

## Discordant Phylogenies within the *rrn* Loci of Rhizobia†

Peter van Berkum,<sup>1\*</sup> Zewdu Terefeework,<sup>2</sup> Lars Paulin,<sup>3</sup> Sini Suomalainen,<sup>3</sup>  
Kristina Lindström,<sup>2</sup> and Bertrand D. Eardly<sup>4</sup>

Soybean Genomics and Improvement Laboratory, Agricultural Research Service, U.S. Department of Agriculture, Beltsville, Maryland 20705<sup>1</sup>; Department of Applied Chemistry and Microbiology, Biocenter 1, Viikinkaari 9,<sup>2</sup> and Institute of Biotechnology, Biocenter 1, Viikinkaari 6,<sup>3</sup> University of Helsinki, Helsinki, Finland; and Penn State Berks-Lehigh Valley College, Berks Campus, Reading, Pennsylvania 19610<sup>4</sup>

Received 25 October 2002/Accepted 3 March 2003

It is evident from complete genome sequencing results that lateral gene transfer and recombination are essential components in the evolutionary process of bacterial genomes. Since this has important implications for bacterial systematics, the primary objective of this study was to compare estimated evolutionary relationships among a representative set of  $\alpha$ -Proteobacteria by sequencing analysis of three loci within their *rrn* operons. Tree topologies generated with 16S rRNA gene sequences were significantly different from corresponding trees assembled with 23S rRNA gene and internally transcribed space region sequences. Besides the incongruence in tree topologies, evidence that distinct segments along the 16S rRNA gene sequences of bacteria currently classified within the genera *Bradyrhizobium*, *Mesorhizobium* and *Sinorhizobium* have a reticulate evolutionary history was also obtained. Our data have important implications for bacterial taxonomy, because currently most taxonomic decisions are based on comparative 16S rRNA gene sequence analysis. Since phylogenetic placement based on 16S rRNA gene sequence divergence perhaps is questionable, we suggest that the proposals of bacterial nomenclature or changes in their taxonomy that have been made may not necessarily be warranted. Accordingly, a more conservative approach should be taken in the future, in which taxonomic decisions are based on the analysis of a wider variety of loci and comparative analytical methods are used to estimate phylogenetic relationships among the genomes under consideration.

Rhizobia are nitrogen-fixing bacterial symbionts of legumes that are of economic importance in low-input sustainable agriculture, agroforestry, and land reclamation. Descriptions of phenotypic and genetic variation among rhizobia have been extensive. Currently there are six genera of rhizobia, *Allorhizobium*, *Azorhizobium*, *Bradyrhizobium*, *Rhizobium*, *Mesorhizobium*, and *Sinorhizobium*, that have been proposed (65), and there is a report of a single species of *Methylobacterium* (of the family *Methylobacteriaceae*), *Methylobacterium nodulans*, which forms a symbiosis with specific species of *Crotalaria* (59). The primary criterion by which these genera are defined is analysis of 16S rRNA gene sequence (65).

Sequencing the 16S rRNA gene has profoundly affected how relationships among the bacteria are portrayed (45). The 16S rRNA gene sequence is useful for this purpose because it is slowly evolving and the gene product is both universally essential and functionally conserved. However, basing bacterial phylogeny on 16S rRNA gene sequence variation not only presupposes that evolution throughout the genome progresses at a constant rate by mutation and Darwinian selection but also assumes that the evolution of the genome and of the 16S rRNA gene is strictly hierarchical. From a practical point of view this approach also requires each genome to harbor a single copy of the 16S rRNA gene or that multiple alleles within single genomes have identical sequences.

Although seemingly convenient for classification purposes, 16S rRNA gene sequences alone may not adequately reflect the relationships among the genomes of the microbes under consideration. For example, several reports have indicated that bacterial genomes may harbor more than a single copy of the 16S rRNA gene (2, 6, 12, 48, 67, 69). Also, there is inferential evidence that ribosomal genes in certain bacteria or archaea undergo lateral transfer and genetic recombination (13, 57, 68, 69, 72). The possibility that 16S rRNA gene sequence divergence in single cells results from lateral transfer and recombination as well as mutation has important implications for evolutionary biology in bacterial populations and also for taxonomy and nomenclature.

Therefore, the primary objective of this study was to expand the number of loci used to portray phylogenetic relationships among a representative set of  $\alpha$ -Proteobacteria with emphasis on rhizobia. To accomplish this objective, we compared the 16S rRNA sequence divergence with that of the 23S rRNA gene and the internally transcribed space (ITS) region. Because we discovered incongruence among estimated phylogenies, further analyses were done to evaluate the possibility that these results may reflect a reticulate evolutionary history of these loci.

### MATERIALS AND METHODS

**Bacterial strains.** The reference rhizobial strains used in this study were obtained from the U.S. Department of Agriculture Agricultural Research Service National *Rhizobium* Germplasm Collection, Beltsville, Md. Other bacterial species and their sources were *Blastobacter denitrificans* IFAM 1005 (LMG 8443) from the Belgian Culture Collection of Microorganisms, *Agrobacterium vitis* ATCC 49767 and *Afipia felis* ATCC 49715 from the American Type Culture Collection, *Mycoplana dimorpha* NRRL B-1091 from the Agricultural Research

\* Corresponding author. Mailing address: Soybean Genomics and Improvement Laboratory, HH-4, Bldg010, BARC-West, USDA, ARS, 10300 Baltimore Blvd., Beltsville, MD 20705. Phone: (301) 504-7280. Fax: (301) 713-0308. E-mail: pberkum@ba.ars.usda.gov.

† For a commentary on this article, see page 2975 in this issue.

Service NRRL Culture Collection, Peoria, Ill., *Agrobacterium tumefaciens* IAM 13129, *Agrobacterium rubi* IAM 13569, *Ochrobactrum anthropi* IAM 14119, and *Phyllobacterium myrsinacearum* IAM 13584 from the Institute of Applied Microbiology (IAM) Culture Collection, and *Rhodobacter sphaeroides* strain 2.4.1 and *Rhodospseudomonas palustris* strain GH, kindly provided by Michael T. Madigan.

**Growth of the bacteria and DNA isolation.** Rhizobial and agrobacterial cultures were grown in 50 ml of modified arabinose gluconate broth (62) for the large-scale isolation of DNA purified by CsCl density centrifugation (40). DNA preparations were made similarly by using 50 ml of yeast phosphate salts-grown cells of *R. sphaeroides* and *R. palustris*. DNA of *B. denitrificans*, *O. anthropi*, and *P. myrsinacearum* were prepared from 10 ml of modified arabinose gluconate broth cultures by use of a tissue and blood DNA extraction kit (Qiagen Inc., Chatsworth, Calif.). This kit was also used to prepare DNA of *M. dimorpha* grown in 10 ml of TGY broth (23) or *A. felis* grown in 10 ml of charcoal-yeast extract broth. Concentrations of DNA in solution were measured spectrophotometrically at 260 nm by using a Response Spectrophotometer (Gilford Instrument Laboratories, Oberlin, Ohio).

**PCR amplification and sequencing analysis.** Primers 16Sa and 16Sb (66) were used for amplification of the 16S rRNA gene locus. The 16S rRNA genes were amplified in 120- $\mu$ l volumes as described before (63) with the exception of the primers and the PCR buffer, which was 60 mM Tris-HCl, 15 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and 3.5 mM MgCl<sub>2</sub> at pH 9.0. The primers 450 and 1440 (66) were used to amplify the ITS region (66) and are located in conserved regions of the 3' end of the 16S rRNA gene and the 5' end of the 23S rRNA gene. The PCR products generated with this primer pair also contained the intervening region (16, 27, 34, 36, 38, 53, 54, 71) of the 23S rRNA gene and permitted sequencing analysis of the 5' ends of the 23S rRNA genes by using primers 1431, 1432, 1439, and 1440 (66). The primer pair 1432 and 23S lowerB were used to amplify most of the 23S rRNA genes. Primer 1432 is located upstream of the intervening region at bases 115 to 130 of the *Bradyrhizobium japonicum* strain USDA 110 23S rRNA gene sequence (GenBank accession no. Z35330). The reverse primer 23S lowerB is located in the 5S rRNA gene at bases 113 to 97 of the *B. japonicum* USDA 110 (GenBank accession no. Z35330). PCR conditions were as those described for the 16S rRNA gene except for an annealing temperature of 52°C in a combination of a buffer containing 60 mM Tris-HCl and 15 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> and of 2.0 mM MgCl<sub>2</sub> at pH 9.0. The PCR products were purified by using QIAquick Spin columns (Qiagen Inc.). A Perkin-Elmer (Foster City, Calif.) 377 DNA sequencer in combination with a Dye Deoxy Terminator Cycle Sequencing kit (Perkin-Elmer) was used for sequencing the purified PCR products. Primers for sequencing were designed by comparing the 23S rRNA genes of *B. japonicum* USDA 110 and *A. vitis* type strain NCPPB3554 (GenBank accession no. Z35330 and U45329, respectively) and are listed in Table 1. Database entries U69638, U35000, AF338176, and AF338177 for *B. japonicum* USDA 6, *Bradyrhizobium elkanii* USDA 76, *B. denitrificans* IFAM 1005, and *A. felis* ATCC 49715, respectively, were updated to include the 23S rRNA gene sequence.

**Analysis of the sequence data.** The sequences were aligned by using the PILEUP program in the Wisconsin package of the Genetics Computer Group (Madison, Wis.). Aligned sequences were checked manually and were edited with Genedoc (Genedoc manual, K. B. Nicholas and H. B. Nicholas, Pittsburgh Supercomputing Center, Pittsburgh, Pa.). Neighbor-joining trees were constructed from Jukes-Cantor distances by using the Molecular Evolutionary Genetics Analysis (MEGA) package version 1.01 (29), or trees were assembled in a stepwise manner with parsimony analysis by using Paup version 4.0b8a (58). Likelihood scores of tree files were generated to determine congruence between tree topologies by using the Shimodaira-Hasegawa test (55). The significance in differences among the likelihood scores was determined with a one-tailed bootstrap test using 1,000 permutations of the data. The Geneconv program version 1.02 (52) was used to test the possibility of a history of recombination among 16S rRNA loci. With this method the distribution of polymorphic nucleotide positions along aligned sequences is examined to estimate the likelihood that distinct segments have differing phylogenies.

