

Phylogenetic Analysis of the Spirochetes

B. J. PASTER,^{1*} F. E. DEWHIRST,¹ W. G. WEISBURG,² L. A. TORDOFF,¹ G. J. FRASER,¹ R. B. HESPELL,³
T. B. STANTON,⁴ L. ZABLEN,⁵ L. MANDELCO,⁶ AND C. R. WOESE⁶

Forsyth Dental Center, Boston, Massachusetts 02115¹; GENE-TRAK Systems, Framingham, Massachusetts 01701²;
Northern Regional Center, U.S. Department of Agriculture, Peoria, Illinois 61604³; National Animal Disease
Center, U.S. Department of Agriculture, Ames, Iowa 50010⁴; Abbott Laboratories, Abbott Park,
Illinois 60064⁵; and Department of Microbiology, University of Illinois, Urbana, Illinois 61801⁶

Received 30 March 1991/Accepted 22 July 1991

The 16S rRNA sequences were determined for species of *Spirochaeta*, *Treponema*, *Borrelia*, *Leptospira*, *Leptonema*, and *Serpula*, using a modified Sanger method of direct RNA sequencing. Analysis of aligned 16S rRNA sequences indicated that the spirochetes form a coherent taxon composed of six major clusters or groups. The first group, termed the treponemes, was divided into two subgroups. The first treponeme subgroup consisted of *Treponema pallidum*, *Treponema phagedenis*, *Treponema denticola*, a thermophilic spirochete strain, and two species of *Spirochaeta*, *Spirochaeta zuelzeriae* and *Spirochaeta stenostrepta*, with an average interspecies similarity of 89.9%. The second treponeme subgroup contained *Treponema bryantii*, *Treponema pectinovorum*, *Treponema saccharophilum*, *Treponema succinifaciens*, and rumen strain CA, with an average interspecies similarity of 86.2%. The average interspecies similarity between the two treponeme subgroups was 84.2%. The division of the treponemes into two subgroups was verified by single-base signature analysis. The second spirochete group contained *Spirochaeta aurantia*, *Spirochaeta halophila*, *Spirochaeta bajacaliforniensis*, *Spirochaeta litoralis*, and *Spirochaeta isovalerica*, with an average similarity of 87.4%. The *Spirochaeta* group was related to the treponeme group, with an average similarity of 81.9%. The third spirochete group contained borrelias, including *Borrelia burgdorferi*, *Borrelia anserina*, *Borrelia hermsii*, and a rabbit tick strain. The borrelias formed a tight phylogenetic cluster, with average similarity of 97%. The borrelia group shared a common branch with the *Spirochaeta* group and was closer to this group than to the treponemes. A single spirochete strain isolated from the shrew constituted the fourth group. The fifth group was composed of strains of *Serpula* (*Treponema*) *hyodysenteriae* and *Serpula* (*Treponema*) *innocens*. The two species of this group were closely related, with a similarity of greater than 99%. *Leptonema illini*, *Leptospira biflexa*, and *Leptospira interrogans* formed the sixth and most deeply branching group. The average similarity within this group was 83.2%. This study represents the first demonstration that pathogenic and saprophytic *Leptospira* species are phylogenetically related. The division of the spirochetes into six major phylogenetic clusters was defined also by sequence signature elements. These signature analyses supported the conclusion that the spirochetes represent a monophyletic bacterial phylum.

The spirochetes are one of the few major bacterial groups whose natural phylogenetic relationships are evident at the level of gross phenotypic characteristics (58, 59). Spirochetes possess a cellular ultrastructure which is unique among eubacteria (13, 24). The helical protoplasmic cylinder is encased by an outer sheath which has some features analogous to the outer membrane of gram-negative bacteria. Spirochetes possess internal organelles of motility called periplasmic flagella (37) which are located between the protoplasmic cylinder and the outer sheath. However, in later stages of growth, the periplasmic flagella of some spirochetal species have been shown to be external (44). Periplasmic flagella are ultrastructurally and chemically similar to the external flagella found in other eubacteria (10, 24). Depending on the spirochete species, the number of flagella varies from 2 to hundreds per cell. Other unifying characteristics of spirochetes are resistance to the antibiotic rifampin (1, 35, 50, 55) and (except for the leptospire) the presence of ornithine in cell wall peptidoglycan (30). Earlier attempts to produce a phylogenetic systematics for the spirochetes, using the rRNA oligonucleotide cataloging method (41), showed the spirochete phylum to fall into five major sub-

classes: the genus *Treponema*, the genus *Spirochaeta*, the borrelias, the leptospire, and *Treponema hyodysenteriae*. Recently, *T. hyodysenteriae* and its nonpathogenic relative *T. innocens* were reclassified in the new genus *Serpula* (51), and henceforth these two species will be discussed as *Serpula* species. The phylogenetic positions of *Cristispira* spp. found in mollusks and spirochetes observed in termite hindguts were not determined because these spirochetes could not be cultured in vitro (7, 8). The phylogenetic structure determined from 16S rRNA cataloging is in approximate agreement with the accepted classical taxonomy for spirochetes (7, 11, 12, 28, 32, 48), in which the order *Spirochaetales* is divided into two families, *Spirochaetaceae* and *Leptospiraceae*, with the family *Spirochaetaceae* comprising the genera *Treponema* (including *Serpula hyodysenteriae* and related species), *Spirochaeta*, *Borrelia*, and *Cristispira* and the family *Leptospiraceae* encompassing the genera *Leptospira* and *Leptonema*.

Analysis of nearly complete 16S rRNA sequences (>95%) is a more powerful and accurate method for determining phylogenetic relationships than the earlier oligonucleotide cataloging methods (58). The goal of this study was to use 16S rRNA sequencing to extend our understanding of spirochetal phylogeny.

* Corresponding author.

