosomal operon has been found on a plasmid in *Bacillus megaterium* [14] and allele differences have been detected in 16S rRNA genes, making these genes a not infallible guide to define evolutionary relationships [31]. Studies of additional genes are required to know more about the evolution of the genome and improve the knowledge of the phylogenetic relationships of rhizobia.

The citrate synthase gene (*gltA*) is present in almost all living cells. It contains conserved and variable regions and its product is the first enzyme in the tricarboxylic acid cycle, and is thus considered a key regulator of intracellular ATP production in both prokaryotic and eukaryotic cells. Based on these characteristics *gltA* gene sequences have been used for estimating phylogenies among some bacterial groups [1, 2, 10]. In the present work we study the usefulness of *gltA* gene sequences for determining phylogenetics relationships among rhizobia and compare it to phylogenies obtained with 16S rRNA gene sequences.

Rhizobia strains were grown in PY medium and *Escherichia coli* strains were cultivated in LB medium. Plasmid purification, genomic DNA extraction and Southern blotting were performed according to published protocols [23]. For sequencing, double stranded DNA was purified and sequencing was performed in an automatic ALF DNA sequencer (Pharmacia Biotech Uppsala) or in an automatic Perkin Elmer/Applied Biosystems 377-18 sequencer. *gltA* genes from *Agrobacterium* sp. Ch-Ag-4 (**AY094145**), *Agrobacterium* sp. K-Ag-3 (**AY094144**), *Rhizobium etli* CFN42 (**AY094142**), *Rhizobium galegae* HAMBI540 (**AY094149**), *Rhizobium* sp. CFN234 (**AY094151**), *Sinorhizobium terangae* USDA4102 (**AY094150**), *Mesorhizobium mediterraneum* USDA3392 (**AY094148**), *Mesorhizobium huakuii* CCBAU2609

0723-2020/04/27/06-703 $ 30.00/0

*(***AY094147**), and *Azorhizobium caulinodans* ORS571 (**AY094143**) were amplified by PCR with forward primers **P231-101** (5′AGAAGAAGGACTTCGACTA3′) or **512–MAP** (5′TACAAGTACCATATCGGCCAGCCC T3′) and reverse primer **OP85-1R** (5′CATGGTCGTGG GGAAGCC3′). They correspond respectively to base positions 299–317, 529–553, and 1123–1140 of the chromosomal *Rhizobium tropici ccsA* gene [7]. Primer pair **512-MAP** and **OP85-1R** was used to amplify 590 bp from *Azorhizobium caulinodans* ORS571 and *Mesorhizobium huakuii* USDA 4779. Primer pair **P231-101** and **OP85-1R** was used to amplify 805 bp for the remaining strains used in this study. PCR fragments were obtained using the following temperature program: initial denaturation at 93 °C for 3 min, then 34 cycles of annealing at 48 °C for 2 min, extension at 72 °C for 2 min and denaturation at 93 °C for 2 min. PCR products were purified from agarose gels with the Geneclean kit (BIO101). PCR products were cloned into the sequencing vector pMos-Blue T (Amersham Life Sciences). In this work we also included previously reported *gltA* sequences of *Rhizobium tropici* CFN299 (L41815, Z34516; [7, 20]) *Sinorhizobium fredii* (AY157738), *Sinorhizobium. meliloti* 1021 (SMc02087; [4]), *Brucella melitensis* 16M (BMEI0836; [3]), *Brucella suis* 1330 (AE014415; [21]), *Agrobacterium tumefaciens* C58 (Atu1392; [27]), *Mesorhizobium loti* MAFF303099 (m1r0629; [11]), *Bradyrhizobium japonicum* (blr4839; [12]), and *Rhodopseudomonas palustris* (BX572602; [15]). Gene accession numbers are shown in parenthesis, sequences in bold were determined in this study.