**Nucleotide sequence accession numbers.** The sequences determined in this study have been deposited in the GenBank database under accession numbers AY244358 to AY244382.

## RESULTS

The molecular sizes of the PCR products obtained with amplification reactions using primers for the 16S rRNA gene were relatively uniform, varying only by approximately 100 bp. The 16S rRNA gene sequences obtained with *A. tumefaciens*,

TABLE 1. Primers used to sequence the 23S rRNA gene generated with primers 1432 (66) and 23S lowerB

Primer	Primer sequence (5'-3')	Location in GenBank accession no. Z35330 (nt)
1433	GTT GGC TTR GAR GCA GC	4306-4322
1434	CCT TGY CGG GTA AGT T	5065-5080
1435	AGT AVG GCG GRA CAC GTG	3615-3632
1436	AAA CCG ACA CAG GTR G	4758-4773
1437	GCT GAA RGC ATC TAA G	5869-5884
1438	GCC AAG GCA TCC RYC	3176-3162
1439	GGG TTN CCC CAT TCG G	3268-3253
1440	CAC GTG TYC CGC CBT ACT	3632-3615
1441	GCT GCY TCY AAG CCA AC	4322-4306
1442	CYA CCT GTG TCG GTT T	4773-4758
1443	AAC TTA CCC GRC AAG G	5080-5065
1445b	GTR CCT TTT GTA KAA TG	3808-3823
1446b	CAT TMT ACA AAA GGY AC	3824-3808
1447	ATA GCT GGT TCT YBC CG	4050-4066
1448	TCG VRA GAA CCA GCT AT	4066-4050
1449b	CAT GAG TAR CGA HAA	4515-4529
1450	TTD TCG YTA CTC ATG	4529-4515
1451	GAC GGA AAG ACC CCR TG	5177-5193
1452	CAY GGG GTC TTT CCG TC	5193-5177
1453	AGT TTG ACT GGG GCG GT	5366-5382
1454	ACC GCC CCA GTC AAA CT	5382-5366
1455	CCT CGA TGT CGG CTC	5623-5637
1456b	GAG CCG ACA TCG AGG	5637-5623
23S lowerB	GGC AGC GAC CTA CTC TC	6231-6215

*A. rubi*, *A. vitis*, *Agrobacterium rhizogenes*, *A. felis*, *B. denitrificans*, *M. dimorpha*, *O. anthropi*, *P. myrsinacearum*, and *R. palustris* were consistent with sequences for these taxa in the GenBank database. The molecular sizes of PCR products generated with primers 450 and 1440 to amplify the ITS region have been reported by van Berkum and Fuhrmann (66) and van Berkum and Eardly (64).

The molecular sizes of products generated with primers 1432 and 23S lowerB to amplify a large portion of the 23S rRNA gene and the ITS region between the 23S rRNA and 5S rRNA genes were approximately 3.0 kb. Entire sequences for the 23S rRNA genes were obtained from the analysis of two PCR products, the ITS region and the remainder of the 23S rRNA gene. Sequencing results obtained with *B. japonicum* USDA 110 and the 23S rRNA gene sequence of GenBank accession Z35330 were consistent. The results obtained for the 23S rRNA gene sequences of 31 taxa were used to derive Jukes-Cantor distances to construct a neighbor-joining tree by using the sequence of *R. sphaeroides* (GenBank accession no. x53855) as outgroup (Fig. 1B). Unfortunately, it was not possible to obtain a PCR product with primers 1432 and 23S lowerB by using *Rhizobium mongolense* as a DNA template. This was probably due to the poor sequence homology of primer 23S lowerB and the corresponding region in the 5S rRNA gene, since 1432 and the 23S rRNA gene were homologous (67).

The neighbor-joining trees obtained from Jukes-Cantor distances of aligned 16S or 23S rRNA gene sequences appeared to differ (Fig. 1A and B). With 23S rRNA gene sequences, *B. elkanii* and *B. japonicum* were placed together in a group, whereas with the 16S rRNA gene sequences they were separated by *B. denitrificans*, *R. palustris*, and *A. felis*. Also, the six species of *Sinorhizobium* were nested within the group formed

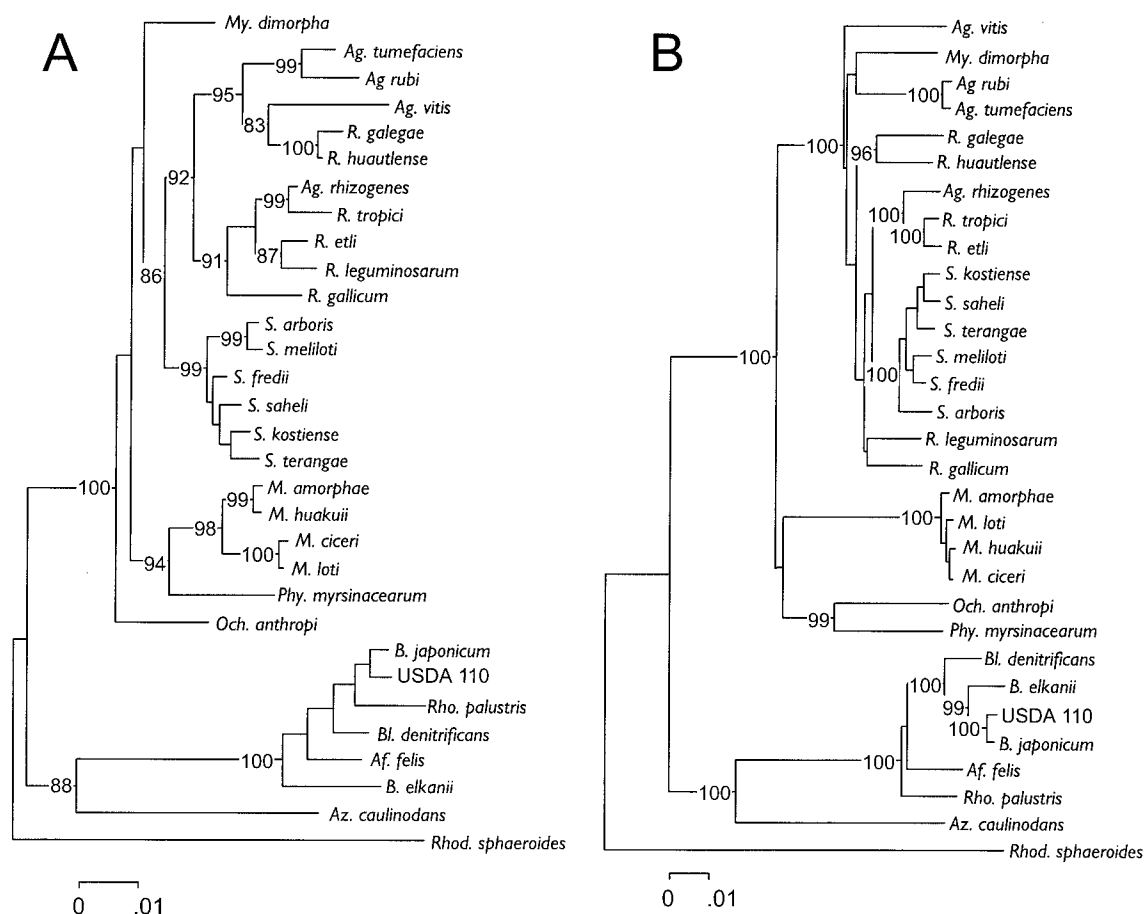


FIG. 1. Phylogenetic relationships among a select group of bacteria belonging to the  $\alpha$ -subdivision of the *Proteobacteria* reconstructed from 16S rRNA gene sequence divergence (A) or 23S rRNA gene sequence divergence (B). Sequences were aligned by using PILEUP of the Wisconsin package of the Genetics Computer Group (Madison, Wis.); aligned sequences were checked manually and were edited with Genedoc (see Materials and Methods) before deriving neighbor-joining trees that were constructed from Jukes-Cantor distances by using the MEGA package version 1.01 (29). Trees assembled in a stepwise manner with parsimony analysis by using Paup version 4.0b8a (58) had the same topologies as the distance trees shown. GenBank accession numbers used for the 16S rRNA gene sequences were as follows: *A. felis*, AF003937; *A. rhizogenes*, D14501; *A. rubi*, D14503; *A. tumefaciens*, D14500; *A. vitis*, D14502; *A. caulinodans*, X94200; *B. denitrificans*, S46917; *B. japonicum*, U69638; *B. japonicum* USDA 110, Z35330; *B. elkani*, U35000; *M. amorphae*, AF041442; *M. ciceri*, U07934; *M. huakuii*, D13431; *M. loti*, X67229; *M. dimorpha*, D12786; *O. anthropi*, D12794; *P. myrsinacearum*, D12789; *R. etli*, U28916; *R. galegae*, X67226; *R. gallicum*, U86343; *R. leguminosarum*, U29386; *R. tropici*, U89832; *Sinorhizobium arboris*, Z78204; *Sinorhizobium fredii*, X67231; *Sinorhizobium kostiense*, Z78203; *S. meliloti*, X67222; *Sinorhizobium saheli*, X68390; *Sinorhizobium terangae*, X68387; *R. sphaeroides*, X53855; and *R. palustris*, D25312. GenBank accession numbers used for the 23S rRNA gene sequences were as follows: *B. japonicum* USDA 110, Z35330; *R. palustris*, X71839; and *R. sphaeroides*, X53855.

by species of *Rhizobium* with analysis of 23S rRNA gene sequence data, whereas with 16S rRNA gene sequence data they formed a group neighboring two groups consisting of species within the genera *Rhizobium* and *Agrobacterium*. Also, there was a difference between the two data sets relative to the placement of *Rhizobium galegae*, *Rhizobium huautlense*, *Rhizobium leguminosarum*, *Rhizobium gallicum*, *M. dimorpha*, *A. vitis*, *A. rubi*, and *A. tumefaciens* (Fig. 1A and B).