TABLE 1. Sources and accession numbers of the strains sequenced

Bacterial strain	Source ^a	Reference(s)	GenBank accession no.
<i>Spirochaeta aurantia</i> J1	ATCC 25082 t	14	M57740
<i>S. bajacaliforniensis</i> BA2	ATCC 35968 t	16	M71239
<i>S. halophila</i> RS1	ATCC 29478 t	18	M34262
<i>S. isovalerica</i> MA-2	ATCC 33939 t	20	M34260
<i>S. litoralis</i> R1	ATCC 27000 t	22	M34263
<i>S. stenostrepta</i> Z1	ATCC 25083 t	21	M34264
<i>S. zuelzeriae</i>	ATCC 19044 t	53	M34265
<i>Treponema bryantii</i> RUS1	ATCC 33254 t	39, 50	M57737
<i>Treponema</i> strain CA	Hespell	39	M59294
<i>T. denticola</i> W	ATCC 33520	23	M71236
<i>T. denticola</i> TD10	Hespell	23	Not deposited
<i>T. pectinovorum</i> D-36-DR-2	ATCC 33768 t	39	M71237
<i>T. pallidum</i> Nichols strain	Fitzgerald	48	M34266
<i>T. phagedenis</i> strain K5	Stanton	48	M57739
<i>T. saccharophilum</i> PB	ATCC 43261 t	39	M71238
<i>T. succinifaciens</i> 6091	ATCC 33096 t	39	M57738
<i>Serpula hyodysenteriae</i> B78	ATCC 27164 t	51	M57741
<i>S. innocens</i> B256	ATCC 29796 t	51	M57744
<i>Borrelia hermsii</i> M1001	Johnson	4	M72398
<i>B. burgdorferi</i> B31	ATCC 35210 t	4	M59293
<i>B. anserina</i> strain ES	Johnson	15	M72397
Rabbit tick spirochete 19941	Johnson	3	M72396
<i>Leptospira</i> turtle strain A-183	Charon	6	M34261
<i>L. biflexa</i> serovar patoc strain Patoc I	ATCC 23582	6	Not deposited
<i>L. interrogans</i> serovar pomona strain Kennewicki	Charon	29	M71241
<i>L. interrogans</i> serovar canicola strain Moulton	ATCC 23606	17	X17547
<i>Leptonema illini</i> strain 3055	Charon	19	M34118
Shrew spirochete CT11616	ATCC 43811	2	M59179
Thermophilic strain H1	Leschine	This study	M71240

^a Bacterial strains were obtained from the following sources: ATCC (American Type Culture Collection); collections of R. B. Hespell, T. B. Stanton, and N. Charon (University of West Virginia), R. C. Johnson (University of Minnesota, Minneapolis), T. J. Fitzgerald (University of Minnesota, Duluth), and S. B. Leschine (University of Massachusetts, Amherst).

MATERIALS AND METHODS

Bacterial strains and culture conditions. The bacterial strains used are listed in Table 1. Except for the thermophilic strain described below, the culture conditions can be found in the references listed in Table 1. The isolation and culture of thermophilic spirochete strain H1 have not previously been published. It was isolated from Hunter Hot Spring, Oregon. Spring temperature was 53°C. Growth medium contained, per liter, 1.0 g of yeast extract (Difco Laboratories, Detroit, Mich.), 2.0 g of tryptone (Difco), 1 g of cellobiose (Sigma Chemical Co., St. Louis, Mo.), 1 g of maltose (Sigma), 0.15 g of dithiothreitol, and 10 g of agar adjusted to a pH of 7.0. Rifampin was not used as a selective agent for isolation since spirochetes, as judged by morphological criteria, were the predominant flora in specimens taken at the time of sampling. Samples were diluted serially into agar deeps and were incubated anaerobically at 52°C. Strain H1 grew optimally at 50°C but was capable of growth at temperatures ranging from 27 to 58°C.

Isolation and purification of RNA. rRNA was isolated and partially purified by a modification of the procedure of Pace et al. (37) as previously described (40); alternatively, cells were disrupted with a Mini-BeadBeater (Biospec Products, Bartlesville, Okla.) directly into phenol for RNA extraction. This step was followed by purification through cesium trifluoroacetate centrifugal gradients (Pharmacia Fine Chemicals AB) as previously described (57).

16S rRNA sequencing. In most cases, rRNA was sequenced directly, using a modification of the standard

Sanger dideoxynucleotide chain termination technique in which reverse transcriptase was used to elongate primers complementary to conserved regions of the 16S rRNA (34). Seven previously described primers were used to obtain nearly complete sequences (56). To obtain the sequence of the 5' end of the molecule, we used an eighth additional primer (AAGCAUGCAAGUCGARGCGG) that was complementary to positions 51 to 69 of the 16S rRNA (using the nucleotide numbering system of *Escherichia coli*) (9). Further descriptions of these procedures are presented in previous publications (40, 56).

The 16S rRNA gene of *Treponema pallidum* was sequenced from a ribosomal DNA clone. It was cloned initially as a *Sau3AI* partial digest into the *BamHI* site of lambda phage 1059. The resulting clone began at the nearly universal *Sau3AI* site located at position 15 (*E. coli* numbering). For subcloning into M13mp8, M13mp9, M13mp18, and M13mp19, the ribosomal DNA was divided at the *EcoRI* site occurring at position 680 (*E. coli* numbering). The *T. pallidum* sequence was determined from these subclones by methods similar to those used for direct rRNA sequencing (described above) but with additional sequence information gained by the use of forward primers (57). Either Klenow fragment or T7 DNA polymerase (Sequenase; U.S. Biochemical Corp.) was used for chain elongation.