Phylogenetic analysis of *gltA* genes was performed. Partial nucleotide sequences (568 bp) of *gltA* genes were first translated into amino acid sequences, aligned using CLUSTAL W [24] and the nucleotide sequence were aligned against aligned amino acid sequences by using DAMBE [29] and optimised by hand with BIOEDIT [6]. Different analyses to assay data robustness of *gltA* sequences corroborate that our dataset showed a strong phylogenetic signal and there is no evidence of saturation. Nucleotide substitution rate over site analysis confirmed a homogeneous distribution of the variation along the sequence using all three codon positions, 1 and 2 positions and 3rd positions only. Substitution saturation test [28] showed no significant saturation ($Iss < Iss.c$, $p < 0.01$) when all three codon positions, 1 and 2 positions and 3rd positions were analysed. A plot of the number of transitions and transversion versus genetic distances confirmed this result (data not shown). Molecular phylogenies were constructed using the neighbour joining (NJ) method [22] with MEGA 2 software [13]. Statistical tests of nucleotide frequencies and transition/transversion substitution rates confirmed Tamura-Nei (TN93) as an accurate distance-correction model for our datasets. Maximum-parsimony (MP) analysis was also performed. Statistical support of branches in the phylogeny was calculated with 1000 bootstrap replications. Very similar topologies between neighbour joining and MP phylogenies were observed. Identical groups were identified particularly for those clades with a strong bootstrap support.

Analysis of GltA protein. A multiple aligment of partial GltA protein sequences (189 AA) was constructed using t_coffee [18]. We then carried out a Bayesian estimation of the phylogeny for this region with the program mrBayes 3.0b4 [9] using 250000 generations, sampling and printing frequencies of 100 and 4 chains for the

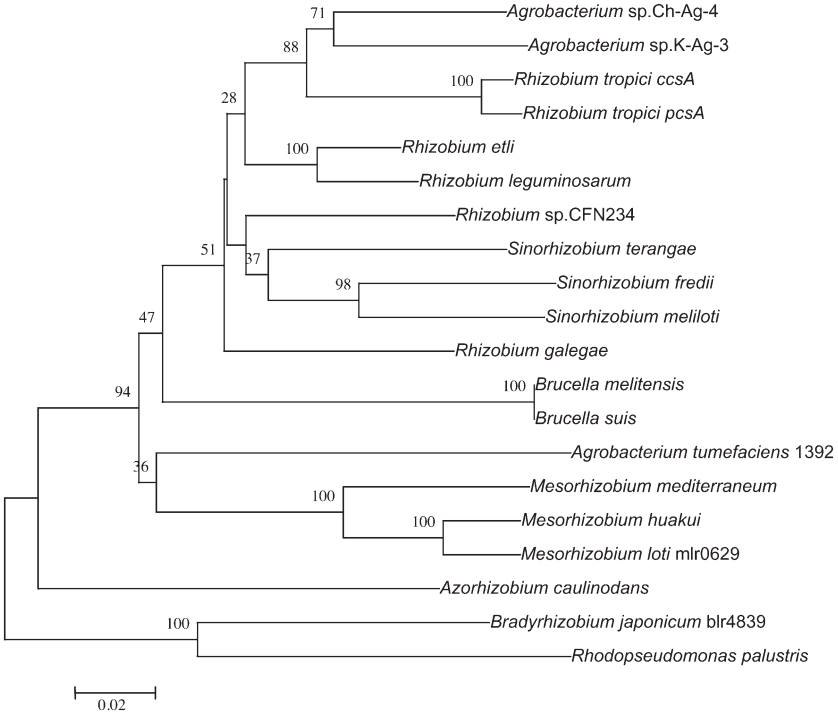

Fig. 1. Phylogenetical relationships of *Rhizobium*, *Sinorhizobium, Mesorhizobium*, *Bradyrhizobium*, *Azorhizobium*, *Agrobacterium, Brucella* and related species based on *gltA* gene sequences. Phylogenetic trees were obtained by neighbour-joining analysis. TN93 distance correction was applied. Percentage bootstrap support (1000 replicates) is indicated at branching points.

Markov Chain Monte Carlo analysis thus generating 2500 trees. As stationarity was reached at around 500 cycles, we used a burnin of 1500 trees so that only 1000 tree samples were used for the determination of a consensus tree and the determination of posterior probabilities for the trees. The consensus tree was displayed with TREEVIEW [19]. We found that topologies generated with DNA or protein showed very similar phylogenetic relationships of rhizobia.