Both data sets were also used to construct trees in a stepwise manner by parsimony analysis. The alignment of the 16S rRNA gene sequences contained a total of 1,509 characters, of which 1,157 were constant, 109 were variable but parsimony uninformative and 243 were parsimony informative. From a heuristic search, the score of best trees was 903 and the number of trees with this score was 27. The strict consensus tree, derived from the 27 most parsimonious trees, had a topology consistent with

the distance tree (Fig. 1A). The alignment of the 23S rRNA gene sequences contained a total of 3,237 characters, of which 2,184 were constant, 274 were variable but parsimony uninformative, and 779 were parsimony informative. From a heuristic search, the score of best trees was 3092 and the number of trees with this score was 11. The strict consensus tree, derived from the 11 most parsimonious trees, had a topology consistent with the distance tree (Fig. 1B).

The Shimodaira-Hasegawa test (55) was used to determine whether the same or different phylogenetic information was obtained from the analysis of the 16S rRNA and 23S rRNA genes. Likelihood scores were obtained for every tree, and the difference between each score and that of the tree with the smallest numerical score (best) was calculated. Subsequently, these values were compared in a one-tailed bootstrap test using 1,000 replications to determine whether scores were signifi-

TABLE 2. Shimodaira-Hasegawa test (55) of the most parsimonious trees generated from aligned sequences of the 23S and 16S rRNA genes

Difference in the likelihood scores of each tree and the best tree for the most parsimonious trees	<i>P</i> <sup>a</sup>
<b>23S rRNA trees</b>	
10.26911 (best)	0.839
8.36277	0.885
12.15221	0.827
14.42721	0.810
1.89246	0.920
13.28196	0.807
10.58340	0.852
10.84981	0.849
15.90222	0.776
15.85598	0.764
<b>16S rRNA trees</b>	
1013.71527	0.000
1184.58247	0.000
1023.51223	0.000
1023.28739	0.000
1071.46505	0.000
1071.46505	0.000
1241.66444	0.000
1241.66444	0.000
1191.59262	0.000
1191.38500	0.000
1081.26676	0.000
1081.26676	0.000
1081.04208	0.000
1081.04208	0.000
1248.67340	0.000
1248.46599	0.000
1248.67340	0.000
1248.46599	0.000
1278.20933	0.000
1278.20933	0.000
1288.32987	0.000
1288.11444	0.000
1229.25939	0.000
1229.04393	0.000
1288.32987	0.000
1288.11444	0.000
1219.13446	0.000

<sup>a</sup> Significance was determined by a one-tailed bootstrap test with 1,000 replications. A *P* value of <0.05 is significant, indicating that tree topologies were different.

cantly different at a *P* value of 0.05 (Table 2). The best likelihood score was obtained with one of the trees resulting from parsimony analysis of the 23S rRNA gene data. The mean difference in the likelihood scores of the best tree and the remaining 10 most parsimonious trees from the 23S rRNA sequence data were not significantly different (Table 2). Therefore, the topologies among the 11 most parsimonious trees obtained with the 23S rRNA gene sequences were not significantly different. The mean difference in the likelihood scores of the best tree (23S) and the 27 most parsimonious trees from the 16S rRNA sequence data was significant (Table 2). Therefore, the topologies of the parsimony trees obtained by analysis of the 16S and the 23S rRNA genes were significantly different. The likelihood score among the two distance trees (Fig. 1A and B) was the smallest with the reconstruction using the 23S

rRNA gene sequence data. The difference in the likelihood scores among the two distance trees was 948.41, and this difference in a bootstrap test was significant (*P* = 0.000). Therefore, the distance tree topologies reconstructed from the 16S rRNA and 23S rRNA data were also significantly different. Thus, the conclusion was that the 16S rRNA gene and the 23S rRNA gene sequences gave different phylogenetic information irrespective of whether a distance or parsimony method had been used to do the analysis.

A similar comparison was made with distance trees reconstructed from 16S rRNA gene and ITS region sequences (Fig. 2A and B). These two trees also appeared to have differing topologies. In particular, the placement of the serogroups belonging to *B. elkanii* (USDA 76) relative to those of *B. japonicum* (USDA 6) varied between the two reconstructions. In the case of the 16S rRNA gene data, *B. elkanii* was placed proximal to *Nitrobacter hamburgensis*, *A. felis*, *R. palustris*, and *B. denitrificans* and distal to *B. japonicum* (Fig. 2A). However, in the case of reconstructions with the ITS data, *B. japonicum* and *B. elkanii* were placed together in a group (Fig. 2B). The distance tree obtained from analysis of the ITS region had the best likelihood score. This score differed significantly (*P* = 0.000) from that obtained with the 16S rRNA gene sequence data by using the Shimodaira and Hasegawa test (55).

Parsimony analyses of these two data sets were also done. The alignment of the 16S rRNA gene sequences contained a total of 1,508 characters, of which 1,185 were constant, 152 were variable but parsimony uninformative, and 171 were parsimony informative. From a heuristic search, the score of best trees was 469 and the number of trees with this score was 2,184. The alignment of the ITS region sequences contained a total of 2,130 characters, of which 1,080 were constant, 476 were variable but parsimony uninformative, and 574 were parsimony informative. From a heuristic search, the score of the best trees was 2079 and the number of trees with this score was 26.

Likelihood scores were generated for the first 174 trees obtained with parsimony analysis of the 16S rRNA gene sequence data and all 26 trees from the ITS data. Subsequently the values of 200 trees were used to compare topologies by the Shimodaira-Hasegawa test (55). It was considered unnecessary to include all of the most parsimonious trees obtained with the aligned 16S rRNA sequence data. The best likelihood score was obtained with one of the trees resulting from parsimony analysis of the ITS data. The topologies among the 26 most parsimonious trees obtained with the ITS sequences were not significantly different (mean *P* value, 0.754). The differences in the likelihood scores of the best tree (ITS) and 174 of the most parsimonious trees obtained with the aligned 16S rRNA gene sequences were significantly different (*P* = 0.000). Therefore, the 16S rRNA gene and ITS region sequences also gave different phylogenetic information.

In an effort to explain the incongruence that we observed among our three data sets, we compared the distribution of polymorphic nucleotide positions along 16S rRNA gene alleles. The first analysis was focused on *B. elkanii* because it grouped with *B. japonicum* in reconstructions of evolutionary history of the 23S rRNA gene and the ITS region. Therefore, aligned 16S rRNA gene sequences of *B. elkanii* (USDA 76), *B. japonicum* (USDA 6), and *Mesorhizobium mediterraneum* (Upm-Ca 36) were examined by using Genconv version 1.02

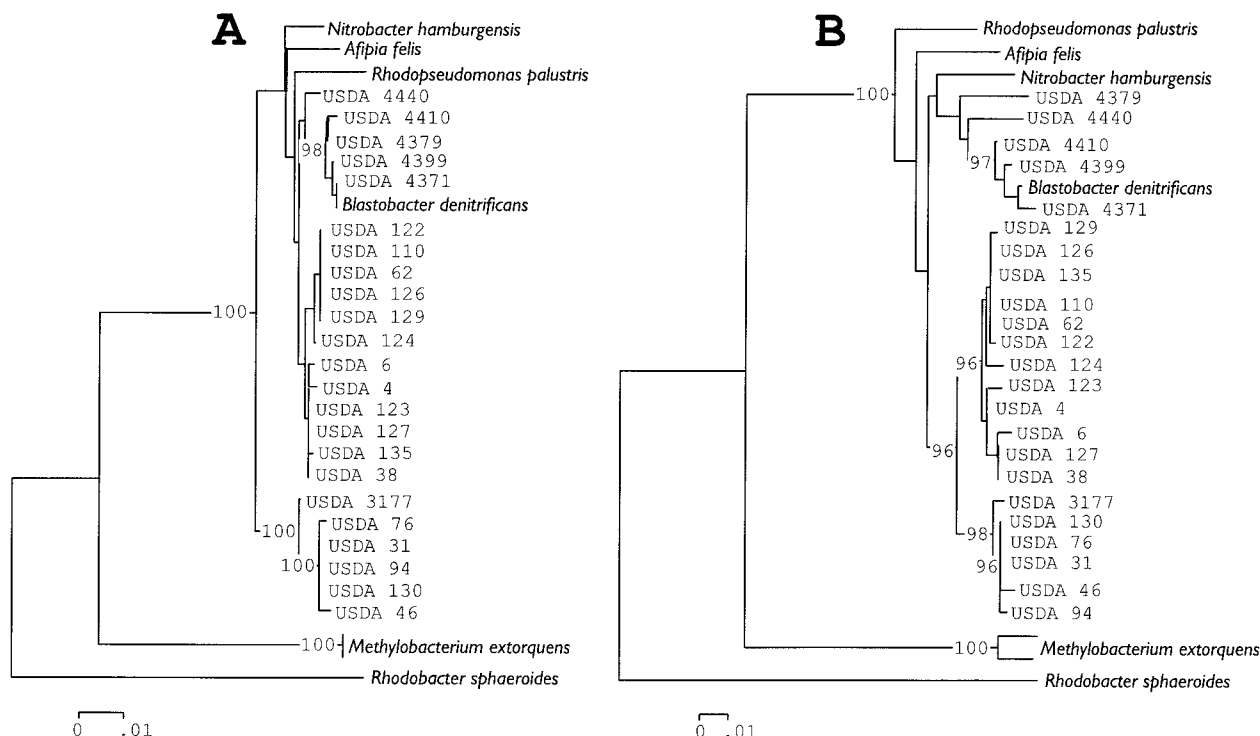


FIG. 2. Phylogenetic relationships among a select group of bacteria belonging to the  $\alpha$ -subdivision of the *Proteobacteria* reconstructed from the 16S rRNA gene sequence divergence (A) or the ITS region sequence divergence (B). Sequences were aligned by using PILEUP of the Wisconsin package of the Genetics Computer Group (Madison, Wis.); aligned sequences were checked manually and were edited with Genedoc (see Materials and Methods) before deriving neighbor-joining trees that were constructed from Jukes-Cantor distances using the MEGA package version 1.01 (29). Trees assembled in a stepwise manner with parsimony analysis by using Paup version 4.0b8a (58) had the same topologies as the distance trees shown.