Spirochaeta halophila was sequenced partially from a clone and partially by direct reverse transcriptase sequencing. The portion of the gene 3' to the *EcoRI* site at position 680 was derived from a DNA sequence, and the 5' 680

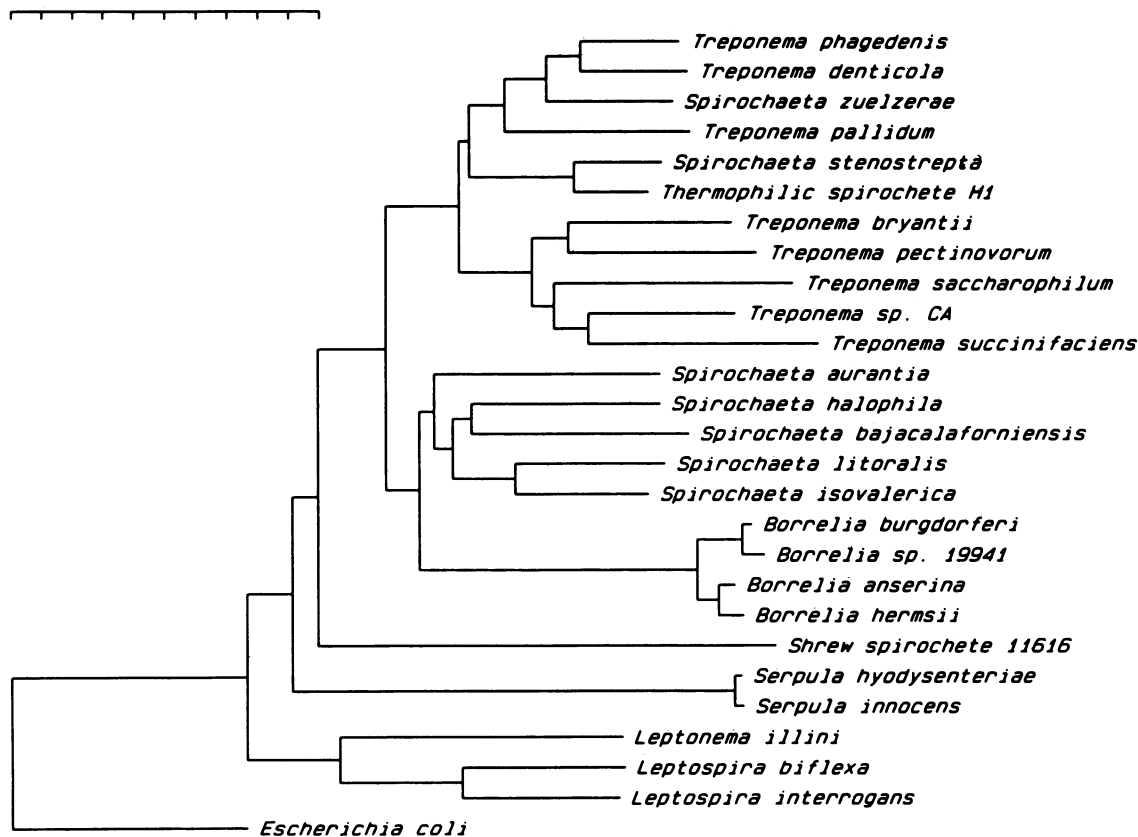


FIG. 1. Phylogenetic tree of the spirochetes, with *E. coli* as the outgroup. The scale bar represents a 10% difference in nucleotide sequence as determined by taking the sum of all of the horizontal lines connecting two species.

nucleotides were derived from rRNA templates. The cloned sequence resulted from an initial *Bam*HI clone in lambda L47 which was subcloned into M13mp18 (46).

Data analysis. A program set for data entry, editing, sequence alignment, secondary structure comparison, similarity matrix generation, and tree construction for 16S rRNA data was written in Microsoft QuickBASIC for use on IBM PC-AT and compatible computers. RNA sequences were entered and aligned as previously described (40). Only those base positions which were determined for every sequence and which could be unequivocally aligned were used in the distance matrix calculations. In addition, hypervariable regions for spirochete sequences were also omitted for these calculations. Using the numbering nomenclature for the 16S rRNA of *E. coli* (9), these regions correspond to base positions 69 to 100, 185 to 195, 201 to 219, 455 to 478, and 840 to 848. The total number of base comparisons for calculation of distance matrices was 1,176. Presently, our data bases contain approximately 350 sequences, including those determined in our laboratories, published sequences, and unpublished sequences provided to us by other investigators. Similarity matrices were corrected for multiple base changes at single locations by the method of Jukes and Cantor (31). Dendrograms were constructed by using the neighbor-joining method of Saitou and Nei (45).

Nucleotide sequence accession numbers. The sequences for the bacterial strains examined are available for electronic retrieval from the EMBL, GenBank, and DDBJ nucleotide sequence data bases under the accession numbers listed in Table 1.

RESULTS AND DISCUSSION

The sequence of approximately 95% of the 16S rRNA molecule (1,400 to 1,500 bases) was determined for each of the spirochete strains listed in Table 1. Sequences for *Leptospira biflexa* serovar patoc and turtle strain A-183 were identical. The relatively complete sequence of *Leptospira interrogans* serovar pomona (1,136 bases) differs by one base from the previously published 16S rRNA sequence of *L. interrogans* serovar canicola. The gap in serovar canicola sequence at *E. coli* base position 386 (after position 370 in reference 17) is a C in the serovar pomona sequence. Since the complete 16S rRNA sequence was available for serovar canicola (17), it was used for phylogenetic analysis. The two *Treponema denticola* strains were also identical.

A sequence similarity matrix for 16S rRNAs of the spirochetes and of *E. coli* as an outgroup is presented in Table 2. A dendrogram derived from the similarity matrix using the neighbor-joining method is shown in Fig. 1. The branching order of the six major spirochete phylogenetic groups remained stable when hypervariable regions were included in the distance matrix calculations. The phylogenetic relationships obtained from the 16S rRNA sequence analysis are in overall agreement with previous catalog studies (41) and comparative sequence analysis of the *trpE* gene (5). Characteristics of known genera of spirochetes with respect to physiology, ecological niche, DNA base composition, and other traits are presented in Table 3.