The analysis of 590 bp of *gltA* genes from rhizobia, shows that the sequence identity of this fragment ranged from 75% to 95%, thus the sequence of *gltA* is much variable among *Rhizobium* species than are 16S rRNA genes (87% to 99%). The 590 bp analysed also showed 261 variable sites and 195 parsimony informative sites, as well as conservation of 16 NADH binding residues and 5 active site residues. Phylogenies of rhizobia based on *gltA* genes sequences (Fig. 1) showed distinct phylogenetic linages. *R. tropici* was clearly separated from *Agrobacterium* sp K-Ag-3 and Ch-Ag-4 with the identity among those strains being 90%. Previously, Willems et al showed that by 16S rRNA sequences, *R. tropici* and *Agrobacterium* spp were highly related, with an identity of 99%. In this, case *gltA* sequences appeared to be more reliable at determining relationships among *Rhizobium* and *Agrobacterium* spp than sequences of 16S rRNA. Furthermore total DNA hibridization experiments showed that *R. tropici* and *Agrobacterium* sp Ch-Ag-4 were not highly related [16], these data support the relationships based on *gltA* genes sequences. Previously we reported [8] that *R.* sp CFN234 belonged to *R. etli* linage (based on 16S rRNA gene sequences), however by phylogentic analysis of *gltA* genes, *R.* sp CFN234 is separated from *R. etli*. Our results with *gltA* support data generated with total DNA hybridisation experiments, which show a low level of homology between *R. etli* and *R.* sp CFN234 [16]. Phylogenetic trees generated with *gltA* showed a close relationships among *M. mediterraneum, M. huakui* and *M. loti*. In this cluster the *gltA* gene of *M. huakuii* is closely related to the *gltA* (mlr0629) of *M. loti* MAFF303099. This observation confirms previous data which suggest that *M. loti* MAFF303099 might belong to *M. huakuii* [25]. *R. galegae* is normally found clustered to *Agrobacterium* rather than *Rhizobium* in the 16S rRNA phylogeny, however phylogentic relationships with *gltA* genes clearly separate *R. galegae* from *A. tumefaciens*. This finding is in agreement with phylogenies obtained with *atpD* or *recA*, [5] and so the use of three molecular markers support *R. galegae* being as distant from *Agrobacterium* as it is from rhizobia. Phylogenetic relationships of *R. etli*, *R. leguminosarum*, *S. terangae*, *S. fredii*, *S. meliloti*, *B. melitensis*, *B. suis*, *B. japonicum*, *A. caulinodans* and *R. palustis* derived from *gltA* sequences are congruent to those obtained with 16S rRNA genes sequences.

In previous studies we reported the presence of two *gltA* genes in *Rhizobium tropici* [7]. To determine if the presence of multiples copies is common in rhizobia, hybridisation experiments were performed. Total DNA of the *Rhizobium* strain analysed were purified and digested with two different restriction enzymes. Our results showed that only *Rhizobium tropici* contain two *gltA* genes, therefore the presence of multiples *gltA* genes is not general in rhizobia. *gltA* gene sequences have been used in several bacterial groups for taxonomy. In *Bartonella,* phylogenetic trees reflect a much higher sequence diversity in *gltA* genes than that observed for the 16S rRNA genes [2]. In *Ehrlichia* the architecture of the trees constructed with *gltA* was similar to that derived from the 16S rRNA gene sequences but showed more significant bootstrap values [10]. In the case of rhizobia, we found that *gltA* is useful for estimating evolutionary relationships of closely related taxa, We propose *gltA* sequences as an additional tool for molecular phylogenetic studies.

Acknowledgements

We thank J. A. Ramírez-Trujillo, M. A. Gaitan, J. Martínez, J. Caballero-Mellado, P. Mavingui, Entao Wang and M. Dunn, for technical help and discussion, as well as J. P. W. Young and Harry Duckworth for useful scientific comments to improve the manuscript.