(52). There were 1,485 aligned bases with 175 polymorphisms. Evidence for possible gene conversion events between the 16S rRNA gene sequences of USDA 76 and Upm-Ca 36 was identified by two globally significant ( $P < 0.05$ ) inner fragments with global permutation  $P$  values of 0.0000 and pairwise  $P$  values of 0.0000. These two inner fragments were 20 and 43 bp in length and were adjacent to each other (Fig. 3). The effect of this 81-bp region (Fig. 3) on tree topology was examined by

removing it from aligned sequences of *B. japonicum*, *B. elkanii*, *R. palustris*, *B. denitrificans*, *A. felis*, *Nitrobacter winogradskyi*, *Azorhizobium caulinodans*, *M. mediterraneum*, *S. meliloti*, and *R. leguminosarum* and comparing a neighbor-joining tree obtained from Jukes-Cantor distances with a similar tree (control) using the aligned sequences without deleting this small region. Because more sequences were used for tree construction than in the search for gene conversion, more gaps were

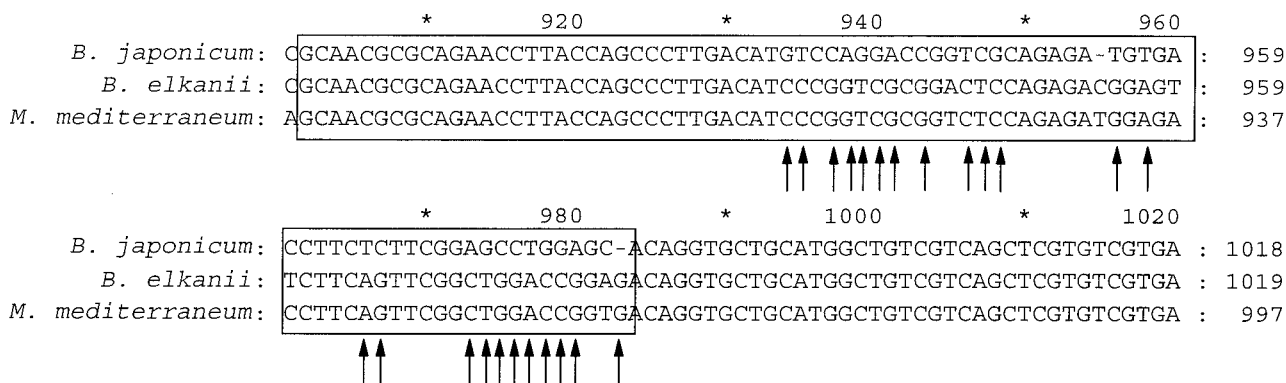


FIG. 3. Comparison of the distribution of polymorphic nucleotide positions along 16S rRNA gene genes of *B. elkanii* (USDA 76); *B. japonicum* (USDA 6), and *M. mediterraneum* (Upm-Ca 36) identifying possible gene conversion events between alleles of USDA 76 and Upm-Ca36 from evidence gathered with Genconv version 1.02 (52). The arrows indicate base pair matches between *B. elkanii* and *M. mediterraneum* for which there was a corresponding mismatch with *B. japonicum*.

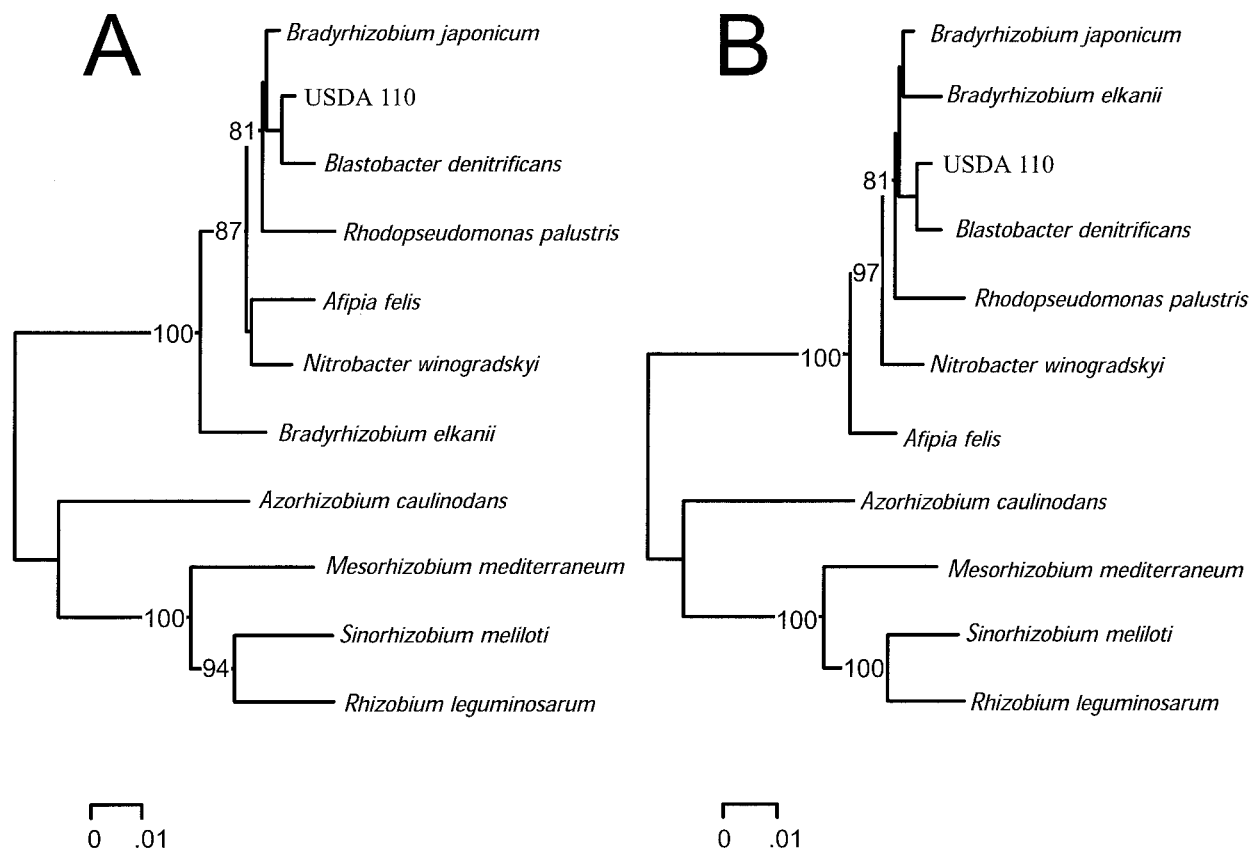


FIG. 4. Changes in 16S rRNA gene tree topology following removal of a region identified as a possible gene conversion event between alleles of *B. elkanii* and *M. mediterraneum*. Distance control trees (A) were reconstructed from aligned 16S rRNA gene sequence of *B. japonicum* (U69638), *B. elkanii* (U35000), *R. palustris* (D25312), *B. denitrificans* (S46917), *A. felis* (AF003937), *N. winogradskyi* (L35506), *A. caulinodans* (X94200), *M. mediterraneum* (L38825), *S. meliloti* (X67222), and *R. leguminosarum* (U29386) as described for the legend to Fig. 1 and were compared with a similar tree by using the aligned sequences from which the small region had been removed (B). Trees assembled in a stepwise manner with parsimony analysis using Paup version 4.0b8a (58) had the same topologies as the distance trees shown.

created in the alignment resulting in the deletion of 84 bp. The deletion resulted in the shift in placement of *B. elkanii* from adjacent to the group formed by *B. japonicum*, *R. palustris*, *B. denitrificans*, *A. felis*, and *N. winogradskyi* to a position neighboring *B. japonicum* (Fig. 4). Tree topologies derived from both alignments were significantly different ( $P = 0.008$ ) by the Shimodaira-Hasegawa test (55). Trees were also constructed in a stepwise manner by using parsimony analysis. The control alignment had 1,496 characters, of which 1,250 were constant, 65 were variable but parsimony uninformative, and 181 were parsimony informative, and these 181 in a branch and bound search resulted in two most parsimonious trees with a score of 402. The alignment with the deletion had 1,412 characters, of which 1,204 were constant, 55 were variable but parsimony uninformative, and 153 were parsimony informative, and these 153 in a branch and bound search resulted in one most parsimonious tree with a score of 314 and with a topology identical with the distance tree (Fig. 4B). The tree with the best likelihood score was obtained with the alignment containing the deletion and this score and those of the control trees by the Shimodaira-Hasegawa test differed significantly at the 5% level (average  $P$  value, 0.044). Therefore, removal of the two inner fragments identified as a possible gene conversion event be-

tween the 16S rRNA gene sequences of USDA 76 and Upm-Ca 36 resulted in a significant change in tree topology.