Members of the genus *Treponema*, *Spirochaeta zuelzeræ*, *S. stenostrepta*, and thermophilic spirochete strain H1 form

TABLE 3. Characteristics of spirochetal genera^a

Trait	Spirochete genus					
	<i>Treponema</i>	<i>Spirochaeta</i>	<i>Borrelia</i>	<i>Leptonema/Leptosira</i>	<i>Serpula</i>	<i>Cristispira</i>
Mol% G+C	36-54	51-65	27-32	37-53	26	?
Relationship to O ₂	Obligate anaerobe	Obligate and facultative anaerobe	Microaerophilic	Aerobic	Microaerophilic	?
Size (μm)	0.3-0.7 by 5-20	0.3 by 10-20 ^b	0.2-0.5 by 8-30	0.2 by 10	0.3-0.4 by 7-9	0.5-3 by 20-100
No. of flagella	2-32	2	15-30	2	18	>100
Habitat	Host-associated (humans and other mammals)	Free-living, fresh and salt H ₂ O	Host associated (insects; humans and other mammals)	Free-living and host associated (humans and other mammals; reptiles)	Host associated (swine)	Digestive tract of mollusks
Pathogenic species	+ and -	-	+	+ and -	+ and -	?
Resistance to rifampin	+	+	+	+	+	?
Ornithine in peptidoglycan	+	+	+	-	?	?
Optimum growth temp (°C)	37-39	15-40 ^c	37	15-37	36-42	?

^a Information obtained from references 4, 7, 11, 12, 13, 47, and 50.

^b *Spirochaeta plicatilis*, which has not been cultured in vitro, is 0.75 by 80 to 250 μm.

^c Thermophilic strain H1 (this study) and other thermophilic spirochetes have been reported (42), and optimum growth temperatures range from 45 to 60°C. These spirochetes are presently classified as belonging to the genus *Spirochaeta* since they are considered free-living. However, strain H1 is related more closely to the treponemes than to other free-living spirochetes (see text).

a phylogenetic cluster, referred to as the treponemes, with an average similarity of 86.1%. The tree topology shows that this cluster is divided into two subgroups. The first subgroup is composed of *T. pallidum*, *T. phagedenis*, the human oral *T. denticola*, *S. stenostrepta*, thermophilic strain H1, and *S. zuelzeriae*, with an average similarity of 89.9%. The second subgroup contains the rumen spirochetes, *Treponema bryantii*, *T. saccharophilum*, and strain CA, the pectinolytic human oral *T. pectinovorum*, and the swine intestinal *T. succinifaciens*, with an average similarity of 86.2%. The average level of similarity between these two subgroups is 84.2%. The separation of the treponemes into two subgroups was verified by base signature analysis. Bases common to each of the two subgroups are shown in Table 4. The significance of this subdivision within the treponemes remains uncertain.

Percent similarity and single-base signature analysis demonstrated that the free-living spirochetes, *S. stenostrepta*, thermophilic strain H1, and *S. zuelzeriae*, are more closely related to the treponemes than to other members of the

genus *Spirochaeta* (Table 5; Fig. 1). As shown in Table 5, *S. stenostrepta*, strain H1, and *S. zuelzeriae* possess more bases in common with the sequences of the treponemes than with the sequences of *Spirochaeta* species. It is tempting to speculate that the free-living spirochetes, *S. stenostrepta*, strain H1, and *S. zuelzeriae*, may represent descendants of the precursors to host-associated treponemes. However, since each species represents a single strain that was isolated from aquatic sources, it is possible that they are fecal contaminants from an unknown host and are true treponemes. To resolve this issue, many strains of each species should be isolated from either aquatic or host environments.

TABLE 4. Signature analysis of treponeme subgroups

Position of base or pair	Composition in:	
	<i>T. phagedenis</i> subgroup ^a	<i>T. bryantii</i> subgroup ^b
133-230	C-G	G-C
421	A	G
478	A	G
502-543	C-G	R-Y
661-744	R-Y	Y-R
733	G	A
835-851	G-C	C-G
1310-1327	A-U	G-C

^a Contains *T. phagedenis*, *T. denticola*, *S. zuelzeriae*, *T. pallidum*, *S. stenostrepta*, and strain H1.

^b Contains *T. bryantii*, *T. pectinovorum*, *T. saccharophilum*, strain CA, and *T. succinifaciens*.

TABLE 5. Relationship of *S. stenostrepta*, thermophilic strain H1, and *S. zuelzeriae* to the treponemes

Base position	Base composition ^a in:			
	Treponemes	Strain H1 and <i>S. stenostrepta</i>	<i>S. zuelzeriae</i>	Other <i>Spirochaeta</i> spp.
108	G	G	G	A
111	U	U	U	G
241	C	C	C	U
242	Y	C	C	G
285	G	G	G	A
289	Y	G	U	G
311	R	<u>C</u>	A	<u>C</u>
422	A	<u>A</u>	A	<u>G</u>
768	A	<u>G</u>	<u>G</u>	<u>A</u>
986	Y	<u>U</u>	<u>U</u>	<u>A</u>
1001	R	G	G	Y
1118	C	C	C	U
1155	G	G	G	A
1157	A	A	A	G
1219	R	A	A	U
1335	Y	G	U	G
1414	U	<u>U</u>	U	<u>C</u>

^a Underlined bases represent signature for other *Spirochaeta* spp.

TABLE 6. Spirochetal base signatures^a

Position of base or pair	Composition in spirochetes	Composition in other bacteria	Exception(s) ^b
28-555	A·U	G·C	A·U prpl, plnc G·U myco, flavo U·G (1); C·G (1)
45-396	C·G ex. G·C Tsa U·A Tpe	G·C or U·A	
47	U	C	G plnc, chlām
50	U	A	U prpl, plnc (2), ttoga G chlām
52-359	A·U ex. G·C shr	C·G or U·A	G·U dein
53-358	G·C ex. A·U lepto	A·U	G·U plnc, chlām G·C <i>Chloroflexus</i>
361	A	G	A plnc, chlām
1415-1485	C·G	G·Y	Y·R plnc, myco C·G isocy, ttoga (4)

^a Abbreviations: prpl, purple bacteria (proteobacteria); plnc, planctomyces; myco, mycoplasma; flavo, flavobacterium; chlām, chlamydia; dein, deinococcus; ttoga, *Thermotoga* sp.; isocy, *Isocystis pallida*; lepto, *Leptonema* and *Leptospira* spp.; ex., exception(s).