References

1. Birtles, R. J., Raoult, D.: Comparison of partial citrate synthase gene (*gltA*) sequences for phylogenetic analysis of *Bartonella* species. Int. J. Syst. Bacteriol. 46, 891–897 (1996).
2. Dehio, C., Lanz, C., Pohl, R., Behrens, P., Bermond, D., Piémont, Y., Pelz, K., Sander, A.: *Bartonella schoenbuchii* sp. nov., isolated from the blood of wild roe deer. Int. J. Syst. Bacteriol. 51, 1557–1565 (2001).
3. DelVecchio, V. G., Kapatral, V., Redkar, R. J., Patra, G., Mujer, C., Los, T., Ivanova, N., Anderson, I., Bhattacharyya, A., Lykidis, A., Reznik, G., Jablonski, L., Larsen, N., D'Souza, M., Bernal, A., Mazur, M., Goltsman, E., Selkov, E., Elzer, P. H., Hagius, S., O'Callaghan, D., Letesson, J. J., Haselkorn, R., Kyrpides, N., Overbeek R.: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*. Proc. Natl. Acad. Sci. U.S.A. 99, 443–448 (2002).
4. Galibert, F., Finan, T. M., Long, S. R., Pühler, A., Abola, A., Ampe, F., Barloy-Hubler, F., Barnett, M. J., Becker, A., Boistard, P., Bothe, M., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R. W., Dréano, S., Federspiel, N. A., Fisher, R. F., Gloux, S., Godrie, T., Goffeau, A., Golding, B., Gouzy, J., Gurjal, M., Hernández-Lucas, I., Hong, A., Huizar, L., Hyman, R. W., Jones, T., Kahn, D., Kahn, M. L., Kalman, S., Keating, D. H., Kiss, E., Komp, C., Lelaure, V., Masuy, D., Palm, C., Peck, M. C., Pohl, T. M., Portetelle, D., Purnelle, B., Ramsperger, U., Surzycki, R., Thébault, P., Vandenbol, M., Vorhölter, F. J., Weidner, S., Wells, D. H., Wong, K., Chen Yeh, K., Batut, J.: The composite genome of the legume symbiont *Sinorhizobium Meliloti*. Science 293, 668–72 (2001).
5. Gaunt, M. W., Turner, S. L., Rigottier-Gois, L., Lloyd-Macgilp, S. A., Young, J. P.: Phylogenies of *atpD* and *recA* support the small subunit rRNA-based classification of rhizobia. Int. J. Syst. Evol. Microbiol. 51, 2037–48 (2001).
6. Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.: Nucl. Acids. Symp. Ser. 41, 95–98 (1999).