The second analysis for the possibility of transfer and genetic recombination among divergent 16S rRNA alleles was focused on *Sinorhizobium*, because species of this genus were nested within the genus *Rhizobium* in reconstructions of evolutionary history of the 23S rRNA genes (Fig. 1B). Therefore, aligned 16S rRNA gene sequences of *Sinorhizobium fredii*, *S. meliloti*, *Mesorhizobium ciceri*, *Mesorhizobium loti*, *Mesorhizobium amorphae*, *M. mediterraneum*, and *Mesorhizobium huakuii* were examined by using Genconv version 1.02 (52). There were 1,475 aligned bases with 88 polymorphisms. Evidence for possible gene conversion events between the 16S rRNA gene sequences of *S. fredii* and *M. amorphae* or *M. huakuii* were identified by a globally significant inner fragment of 168 bp with global permutation  $P$  values of 0.0036 or 0.0104 and pairwise  $P$  values of 0.0013 or 0.0004 (Fig. 5, box B). A second region of 56 bp was identified in a pairwise comparison by pairwise  $P$  values of 0.0107 or 0.0078 (Fig. 5, box A). An additional 84-bp fragment, which was located within the 168-bp fragment detected with *S. fredii*, was identified in a comparison of the 16S rRNA genes of *S. meliloti* and *M. amor-*

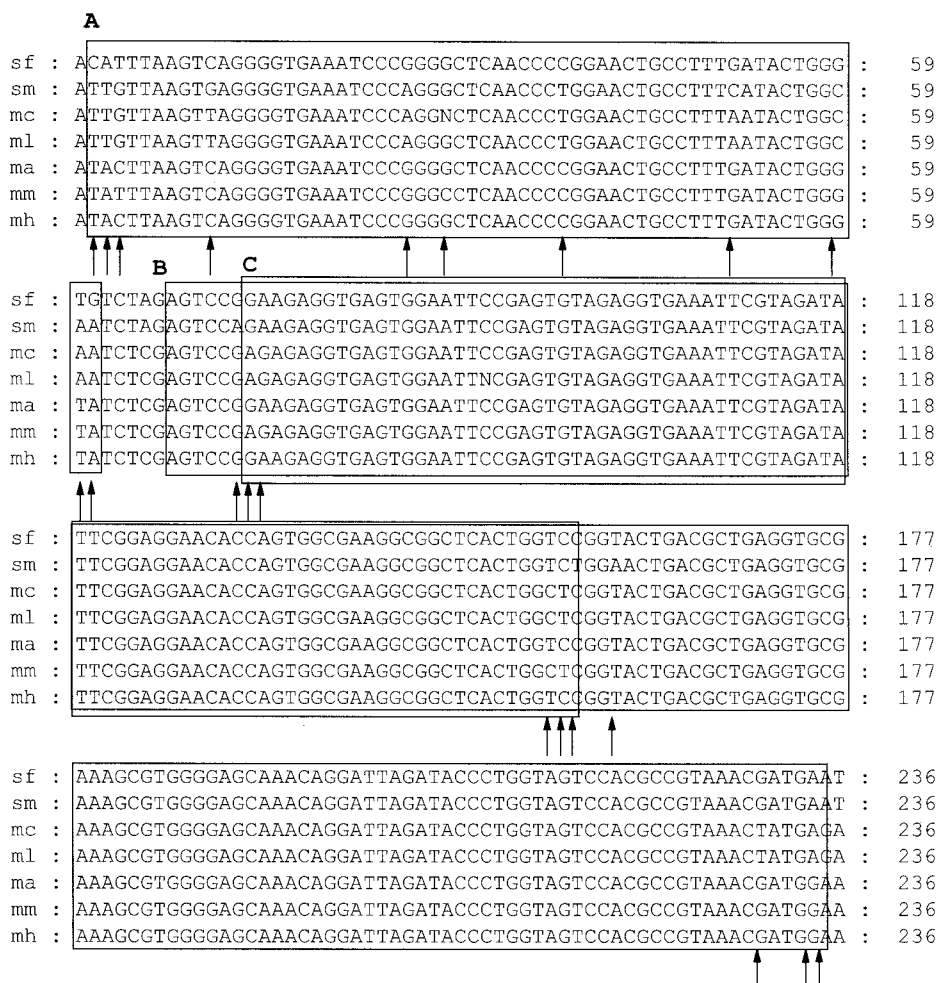


FIG. 5. Comparison of the distribution of polymorphic nucleotide positions along 16S rRNA genes of *S. fredii* (USDA 205), *S. meliloti* (USDA 1002), *M. ciceri* (UPM-Ca7), *M. loti* (NZP 2213), *M. amorphae* (ACCC 19665), *M. mediterraneum* (Upm-Ca36), and *M. huakuii* (CCBAU2609) identifying possible gene conversion events between alleles from evidence gathered with Genconv version 1.02 (52). The arrows indicate base pair mismatches between *S. fredii* and *S. meliloti* for which there was a corresponding match between *Mesorhizobium* and *Sinorhizobium*.

*phae* or *M. huakuii* with inner pairwise *P* values of 0.0335 or 0.0177 (Fig. 5, box C).

The effect of the combined 232-bp (Fig. 5) and 81-bp (Fig. 3) regions (a combined deletion of 317 bp in the alignment using 31 sequences) on tree topology was examined by removing them from the 16S rRNA gene sequence alignment that was originally used for comparison with an analysis of the 23S rRNA gene (Fig. 1A). Neighbor-joining trees obtained from Jukes-Cantor distances were compared with a similar tree using the aligned sequences without (control) deleting this small region. Tree topologies appeared to differ in the placement of *B. elkanii* and species within the genus *Sinorhizobium* (Fig. 6A). The placement of *B. elkanii* changed from adjacent to the group formed by *B. japonicum*, *R. palustris*, *B. denitrificans*, and *A. felis* (Fig. 1A) to a position within *B. japonicum* (Fig. 6A). The placement of six species of *Sinorhizobium* changed from adjacent to the group formed by species of *Agrobacterium* and *Rhizobium* (Fig. 1A) to a position that formed a group with *R. leguminosarum*, *Rhizobium etli*, *R. gallicum*, *Rhizobium tropici*, and *A. rhizogenes* (Fig. 6A). Trees were also constructed in a

stepwise manner by using parsimony analysis. The alignment with the two deletions had 1,192 characters, of which 941 were constant, 88 were variable but parsimony uninformative, and 163 were parsimony informative, and these 163 in a heuristic search resulted in 84 most parsimonious trees with a score of 539. The tree with the best likelihood score was obtained with the control alignment. The difference in this score and those generated with the trees obtained from alignments containing the deletions in a bootstrap analysis were significantly different at the 5% level (*P* < 0.01) by the Shimodaira-Hasegawa test (55). A similar analysis using the two distance trees also resulted in a significant difference in the scores (*P* < 0.000). Therefore, removal of two portions of the alignment identified as three possible gene conversion events between 16S rRNA gene sequences of species of *M. huakuii*, *M. amorphae*, *S. fredii*, and *S. meliloti* (Fig. 5) as well as *M. mediterraneum* and *B. elkanii* (Fig. 3) resulted in a significant change in tree topology.

Jukes-Cantor distances were also used to construct a neighbor-joining tree with a combination of the two regions that were deleted from the original alignment of the 16S rRNA

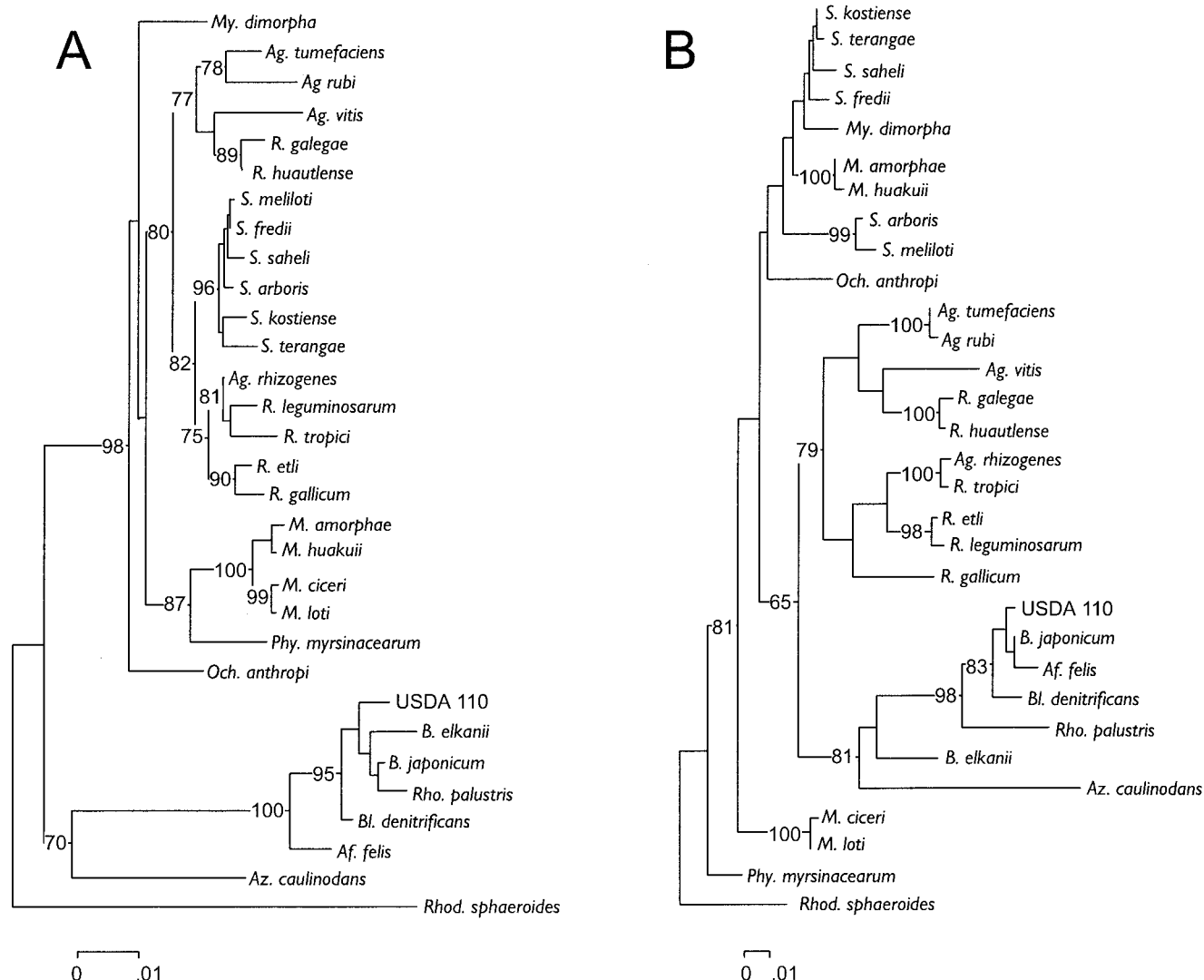


FIG. 6. Changes in 16S rRNA gene tree topology following removal of two regions identified as a possible gene conversion event between alleles of two species of *Sinorhizobium* and four species of *Mesorhizobium*. The distance control tree is presented in Fig. 1A, which is compared with the distance tree after removal of the 317-bp fragment (A). Also shown is the tree topology obtained using the combined region that was deleted from the alignment (B). Trees assembled in a stepwise manner with parsimony analysis by using Paup version 4.0b8a (58) had the same topologies as the distance trees shown.