^b Numbers in parentheses represent the number of strains that possess the indicated base signatures.

16S rRNA-based DNA probes that are specific for each of these species will facilitate the screening for these spirochetes.

The closest relative of *S. stenostrepta* is the thermophilic strain H1, with 95.1% similarity. Other anaerobic thermophilic spirochetes isolated from thermal springs have been described recently (42). It will be interesting to determine whether these thermophilic strains cluster also with *S. stenostrepta* or whether they occur in multiple branches of the phylogenetic tree.

The phylogenetic position of *T. pallidum*, the type strain of the genus *Treponema*, has been uncertain since DNA-DNA homology studies have found little or no homology between *T. pallidum* and other treponemes (36). This is not surprising, however, since DNA-DNA hybridizations are most useful in determining the genetic relationships of closely related species, i.e., 16S rRNA sequence differences of less than 5%. The 16S rRNA sequences of treponemes differ by an average of almost 10%. For determining phylogenetic relationships between distantly related organisms (even prokaryote to eukaryote), 16S rRNA sequence similarity is the most powerful method currently available (58). As shown in Fig. 1, it is clear that *T. pallidum* does indeed belong within the genus *Treponema*.

Spirochaeta aurantia, *S. halophila*, *S. litoralis*, *S. isovalerica*, and *S. bajacaliforniensis* form a cluster separate from the treponemes at an average level of similarity of 81.9%. Within the *Spirochaeta* cluster, the average level of similarity is 87.4%. It has been suggested that *S. bajacaliforniensis* may be a strain of *S. litoralis*, given similarities in physiology and G+C content of their DNAs (13). However, 16S rRNA data suggest that *S. bajacaliforniensis* warrants a separate species designation within the genus *Spirochaeta*.

The borrelias are more closely related to members of the *Spirochaeta* cluster (82.2% similarity) than to the treponemes (77.9% average similarity) and share a short common branch with the *Spirochaeta* cluster (Fig. 1). These results

TABLE 7. Differentiation of major spirochetal phylogenetic groups by base signature analysis

Position of base or pair	Base composition ^a in:					
	Treponemes	<i>Spirochaeta</i>	<i>Borrelia</i>	Shrew	<i>Serpula</i>	Leptospire
108	<u>G</u>	A	A	G	A	A
111	<u>U</u>	G	G	G	G	U
1025-1035	<u>G·C</u>	Y·G	C·G	—	—	—
126-235	<u>G·Y</u>	<u>A·U</u>	G·U	A·U	A·U	R·Y
986-1219	Y·R	<u>A·U</u>	A·U	A·U	Y·R	Y·R
7	G	<u>G</u>	<u>A</u>	G	G	G
41-401	G·C	G·C	<u>A·U</u>	A·N	G·C	G·C
1245-1292	R·Y	G·C	<u>U·A</u>	C·G	A·U	G·C
52-359	A·U	A·U	<u>A·U</u>	<u>G·C</u>	A·U	A·U
783-799	C·G	C·G	C·G	<u>U·A</u>	C·G	C·G
36-548	C·G	C·G	C·G	<u>C·G</u>	<u>U·A</u>	C·G
787	A	A	A	A	C	A
655-751	A·U	A·U	A·U	A·U	<u>G·C</u>	A·U
562	U	U	U	U	<u>U</u>	C
53-358	G·C	G·C	G·C	G·C	G·C	<u>A·U</u>
369-392	C·G	C·G	C·G	C·G	C·G	<u>G·C</u>

^a Underlined base(s) represents signature for each phylogenetic group. —, gap at the indicated base position.

differ slightly from previous rRNA oligonucleotide catalog comparisons in which the borrelias were found to be equally distant from the treponemes and the *Spirochaeta* species (41). The borrelias comprise a very tight cluster, with average interspecies similarities of greater than 97%. These results are in agreement with DNA-DNA reassociation studies which demonstrated 30 to 70% homologies among *Borrelia* species (4, 26, 47).

Previous studies demonstrated that spirochetes isolated from the shrew and white-footed mouse are serologically and ultrastructurally distinct from species of *Treponema*, *Borrelia*, *Leptospira*, and *Spirochaeta* (2). On the basis of these data, the authors suggested that these spirochetes may represent a new genus. This suggestion is supported by 16S rRNA sequence analysis in that the shrew isolate branched separately from other spirochetes (Fig. 1).

The strains of *S. hyodysenteriae* and *S. innocens* form a very tight cluster (e.g., greater than 99% sequence similarity) which is only distantly related to the other spirochetes. Because these strains do not branch with other *Treponema* species, it was suggested (41) that they were clearly misclassified as members of the genus *Treponema* and were renamed as species of the genus *Serpula* (51). The sequence analysis is consistent with DNA-DNA hybridization studies and analyses of protein profiles which failed to demonstrate a relationship of *S. hyodysenteriae* and *S. innocens* to other treponemes (36, 51).

Lastly, the leptospire form the most deeply branching cluster within the spirochetal phylum and are divided into two subgroups. The first subgroup contains the saprophytic leptospire (*L. biflexa* serovar patoc and turtle strain A-183 [which had identical sequences]) and the parasitic leptospire (*L. interrogans* serovar canicola [17] and serovar pomona [which had nearly identical sequences]), with 90.3% similarity. The identical sequences of the patoc serovar and strain A-183 were not surprising inasmuch as the patoc strain had approximately 50% DNA relatedness with strain A-183 (6). The other subgroup consists of the single species, *Leptonema illini*, with an average level of 83.2% sequence similarity with members of the first subgroup. As suggested

TABLE 8. 5' extension of the 16S rRNA molecule for *Spirochaeta* spp., *Treponema* spp., and leptospire