7. Hernández-Lucas, I., Pardo, M., Segovia, L., Miranda, J., Martínez, E.: *Rhizobium tropici* chromosomal citrate synthase gene. Appl. Environ. Microbiol. 61, 3992–3997 (1995a).
8. Hernández–Lucas, I., Segovia, L., Martínez-Romero, E., Pueppke, S. G.: Phylogenetic relationships and host range of *Rhizobium* spp. that nodulate bean, *Phaseolus vulgaris* L. Appl. Environ. Microbiol. 61, 2775–2779 (1995b).
9. Huelsenbeck, J. P., Ronquist F.: MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754–755 (2001).
10. Inokuma, H., Brouqui, P., Drancourt, M., Raoult, D.: Citrate synthase gene sequence: a new tool for phylogenetic analysis and identification of *Ehrlichia*. J. Clin. Microbiol. 39, 3031–3039 (2001).
11. Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S., Watanabe, A., Idesawa, K., Ishikawa, A., Kawashim, K., Kimura, T., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimpo, S., Sugimoto, M., Takeuchi, C., Yamada, M., and Tabata, S.: Complete genome sequence of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. DNA Res. 7, 331–338 (2000).
12. Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Idesawa, K., Iriguchi, M., Kawashima, K., Kohara, M., Matsumoto, M., Shimpo, S., Tsuruoka, H., Wada, T., Yamada, M. and Tabata, S.: Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110.: DNA Research 9, 189–197 (2002).
13. Kumar, S., Tamura, K., Jakobsen, I., Nei, M.: MEGA2: Molecular evolutionary genetic análisis. Bioinformatics 17, 1244–1255 (2001).
14. Kunnimalaiyaan, M., Stevenson, D. M., Zhou, Y., Vary, P. S.: Analysis of the replicon region and identification of an rRNA operon on pBM400 of *Bacillus megaterium* QM B1551. Mol. Microbiol. 39, 1010–1021 (2001).
15. Larimer, F. W., Chain, P., Hauser, L., Lamerdin, J., Malfatti, S., Do, L., Land, M. L., Pelletier, D. A., Beatty, T. J., Lang, A. S., Tabita, F. R., Gibson, J. L., Hanson, T. E., Torres y Torres, J., Peres, C., Harrison, F. H., Gibson, J., Harwood, C. S.: Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris*. Nat. Biotechnol. 22, 55–61 (2004).
16. Martínez-Romero E.: Recent developments in *Rhizobium* taxonomy. Plant and Soil. 161, 11–20 (1994).
17. Martínez-Romero, E., Caballero-Mellado, J.: *Rhizobium* phylogenies and bacterial genetic diversity. Crit. Rev. Plant. Sci. 15, 113–140 (1996).
18. Notredame, C., Higgins, D., Heringa, J.: T-Coffee: A novel method for multiple sequence alignments. J. Mol. Biol. 302, 205–217 (2000).
19. Page, R. D. M.: TREEVIEW: An application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 12, 357–358 (1996).
20. Pardo, M. A., Lagunez, J., Miranda, J., Martínez, E.: Nodulating ability of *Rhizobium tropici* is conditioned by a plasmid-encoded citrate synthase. Mol. Microbiol. 11, 315–321 (1994).
21. Paulsen, I., Seshadri, R., Nelson, K. E., Eisen, J. A., Heidelberg, J. F., Read, T. D., Dodson, R. J., Umayam, L. A., Brinkac, L. M., Beanan, M. J., Daugherty, S. C., Deboy, R., Durkin, A. S., Kolonay, J. F., Madupu, R., Nelson, W. C., Ayodeji, B., Kraul, M., Shetty, J., Malek, J. A., Van Aken, S. E., Riedmuller, S., Tettelin, H., Gill, S., White, O., Salzberg, S. L., Hoover, D. L., Lindler, L., Halling, S. M., Boyle, S. M., Fraser, C. M.: The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts. Proc. Natl. Acad. Sci. U.S.A. 99, 13148–13153 (2002).
22. Saitou, N., Nei, M.: The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425 (1987).
23. Sambrook, J., Fritsch, E. F., Maniatis, T.: Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y (1989).
24. Thompson, J. D., Higgins, D. G. and Gibson, T. J.: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic. Acids. Research. 22, 4673–4680 (1994).
25. Turner, S. L., Zhang, X. X., Di Li, F., Young J. P. W.: What does a bacterial genome sequence represent? Mis-assignment of MAFF 303099 to the genospecies *Mesorhizobium loti*. Microbiology 148, 3330–3331 (2002).
26. Willems, A., Collins, M. D.: Phylogenetic analysis of rhizobia and agrobacteria based on 16S rRNA gene sequences. Int. J. Syst. Bacteriol. 43, 305–313 (1993).
27. Wood, D. W., Setubal, J. C., Kaul, R., Monks, D., Chen, L., Wood, G. E., Chen, Y., Woo, L., Kitajima, J. P., Okura,V. K., Almeida Jr., N. F., Zhou, Y., Bovee Sr., D., Chapman, P., Clendenning, J., Deatherage, G., Gillet, W., Grant, C., Guenthner, D., Kutyavin, T., Levy, R., Li, M., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z., Gordon, D., Eisen, J. A., Paulsen, I., Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z., Dolan, M., Tingey, S. V., Tomb, J., Gordon, M. P., Olson, M. V. and Nester, E. W.: The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. Science 294, 2317–2323 (2001).
28. Xia, X., Xie, Z., Salemi, M., Chen, L., Wang, Y.: An index of substitution saturation and its application. Molecular Phylogenetics and Evolution. 26, 1–7 (2003).
29. Xia, X., Xie, Z.: DAMBE: data analysis in molecular biology and evolution. J. Hered. 92, 371–373 (2001).
30. Xia, X.: Data analysis in molecular biology and evolution. Kluwer Academic Publishers. Boston/Dordrecht/London. (2000).
31. Young, J. P. W., Haukka, K.: Diversity and phylogeny of rhizobia. New. Phytol. 133, 87–94 (1996).

Corresponding author:
Ismael Hernández-Lucas, Centro de Investigación sobre Fijación de Nitrógeno, Universidad Nacional Autónoma de México, Apdo. Postal 565-A, Cuernavaca, Morelos, México
Tel.: ++777 3-13-16-97; Fax: ++777 3-17-55-81;
e-mail: ismael@cifn.unam.mx