genes (Fig. 3 and 5). The alignment was 317 bp in length and resulted in a tree topology that visually was very different from all trees obtained with the 16S rRNA gene sequence data (Fig. 6B). Most notable was the change in the placement of *Mesorhizobium* spp., *Sinorhizobium* spp., *M. dimorpha* and *O. anthropi*, which formed a group. This group occupied a position with *P. myrsinacearum* that was closer to the outgroup (*R. sphaeroides*) than the group formed by *B. japonicum*, *B. elkanii*, *A. felis*, *B. denitrificans*, *R. palustris*, and *A. caulinodans* (Fig. 6B) in contrast to the control tree (Fig. 1A). Finally, the topologies of the distance tree obtained with the aligned 16S rRNA gene sequences from which the two regions were deleted and that obtained by analysis of the aligned 23S rRNA gene also differed significantly ( $P = 0.000$ ) when analyzed by the Shimodaira-Hasegawa test (55). Therefore, the two regions (317 bp) removed from the aligned 16S rRNA gene sequences

alone do not account for the significant difference in phylogenetic information obtained from analysis of these two loci.

## DISCUSSION

The focus of this study was to confirm the phylogenetic placement of the six genera *Allorhizobium*, *Azorhizobium*, *B Bradyrhizobium*, *Rhizobium*, *Mesorhizobium*, and *Sinorhizobium* within the  $\alpha$ -Proteobacteria in reconstructions from 16S rRNA gene sequence divergence by the analysis of the ITS region and the 23S rRNA gene. This examination was initiated because it was concluded from partial 23S rRNA gene sequence analysis that differences between *Sinorhizobium* and *Rhizobium* might be insufficient to warrant their separation into different genera (67) and because placement of *M. loti* was uncertain in comparative analyses of partial 23S rRNA and 16S rRNA sequence



data (67). The analysis of the entire 23S rRNA gene was relevant because evidence from partial gene sequence data may be inconclusive, since different portions of the 16S rRNA gene sequence were reported to provide conflicting phylogenetic signals (13). The Shimodaira-Hasegawa test (55) was used to assess which pairs of phylogenetic trees have similar topologies. This test was used in every case, because the alternative test by Kishino and Hasegawa (25) would not be valid for cases in which derivatives of the same alignments were compared (20). Evidence was obtained that the ITS region and the 23S rRNA gene provided phylogenetic signals that differed from those of the 16S rRNA gene.

Gaunt et al. (19) also compared the phylogeny of the 16S rRNA gene with two other loci. They concluded from an analysis of *atpD* and *recA* that there was broad phylogenetic agreement among the loci examined in their study. They based this conclusion on the consistency with which species within each genus formed a group in reconstructions with each of the three loci. It is important to note that despite the conclusion of broad phylogenetic agreement among loci, Gaunt et al. (19) indicated that from results of a partition homogeneity test the 16S rRNA gene sequence data could not be combined with those of *atpD* and *recA*. Also, possibly of significance in the investigation of Gaunt et al. (19) is the omission of *B. elkanii*, because species within the genus *Bradyrhizobium* do not form a single group in reconstructions of phylogeny from 16S rRNA gene sequence divergence. In the case of the ITS region and the 23S rRNA gene, we have provided evidence that both these species are placed in a monophyletic group and that the discordance in phylogeny with the 16S rRNA gene locus is in part because of the relative placements of *B. elkanii* and *B. japonicum*.

We suggest that gene conversion in the 16S rRNA gene may be one of the possible mechanisms responsible for the discordant phylogenies among loci. This suggestion is based on a search for regions within specific alleles of the 16S rRNA gene that may have a history of recombination. Most noteworthy were the identification of potential recombination events between short segments of the 16S rRNA genes of *B. elkanii* with species of *Mesorhizobium* and between species of *Sinorhizobium* and species of *Mesorhizobium*. The detection of potential genetic transfer and recombination among divergent alleles of the 16S rRNA gene across genera of the  $\alpha$ -*Proteobacteria* would contradict inferences of the genetic isolation of genera (19).

It is becoming increasingly evident that a major component driving microbial evolution is the exchange of genetic information by lateral transfer and recombination (1, 3–5, 11, 14, 15, 17, 18, 21, 22, 24, 26, 30–33, 35, 37, 41–44, 46, 47, 50, 51, 56, 60, 61). This exchange of genetic information is not confined to closely related genomes but has been detected at the intergenic level (28, 41) and is suggested to occur across major domains (7, 42, 49). Lateral transfer and recombination events across divergent alleles of the small subunit ribosomal genes have also been suggested (8, 13, 57), and substantial evidence has been provided through the analysis of the actinomycete *Thermomonospora chromogena* (72). Against this background we anticipated that gene conversion among the 16S rRNA genes of members of the  $\alpha$ -*Proteobacteria* was a distinct possibility. Of course the detection of gene conversion across highly conserved DNA sequences that are characteristic of ribosomal

genes is difficult, but sequence conservation also provides more opportunity for homologous recombination. We wish to emphasize that even though we presented evidence for gene conversion in the 16S rRNA gene of the  $\alpha$ -*Proteobacteria*, other genes and DNA regions also are likely to be mosaic including the 23S rRNA gene and the ITS region.

In most cases the phylogenetic placement of bacteria derived from 16S rRNA gene sequence analysis is the most dominant character used in bacterial taxonomy. This has been especially true in the case of nomenclature of bacteria that form hyper-trophies on higher plants. Some examples in which rhizobial classification decisions were made largely on the basis of analysis of the 16S rRNA gene include separation of *Sinorhizobium* from *Rhizobium* (10), the proposal of the new genus *Allorhizobium* (9), the suggestion for combining *Agrobacterium* and *Allorhizobium* into the genus *Rhizobium* (73), proposing separate genera for *B. japonicum* and *B. elkanii* (70), the identification of a species of *Methylobacterium* forming a symbiosis with specific species of *Crotalaria* (59), and nodulation of the South African legume *Aspalathus carnosa* by a bacterium placed within the  $\beta$  subdivision of the *Proteobacteria* (39). Such suggestions or proposed changes in nomenclature may not be warranted, since the evidence for phylogenetic placement of these genera is at best inconclusive. As an alternative we suggest that a more conservative approach be taken whereby taxonomic decisions are based on the analysis of a variety of loci and that comparative analytical methods be used to estimate phylogenetic relationships among the species being considered.

#### ACKNOWLEDGMENTS

We thank Patrick Elia and Kenneth Lee Nash for excellent technical assistance.

#### REFERENCES

- Amabile-Cuevas, C. F., M. Cardenas-Garcia, and M. Ludgar. 1995. Antibiotic resistance. *Am. Sci.* **83**:320–329.
- Amann, G., K. O. Stetter, E. Llobet-Brossa, R. Amann, and J. Anton. 2000. Direct proof for the presence and expression of two 5% different 16S rRNA genes in individual cells of *Haloarcula marismortui*. *Extremophiles* **4**:373–376.
- Blattner, F. R., G. I. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, V. J. Collado, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* (Washington, D.C.) **277**:1453–1462.
- Bolotin, A., P. Wincker, S. Mauger, O. Jaillon, K. Malarne, J. Weissenbach, S. D. Ehrlich, and A. Sorokin. 2001. The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. *Genome Res.* **11**:731–753.
- Bourgoin, F., G. Guedon, M. Pebay, Y. Roussel, C. Panis, and B. Decaris. 1996. Characterization of a mosaic ISS1 element and evidence for the recent horizontal transfer of two different types of ISS1 between *Streptococcus thermophilus* and *Lactococcus lactis*. *Gene* **178**:15–23.
- Carbon, C., J. Philips, Z. Y. Fu, C. Squires, and C. L. Squires. 1979. The complete nucleotide sequence of ribosomal 16S rRNA from *Escherichia coli*. *EMBO J.* **11**:4175–4185.
- Chou, I. T., and C. S. Gasser. 1997. Characterization of the cyclophilin gene family of *Arabidopsis thaliana* and phylogenetic analysis of known cyclophilin proteins. *Plant Mol. Biol.* **35**:873–892.
- Cilia, V., B. Lafay, and R. Christen. 1996. Sequence heterogeneities among 16S ribosomal RNA sequences, and their effect on phylogenetic analyses at the species level. *Mol. Biol. Evol.* **13**:451–461.
- de Lajudie, P., E. Laurent-Fulele, A. Willems, U. Torck, R. Coopman, M. D. Collins, K. Kersters, B. Dreyfus, and M. Gillis. 1998. *Allorhizobium undicola* gen. nov., nitrogen fixing bacteria that efficiently nodulate *Neptunia natans*. *Int. J. Syst. Bacteriol.* **48**:1277–1290.
- de Lajudie, P., A. Willems, B. Pot, D. Dewettinck, G. Maestrojuan, M. Neyra, M. D. Collins, B. Dreyfus, K. Kersters, and M. Gillis. 1994. Polyphasic