Organism	5' extension ^a
<i>E. coli</i> ^b	AAAUGAAGAGUU
Treponeme cluster	
<i>T. phagedenis</i>	CAGGCUCUCUUA-GCAA-UAGGAGAGnUUGAAAUA-UAAUAAUGGAGAGUU
<i>T. denticola</i>	nAACCGCUCUUAA-GGGCGGCUU-GAAAUAUAAUGAUGGAGAGUU
<i>S. zuelzeriae</i>	nnUCUUGAUCC-GCAA-GGGAUUUAGA-GnA-UA-UA-UCAUGGAGAGUU
<i>S. stenostrepta</i>	nCUUU-GGUC-UUCG-GnCC-guaA-GAAAUA-UA-UCAUGGnnAgUU
Spirochete strain H1	cUUUCGGGC-UUCG-GGCCcGUAA-GAnAUA-UA-UCAUGGAGAGUU
<i>T. bryantii</i>	AUUAGCC-UUCG-GGCgAUU-GAAAUA-UCAUGGAGAGUU
<i>T. pectinovorum</i>	UUGACC-GCAA-GGGUUUGA-GAnAUA-UCAUGGAGAGUU
<i>T. sacchrophilum</i>	CUUUC-UUCG-GGGnUUA-GAA-UAA-UCAUGGAGAGUU
Treponema sp. strain CA	nnncU-uacG-GGgnnUU-GAA-UAA-UCAUGGAGAGUU
<i>T. succinifaciens</i>	CUCCGGUC-CUCG-GGGCCGAG-GAAAUAGG-CAUGGAGAGUU
Conserved regionGAAUA...u.AUGGAGAGUU
Spirochaeta cluster	
<i>S. aurantia</i>	nCCUUGGA-GCaA-UCCAAGGuau-----AUUGAUGGAGAGUU
<i>S. bajacaliforniensis</i>	UCAACAUC-UUCG-GAUGunAACUA-----CUAUAUGGAGAGUU
<i>S. halophila</i>	nCGUCC-UUCG-GGGACGUACA-----AAUCAUGGAGAGUU
<i>S. litoralis</i>	naggGUC-UUCG-GacnnnnAU-----aUAUCAUGGAGAGUU
<i>S. isovaleric</i>	nnUCGCAC-UUCG-GUGUnnUnnAU-----nAAUCAUGGAGAGUU
Conserved regionAU.AUGGAGAGUU
Leptospire cluster	
<i>Leptonema illini</i>	UAAUCU-AUC-GCAA-GAU-AGAU-----CAUAACGGAGAGUU
<i>Leptospira biflexa</i>	AAAUGGUA-GCAA-UACCAGAU-----CAUAACGGAGAGUU
Conserved region	.AA.U.U.GCAA.A.AGAUU.....CAUAACGGAGAGUU

^a Potential helices are underlined.

^b The sequence represents the 5' end of the 16S rRNA molecule for *E. coli* beginning at position 1 (9).

previously (41), these results are consistent with the leptospiral taxonomy as proposed by Hovind-Hougen (25). The author suggested that the now accepted family *Leptospiraceae* be divided into two genera, *Leptospira* (which would include the parasitic and saprophytic leptospire) and *Leptonema* (which would include *L. illini*). Recent studies have examined the DNA relatedness among the many serovars of *L. biflexa* and *L. interrogans* (43, 60). These studies revealed that many strains within the two species were significantly heterogeneous, and consequently seven new *Leptospira* species were proposed (60). However, by using DNA hybridization techniques, it was impossible to detect the genetic relationship between the saprophytic and parasitic leptospire. The present study represents the first demonstration that the saprophytic and parasitic leptospire are related phylogenetically. Obviously, analysis of 16S rRNA sequences of the newly proposed species and other strains is necessary to establish a more complete phylogenetic structure for this diverse grouping of spirochetes.

Because of the depth of the branching within the spirochete phylum, some spirochete species are as similar to enterobacteria as to other spirochetes. For example, the leptospire, *Serpula* species, and the shrew isolate are more closely related to *E. coli* than to *Treponema succinifaciens* (Table 2), although the overall tree topology indicates that they are derived from spirochetal ancestry. To verify that the spirochetes are of monophyletic origin, individual base signatures that are unique to the spirochetes were identified. Table 6 provides most of these key base signatures for the spirochete sequences. The most significant base signature is the U at position 47, which occurs in all spirochete sequences. In over 350 other eubacterial and archaeobacterial sequences, a C is present at position 47; the exceptions are

chlamydiae and relatives of the planctomyces, which possess a G (56, 58). The other spirochetal base signatures listed in Table 6 are not unique for all spirochetes or are occasionally found in other eubacterial sequences, but they nevertheless present a strong argument for inclusion of all spirochetes into a single monophyletic phylum.

Each of the six major phylogenetic clusters discussed above was also defined by single-base signature analysis. Examples of base signatures for each spirochete cluster are given in Table 7. Not all of the signatures for each cluster are shown. Since the shrew isolate represents the sole member of its phylogenetic cluster, additional members of this group will have to be analyzed to verify its signature elements.

The sequences of the species in the treponemes, *Spirochaeta*, and leptospire clusters are unusual in that they possess a 20- to 30-base 5' extension of the 16S rRNA molecule. As shown in Table 8, the sequences of the extensions are highly variable. Secondary structure analysis suggests that the extensions may form a helix. The potential helices contain 2 to 12 bp with a terminal loop of 4 or 5 nucleotides. The terminal loop generally consisted of the stable sequence UUCG or GCAA (52). Since the sequence of *T. pallidum* was obtained from a clone lacking the 5' end of the molecule, the existence and sequence of such an extension for *T. pallidum* remain unknown. The published sequence of *L. interrogans* serovar canicola did not include sequence information 5' to position 1 of *E. coli* (17).