- taxonomy of rhizobia: emendation of the genus *Sinorhizobium* and description of *Sinorhizobium meliloti* comb. nov. *Sinorhizobium saheli* sp. nov. and *Sinorhizobium teranga* sp. nov. *Int. J. Syst. Bacteriol.* **44**:715–733.
11. Delorme, C., J. J. Godon, S. D. Ehrlich, and P. Renault. 1994. Mosaic structure of large regions of the *Lactococcus lactis* subsp. *cremoris* chromosome. *Microbiology* **140**:3053–3060.
  12. Dreyden, S. C., and S. Kaplan. 1990. Localization and structural analysis of the ribosomal RNA operons of *Rhodobacter sphaeroides*. *Nucleic Acids Res.* **18**:7267–7277.
  13. Eardly, B. D., F.-S. Wang, and P. van Berkum. 1996. Corresponding 16S rRNA segments in *Rhizobiaceae* and *Aeromonas* yield discordant phylogenies. *Plant Soil* **186**:69–74.
  14. Edwards, R. A., G. J. Olsen, and S. R. Maloy. 2002. Comparative genomics of closely related salmonellae. *Trends Microbiol.* **10**:94–99.
  15. Eisen, J. A. 2000. Horizontal gene transfer among microbial genomes: new insights from complete genome analysis. *Curr. Opin. Genet. Dev.* **10**:606–611.
  16. Evguenieva-Hackenberg, E., and S. Selenska-Pobell. 1995. Variability of the 5'-end of the large subunit rDNA and the presence of a new short class of rRNA in *Rhizobiaceae*. *Lett. Appl. Microbiol.* **21**:402–405.
  17. Ferretti, J. J., W. M. McShan, D. Ajdic, D. J. Savic, G. Savic, K. Lyon, C. Primeaux, S. Sezate, A. N. Suvorov, S. Kenton, H. S. Lai, S. P. Lin, Y. Qian, H. G. Jia, F. Z. Najar, Q. Ren, H. Zhu, L. Song, J. White, X. Yuan, S. W. Clifton, B. A. Roe, and R. McLaughlin. 2001. Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. *Proc. Natl. Acad. Sci. USA* **98**:4658–4663.
  18. Garcia, V. S., A. Romeu, and J. Palau. 2000. Horizontal gene transfer in bacterial and archaeal complete genomes. *Genome Res.* **10**:1719–1725.
  19. Gaunt, M. W., S. L. Turner, L. Rigottier-Gois, S. A. Lloyd-Macgilp, and J. P. W. Young. 2001. Phylogenies of *atpD* and *recA* support the small subunit rRNA-based classification of rhizobia. *Int. J. Syst. Bacteriol.* **51**:2037–2048.
  20. Goldman, N., J. P. Anderson, and G. A. Rodrigo. 2000. Likelihood-based tests of topologies in phylogenetics. *Syst. Biol.* **49**:652–670.
  21. Groisman, E. A., M. H. Saier, Jr., and H. Ochman. 1992. Horizontal transfer of a phosphatase gene as evidence for mosaic structure of the *Salmonella* genome. *EMBO J.* **11**:1309–1316.
  22. Groisman, E. A., M. A. Sturmoski, F. R. Solomon, R. Lin, and H. Ochman. 1993. Molecular, functional, and evolutionary analysis of sequences specific to *Salmonella*. *Proc. Natl. Acad. Sci. USA* **90**:1033–1037.
  23. Haynes, W. C., L. J. Wickerham, and C. W. Hesseltine. 1955. Maintenance of cultures of industrially important microorganisms. *Appl. Microbiol.* **3**:361–368.
  24. Herrick, J. B., K. K. G. Stuart, W. C. Ghiorse, and E. L. Madsen. 1997. Natural horizontal transfer of a naphthalene dioxygenase gene between bacteria native to a coal tar contaminated field site. *Appl. Environ. Microbiol.* **63**:2330–2337.
  25. Kishino, H., and M. Hasegawa. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies for DNA sequence data, and the branching order in hominoidea. *J. Mol. Evol.* **29**:170–179.
  26. Koonin, E. V., K. S. Makarova, and L. Aravind. 2001. Horizontal gene transfer in prokaryotes: quantification and classification. *Annu. Rev. Microbiol.* **55**:709–742.
  27. Kordes, E., S. Jock, J. Fritsch, F. Bosch, and G. Klug. 1994. Cloning of a gene involved in rRNA precursor processing and 23S rRNA cleavage in *Rhodobacter capsulatus*. *J. Bacteriol.* **176**:1121–1127.
  28. Kroll, J. S., K. E. Wilks, J. L. Farrant, and P. R. Langford. 1998. Natural genetic exchange between *Haemophilus* and *Neisseria*: intergenic transfer of chromosomal genes between major human pathogens. *Proc. Natl. Acad. Sci. USA* **95**:12381–12385.
  29. Kumar, S., K. Tamura, and M. Nei. 1993. MEGA: molecular evolutionary genetics analysis, 1.01 ed. The Pennsylvania State University, University Park, Pa.
  30. Kunst, F., N. Ogasawara, I. Moszer, A. M. Albertini, G. Alloni, V. Azevedo, M. G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S. C. Brignell, S. Bron, S. Brouillet, C. V. Bruschi, B. Caldwell, V. Capuano, N. M. Carter, S. K. Choi, J. J. Codani, I. F. Conner-ton, N. J. Cummings, R. A. Daniel, F. Denizot, K. M. Devine, A. Duesterhoeft, S. D. Ehrlich, P. T. Emmerson, K. D. Entian, J. Errington, C. Fabret, E. Ferrari, D. Foulger, C. Fritz, M. Fujita, Y. Fujita, S. Fuma, A. Galizzi, N. Galleron, S. Y. Ghim, P. Glaser, A. Goffeau, E. J. Golightly, G. Grandi, G. Giuseppe, B. J. Guy, K. Haga, J. Haiech, C. R. Harwood, A. Henaut, H. Hilbert, S. Holsappel, S. Hosono, L. Jones, B. Joris, D. Karamata, Y. Kasahara, B. M. Klaerr, C. Klein, Y. Kobayashi, P. Koetter, G. Konigstein, S. Krogh, M. Kumano, K. Kurita, A. Lapidus, S. Lardinois, J. Lauber, V. Lazarevic, S. M. Lee, A. Levine, H. Liu, S. Masuda, C. Mauel, C. Medigue, N. Medina, R. P. Mellado, M. Mizuno, D. Moestl, S. Nakai, M. Noback, D. Noone, K. Ogawa, A. Ogiwara, B. Oudega, S. H. Park, V. Parro, T. M. Pohl, D. Portetelle, S. Porwollik, A. M. Presecan, P. Pujic, B. Purnelle, G. Rapoport, M. Rey, et al. 1997. The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. *Nature (London)* **390**:249–256.
  31. Kuroda, M., T. Ohta, I. Uchiyama, T. Baba, H. Yuzawa, I. Kobayashi, L. Cui, A. Oguchi, K. Aoki, Y. Nagai, J. Lian, T. Ito, M. Kanamori, H. Matsumaru, A. Maruyama, H. Murakami, A. Hosoyama, U. Y. Mizutani, N. K. Takahashi, T. Sawano, R. Inoue, C. Kaito, K. Sekimizu, H. Hira-kawa, S. Kuhara, S. Goto, J. Yabuzaki, M. Kanehisa, A. Yamashita, K. Oshima, K. Furuya, C. Yoshino, T. Shiba, M. Hattori, N. Ogasawara, H. Hayashi, and K. Hiramatsu. 2001. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. *Lancet* **357**:1225–1240.
  32. Lawrence, J. G., and H. Ochman. 1997. Amelioration of bacterial genome: rates of change and exchange. *J. Mol. Biol.* **44**:383–397.
  33. Lawrence, J. G., and H. Ochman. 1998. Molecular archaeology of the *Escherichia coli* genome. *Proc. Natl. Acad. Sci. USA* **95**:9413–9417.
  34. Lessie, T. G. 1965. The atypical ribosomal RNA complement of *Rhodospseudomonas sphaeroides*. *J. Gen. Microbiol.* **39**:311–320.
  35. Logsdon, J. M., Jr., and D. M. Faguy. 1999. Evolutionary genomics: Thermotoga heats up lateral gene transfer. *Curr. Biol.* **9**:R747–R751.
  36. Mackay, M. L., B. Zablén, C. R. Woese, and W. F. Doolittle. 1979. Homologies in processing and sequence between the 23S ribosomal ribonucleic acids of *Paracoccus denitrificans* and *Rhodospseudomonas sphaeroides*. *Arch. Microbiol.* **123**:165–172.
  37. Markham, P. F., M. F. Duffy, M. D. Glew, and G. F. Browning. 1999. A gene family in *Mycoplasma imitans* closely related to the PMGA family of *Mycoplasma gallisepticum*. *Microbiology* **145**:2095–2103.
  38. Marrs, B., and S. Kaplan. 1970. 23S precursor ribosomal RNA of *Rhodospseudomonas sphaeroides*. *J. Mol. Biol.* **49**:297–317.
  39. Moulin, L., A. Munive, B. Dreyfus, and C. Boivin-Masson. 2001. Nodulation of legumes by members of the b-subclass of Proteobacteria. *Nature (London)* **411**:948–950.
  40. Navarro, R. B., A. A. T. Vargas, E. C. Schröder, and P. van Berkum. 1993. Uptake hydrogenase (Hup) in common bean (*Phaseolus vulgaris*) symbioses. *Appl. Environ. Microbiol.* **59**:4161–4165.
  41. Nelson, K., and R. K. Selander. 1994. Intergeneric transfer and recombination of the 6-phosphogluconate dehydrogenase gene (*gnd*) in enteric bacteria. *Proc. Natl. Acad. Sci. USA* **91**:10227–10231.
  42. Nelson, K. E., R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, J. D. Peterson, W. C. Nelson, K. A. Ketchum, L. McDonald, T. R. Utterback, J. A. Malek, K. D. Linher, M. M. Garrett, A. M. Stewart, M. D. Cotton, M. S. Pratt, C. A. Phillips, D. Richardson, J. Heidelberg, G. G. Sutton, R. D. Fleischmann, J. A. Eisen, O. White, S. L. Salzberg, H. O. Smith, J. C. Venter, and C. M. Fraser. 1999. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature (London)* **399**:323–329.
  43. Nwosu, V. C. 2001. Antibiotic resistance with particular reference to soil microorganisms. *Res. Microbiol.* **152**:421–430.
  44. Ochman, H., J. G. Lawrence, and E. A. Groisman. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature (London)* **405**:299–304.
  45. Olsen, G. J., C. R. Woese, and R. Overbeek. 1994. The winds of (evolutionary) change: breathing new life into microbiology. *J. Bacteriol.* **176**:1–6.
  46. Perna, N. T., G. Plunkett III, V. Burland, B. Mau, J. D. Glasner, D. J. Rose, G. F. Mayhew, P. S. Evans, J. Gregor, H. A. Kirkpatrick, G. Posfai, J. Hackett, S. Plink, A. Boutin, Y. Shao, L. Miller, E. J. Grotbeck, N. W. Davis, A. Lim, E. T. Dimalanta, K. D. Potamousis, J. Apodaca, T. S. Anantharaman, J. Lin, G. Yen, D. C. Schwartz, R. A. Welch, and F. R. Blattner. 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature (London)* **409**:529–533.
  47. Ragan, M. A. 2001. Detection of lateral gene transfer among microbial genomes. *Curr. Opin. Genet. Dev.* **11**:620–626.
  48. Rainey, F. A., N. L. Ward-Rainey, P. H. Janssen, H. Hippe, and E. Stackebrandt. 1996. *Clostridium paradoxum* DSM 7308T contains multiple 16S rRNA genes with heterogeneous intervening sequences. *Microbiology* **142**:2087–2095.
  49. Rosenthal, B., Z. Mai, D. Caplviski, S. Ghosh, H. de la Vega, T. Graf, and J. Samuelson. 1997. Evidence for the bacterial origin of genes encoding fermentation enzymes of the amitochondriate protozoan parasite *Entamoeba histolytica*. *J. Bacteriol.* **179**:3736–3745.
  50. Ruepp, A., W. Graml, M. M. L. Santos, K. K. Koretke, C. Volker, H. W. Mewes, D. Frishman, S. Stocker, A. N. Lupas, and W. Baumeister. 2000. The genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum*. *Nature (London)* **407**:508–513.
  51. Salanoubat, M., S. Genin, F. Artiguenave, J. Gouzy, S. Mangenot, M. Arlat, A. Billault, P. Brottier, J. C. Camus, L. Cattolico, M. Chandler, N. Choise, R. C. Claudel, S. Cunnac, N. Demange, C. Gaspin, M. Lavie, A. Moisan, C. Robert, W. Saurin, T. Schiex, P. Signier, P. Thebault, M. Whalen, P. Wincker, M. Levy, J. Weissenbach, and C. A. Boucher. 2002. Genome sequence of the plant pathogen *Ralstonia solanacearum*. *Nature (London)* **415**:497–502.
  52. Sawyer, S. A. 1999. GENECONV: a computer package for the statistical detection of gene conversion., 1.02 ed. Department of Mathematics, Washington University, St. Louis, Mo.
  53. Schuch, W., and U. E. Loening. 1975. The ribosomal ribonucleic acid of *Agrobacterium tumefaciens*. *Biochem. J.* **149**:17–22.
  54. Selenska-Pobell, S., and E. Evguenieva-Hackenberg. 1995. Fragmentations