The 5' extension found in spirochete 16S rRNA is unique among procaryotes and eucaryotes. While its origin and function are unknown, it may reflect some defect or alteration in the processing and maturation of the 16S rRNA transcript (33). Procaryotes usually contain the rRNA genes grouped together in 1 to 10 nearly identical operons. The operons in *E. coli* are organized with the small-subunit RNA

gene (16S rRNA) located at the 5' end, a tRNA gene, the large-subunit RNA gene (23S rRNA), and finally the 5S rRNA gene at the 3' end (9). The genes are separated by spacer RNA. The spacer RNA adjacent to the 16S and 23S RNA genes forms characteristic structures that include double-stranded helical stems. The processing of the pre-rRNA transcript to mature rRNAs is complex and involves multiple enzymes (33, 49). RNase III cleaves the double-stranded stems to separate the rRNA subunits from one another and from the tRNA element. Poorly characterized endonucleases affect the final 3' and 5' cleavages of the 16S rRNA. Examination of the rRNA operons of spirochetal species with and without the 5' extension is needed to clarify the origin and function of the 5' extension.

As recommended by the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics, phylogenetically based taxonomic schemes should reflect phenotypic consistency (54). The spirochetal phylum easily satisfies these criteria based on the unique cellular ultrastructure of all spirochetes. Within this phylum, the spirochetes form six distinct but deeply branching groups: *Treponema* (including *S. zuelzeriae*, *S. stenostrepta*, and strain H1), *Spirochaeta*, *Borrelia*, *Serpula*, the shrew isolate, and leptospirae. The difference between spirochete genera is greater than the distance between the families *Enterobacteriaceae* and *Pasteurellaceae*. Despite the depth of branching, we suggest that the phylogenetic clustering of spirochetal groups, rather than a specific percent similarity, be used as a basis for defining spirochetal genera. We would agree with the suggestion that the exact 16S rRNA similarity limits for defining a given taxon will have to be determined for each bacterial phylum (27).

ACKNOWLEDGMENTS

We thank T. Fitzgerald for the growth and purification of *T. pallidum*. We also thank the following investigators for supplying bacterial cell pellets: Nyles Charon, *L. interrogans* serovar pomona; G. Hinkle, *S. bajacaliforniensis*; and M. Pohlschröder, thermophilic spirochete strain H1.

This work was supported by grants DE-04881 and DE-08303 from the National Institutes of Health and by grant BSR 87-05352 from the National Science Foundation.

ADDENDUM IN PROOF

The genus designation *Serpula* should be changed to *Serpulina*; a correction will be submitted.

REFERENCES

- Allan, B., E. P. Greenberg, and A. Kropinski. 1986. DNA-dependant RNA polymerase from *Spirochaeta aurantia*. FEMS Microbiol. Lett. 35:205-210.
- Anderson, J., R. C. Johnson, L. A. Magnarelli, F. W. Hyde, and T. G. Andreadis. 1987. New infectious spirochete isolated from short-tailed shrews and white-footed mice. J. Clin. Microbiol. 25:1490-1494.
- Anderson, J., L. A. Magnarelli, R. B. LeFebvre, T. G. Andreadis, J. B. McAninch, G.-C. Perng, and R. C. Johnson. 1989. Antigenically variable *Borrelia burgdorferi* isolated from cottontail rabbits and *Ixodes dentatus* in rural and urban areas. J. Clin. Microbiol. 27:13-20.
- Barbour, A. G., and S. F. Hayes. 1986. Biology of *Borrelia* species. Microbiol. Rev. 50:381-400.
- Brahamsa, B., C.-Y. Han, I. P. Crawford, and E. P. Greenberg. 1991. Nucleotide sequence and analysis of a gene encoding anthranilate synthase component I in *Spirochaeta aurantia*. J. Bacteriol. 173:541-548.
- Brendle, J. J., M. Rogul, and A. D. Alexander. 1974. Deoxyribonucleic acid hybridization among selected leptospiral serotypes. Int. J. Syst. Bacteriol. 24:205-214.
- Breznak, J. A. 1984. Genus II. *Cristispira* Gross 1910, 44^{AL}, p. 46-49. In N. R. Krieg and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 1. The Williams & Wilkins Co., Baltimore.
- Breznak, J. A. 1984. Hindgut spirochetes of termites and *Cryptocercus punctulatus*, p. 67-70. In N. R. Krieg and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 1. The Williams & Wilkins Co., Baltimore.
- Brosius, J., T. J. Dull, D. D. Sleeter, and H. F. Noller. 1981. Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. J. Mol. Biol. 148:107-127.
- Canale-Parola, E. 1977. Physiology and evolution of spirochetes. Bacteriol. Rev. 41:181-204.
- Canale-Parola, E. 1984. Order I. *Spirochaetales* Buchanan 1917, 163^{AL}, p. 38-39. In N. R. Krieg and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 1. The Williams & Wilkins Co., Baltimore.
- Canale-Parola, E. 1984. Genus I. *Spirochaeta* Ehrenberg 1835, 313^{AL}, p. 39-46. In N. R. Krieg and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 1. The Williams & Wilkins Co., Baltimore.
- Canale-Parola, E. 1991. Free-living saccharolytic spirochetes: the genus *Spirochaeta*, in press. In A. Balows, H. G. Trüper, M. Dworkin, W. Harder, and K. H. Schleifer (ed.), The prokaryotes: a handbook on the biology of bacteria ecophysiology, isolation, identification, applications, 2nd ed. Springer-Verlag, New York.
- Canale-Parola, E., S. C. Holt, and Z. Udriș. 1976. Isolation of free-living, anaerobic spirochetes. Arch. Mikrobiol. 59:41-48.
- DaMassa, A. J., and H. E. Adler. 1979. Avian spirochetosis: natural transmission by *Argas* (Persicargas) *sanchezi* (Ixodoidea: Argasidae) and existence of different serologic and immunologic types of *Borrelia anserina* in the United States. Am. J. Vet. Res. 40:154-157.
- Francek, S. P., Jr., and J. F. Stolz. 1985. *Spirochaeta bajacaliforniensis* sp. n. from a microbial mat community at Laguna Figueroa, Baja California Norte, Mexico. Arch. Microbiol. 142:317-325.
- Fukunaga, M., I. Horie, N. Okuzako, and I. Mifuchi. 1990. Nucleotide sequence of a 16S rRNA gene for *Leptospira interrogans* serovar *canicola* strain Moulton. Nucleic Acids Res. 18:366.
- Greenberg, E. P., and E. Canale-Parola. 1976. *Spirochaeta halophila*, sp. nov., a facultative anaerobe from a high salinity pond. Arch. Microbiol. 110:185-194.
- Hanson, L. E., D. N. Tripathy, L. B. Evans, and A. D. Alexander. 1974. An unusual leptospira, serotype *illini* (a new serotype). Int. J. Syst. Bacteriol. 24:355-357.
- Harwood, C. S., and E. Canale-Parola. 1983. *Spirochaeta isovalerica* sp. nov., a marine anaerobe that forms branched-chain fatty acids as fermentation products. Int. J. Syst. Bacteriol. 33:573-579.
- Hespell, R. B., and E. Canale-Parola. 1970. Carbohydrate metabolism in *Spirochaeta stenostrepta*. J. Bacteriol. 103:216-226.
- Hespell, R. B., and E. Canale-Parola. 1970. *Spirochaeta litoralis* sp. n., a strictly anaerobic marine spirochete. Arch. Mikrobiol. 74:1-18.
- Hespell, R. B., and E. Canale-Parola. 1971. Amino acid and glucose fermentation by *Treponema denticola*. Arch. Mikrobiol. 78:234-251.
- Holt, S. C. 1978. Anatomy and chemistry of spirochetes. Microbiol. Rev. 42:114-160.
- Hovind-Hougen, K. 1979. *Leptospiraceae*, a new family to include *Leptospira* Noguchi 1917 and *Leptonema* gen. nov. Int. J. Syst. Bacteriol. 29:245-251.
- Hyde, F. W., and R. C. Johnson. 1984. Genetic relationship of Lyme disease spirochetes to *Borrelia*, *Treponema*, and *Leptospira* spp. J. Clin. Microbiol. 20:151-154.
- Johnson, J. L. 1984. Nucleic acids in bacterial classification, p. 8-11. In N. R. Krieg and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 1. The Williams & Wilkins Co.,