- of the large-subunit rRNA in the family *Rhizobiaceae*. *J. Bacteriol.* **177**:6993–6998.
55. Shimodaira, H., and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* **16**:1114–1116.
  56. Simpson, A. J. G., F. C. Reinach, P. Arruda, F. A. Abreu, M. Acencio, R. Alvarenga, L. M. C. Alves, J. E. Araya, G. S. Baia, C. S. Baptista, M. H. Barros, E. D. Bonaccorsi, S. Bordin, J. M. Bove, M. R. S. Briones, M. R. P. Bueno, A. A. Camargo, L. E. A. Camargo, D. M. Carraro, H. Carrer, N. B. Colauto, C. Colombo, F. F. Costa, M. C. R. Costa, N. C. M. Costa, L. L. Coutinho, M. Cristofani, N. E. Dias, C. Docena, D. H. El, A. P. Facincani, A. J. S. Ferreira, V. C. A. Ferreira, J. A. Ferro, J. S. Fraga, S. C. Franca, M. C. Franco, M. Frohme, L. R. Furlan, M. Garnier, G. H. Goldman, M. H. S. Goldman, S. L. Gomes, A. Gruber, P. L. Ho, J. D. Hoheisel, M. L. Junqueira, E. L. Kemper, J. P. Kitajima, J. E. Krieger, E. E. Kuramae, F. Laigret, M. R. Lambais, L. C. C. Leite, E. G. M. Lemos, M. V. F. Lemos, S. A. Lopes, C. R. Lopes, J. A. Machado, M. A. Machado, A. M. B. N. Madeira, H. M. F. Madeira, C. L. Marino, M. V. Marques, E. A. L. Martins, E. M. F. Martins, A. Y. Matsukuma, C. F. M. Menck, E. C. Miracca, C. Y. Miyaki, V. C. B. Monteiro, D. H. Moon, M. A. Nagai, A. L. T. O. Nascimento, L. E. S. Netto, A. Nhani, Jr., F. G. Nobrega, L. R. Nunes, M. A. de Oliveira, M. C. de Oliveira, O. R. C. de Oliveira, D. A. Palmieri, A. Paris, B. R. Peixoto, G. A. G. Pereira, H. A. Pereira, Jr., J. B. Pesquero, R. B. Quaggio, P. G. Roberto, V. Rodrigues, et al. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. *Nature (London)* **406**:151–157.
  57. Sneath, P. H. A. 1993. Evidence from *Aeromonas* for genetic crossing-over in ribosomal sequences. *Int. J. Syst. Bacteriol.* **43**:626–629.
  58. Swofford, D. L. 1999. PAUP\*. Phylogenetic analysis using parsimony (\*and other methods), 4.0b8a ed. Sinauer Associates, Sunderland, Mass.
  59. Sy, A., E. Giraud, P. Jourand, N. Garcia, A. Willems, P. de Lajudie, Y. Prin, M. Neyra, M. Gillis, C. Boivin-Masson, and B. Dreyfus. 2001. Methylo-trophic *Methylobacterium* bacteria nodulate and fix nitrogen in symbiosis with legumes. *J. Bacteriol.* **183**:214–220.
  60. Tettelin, H., N. J. Saunders, J. Heidelberg, A. C. Jeffries, K. E. Nelson, J. A. Eisen, K. A. Ketchum, D. W. Hood, J. F. Peden, R. J. Dodson, W. C. Nelson, M. L. Gwinn, R. DeBoy, J. D. Peterson, E. K. Hickey, D. H. Haft, S. L. Salzberg, O. White, R. D. Fleischmann, B. A. Dougherty, T. Mason, A. Ciecko, D. S. Parksey, E. Blair, H. Citti, E. B. Clark, M. D. Cotton, T. R. Utterback, H. Khouri, H. Qin, J. Vamathevan, J. Gill, V. Scarlato, V. Masignani, M. Pizza, G. Grandi, L. Sun, H. O. Smith, C. M. Fraser, E. R. Moxon, R. Rappuoli, and J. C. Venter. 2000. Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58. *Science (Washington, D.C.)* **287**:1809–1815.
  61. Teuber, M. 1999. Spread of antibiotic resistance with foodborne pathogens. *Cell. Mol. Life Sci.* **56**:755–763.
  62. van Berkum, P. 1990. Evidence for a third uptake hydrogenase phenotype among the soybean bradyrhizobia. *Appl. Environ. Microbiol.* **56**:3835–3841.
  63. van Berkum, P., D. Beyene, and B. D. Eardly. 1996. Phylogenetic relationships among *Rhizobium* species nodulating the common bean *Phaseolus vulgaris* L. *Int. J. Syst. Bacteriol.* **46**:240–244.
  64. van Berkum, P., and B. D. Eardly. 2002. The aquatic budding bacterium *Blastobacter denitrificans* is a nitrogen-fixing symbiont of *Aeschynomene indica*. *Appl. Environ. Microbiol.* **68**:1132–1136.
  65. van Berkum, P., and B. D. Eardly. 1998. Molecular evolutionary systematics of the *Rhizobiaceae*. In H. Spaink, A. Kondorosi, and P. Hooykaas (ed.), *The Rhizobiaceae*. Kluwer Academic Pub., Dordrecht, The Netherlands.
  66. van Berkum, P., and J. J. Fuhrmann. 2000. Evolutionary relationships among the soybean bradyrhizobia reconstructed from 16S rRNA gene and internally transcribed spacer region sequence divergence. *Int. J. Syst. Evol. Microbiol.* **50**:2165–2172.
  67. van Berkum, P., F. Ruihua, T. A. Campbell, and B. D. Eardly. 1999. Some issues of relevance in the taxonomy of rhizobia, p. 267–270. In E. Martinez and G. Hernandez (ed.), *Highlights of nitrogen fixation research*. Plenum Publishing Corp., New York, N.Y.
  68. Wang, Y., and Z. Zhang. 2000. Comparative sequence analyses reveal frequent occurrence of short segments containing an abnormally high number of non-random base variations in bacterial rRNA genes. *Microbiology* **146**:2845–2854.
  69. Wang, Y., Z. Zhang, and N. Ramanan. 1997. The actinomycete *Thermobispora bispora* contains two distinct types of transcriptionally active 16S rRNA genes. *J. Bacteriol.* **179**:3270–3276.
  70. Willems, A., R. Coopman, and M. Gillis. 2001. Phylogenetic and DNA-DNA hybridization analyses of *Bradyrhizobium* species. *Int. J. Syst. Bacteriol.* **51**:111–117.
  71. Winkler, M. E. 1979. Ribosomal ribonucleic acid isolated from *Salmonella typhimurium*: absence of the intact 23S species. *J. Bacteriol.* **139**:842–849.
  72. Yap, W. H., Z. Zhang, and Y. Wang. 1999. Distinct types of rRNA operons exist in the genome of the actinomycete *Thermomonospora chromogena*. *J. Bacteriol.* **181**:5201–5209.
  73. Young, J. M., L. D. Kuykendall, E. Martinez-Romero, A. Kerr, and H. Sawada. 2001. A revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the inclusion of all species of *Agrobacterium* Conn 1942 and *Allorhizobium undicola* de Lajudie et al. 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi*, *R. undicola* and *R. vitis*. *Int. J. Syst. Bacteriol.* **51**:89–103.