- Baltimore.
28. Johnson, R. C., and S. Faine. 1984. Genus I. *Leptospira* Noguchi 1917, 755^{AL}, p. 62–67. In N. R. Krieg and J. G. Holt (ed.), *Bergey's manual of systematic bacteriology*, vol. 1. The Williams & Wilkins Co., Baltimore.
 29. Johnson, R. C., and V. G. Harris. 1967. Differentiation of pathogenic and saprophytic leptospire. I. Growth at low temperatures. *J. Bacteriol.* **94**:27–31.
 30. Joseph, R., S. C. Holt, and E. Canale-Parola. 1973. Peptidoglycan of free-living anaerobic spirochetes. *J. Bacteriol.* **115**:426–435.
 31. Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules, p. 21–132. In H. N. Munro (ed.), *Mammalian protein metabolism*, vol. 3. Academic Press, Inc., New York.
 32. Kelly, R. T. 1984. Genus IV. *Borrelia* Swellengrebel 1907, 582^{AL}, p. 57–62. In N. R. Krieg and J. G. Holt (ed.), *Bergey's manual of systematic bacteriology*, vol. 1. The Williams & Wilkins Co., Baltimore.
 33. King, T. C., R. Sirdeskumukh, and D. Schlessinger. 1986. Nucleolytic processing of ribonucleic acid transcripts in procaryotes. *Microbiol. Rev.* **50**:428–451.
 34. Lane, D. J., B. Pace, G. J. Olsen, D. A. Stahl, M. L. Sogin, and N. R. Pace. 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc. Natl. Acad. Sci. USA* **82**:6955–6959.
 35. Leschine, S. B., and E. Canale-Parola. 1986. Rifampin-resistant RNA polymerase in spirochetes. *FEMS Microbiol. Lett.* **35**:199–204.
 36. Miao, R. M., A. H. Fieldsteel, and D. L. Harris. 1978. Genetics of *Treponema*: characterization of *Treponema hyodysenteriae* and its relationship to *Treponema pallidum*. *Infect. Immun.* **22**:736–739.
 37. Pace, B., E. A. Matthews, K. D. Johnson, C. R. Cantor, and N. R. Pace. 1982. Conserved 5S rRNA complement to tRNA is not required for protein synthesis. *Proc. Natl. Acad. Sci. USA* **79**:36–40.
 38. Paster, B. J., and E. Canale-Parola. 1980. Involvement of periplasmic fibrils in motility of spirochetes. *J. Bacteriol.* **141**:359–364.
 39. Paster, B. J., and E. Canale-Parola. 1982. Physiological diversity of rumen spirochetes. *Appl. Environ. Microbiol.* **43**:686–693.
 40. Paster, B. J., and F. E. Dewhirst. 1988. Phylogeny of campylobacters, wolinellas, *Bacteroides gracilis*, and *Bacteroides ureolyticus* by 16S ribosomal ribonucleic acid sequencing. *Int. J. Syst. Bacteriol.* **38**:56–62.
 41. Paster, B. J., E. Stackebrandt, R. B. Hespell, C. M. Hahn, and C. R. Woese. 1984. The phylogeny of the spirochetes. *Syst. Appl. Microbiol.* **5**:337–351.
 42. Patel, G. K. C., H. W. Morgan, and R. M. Daniel. 1985. Thermophilic anaerobic spirochetes in New Zealand hot springs. *FEMS Microbiol. Lett.* **26**:101–106.
 43. Ramadass, P., B. D. W. Jarvis, R. J. Corner, M. Cinco, and R. B. Marshall. 1990. DNA relatedness among strains of *Leptospira biflexa*. *Int. J. Syst. Bacteriol.* **40**:231–235.
 44. Ruby, J. D., J. Curci, S. Goldstein, and N. W. Charon. 1991. Evidence that the periplasmic flagella of *Treponema denticola* and *Treponema phagedenis* can be external, I-86, p. 204. *Abstr. 91st Gen. Meet. Am. Soc. Microbiol.* 1991.
 45. Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
 46. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed., p. 2.80–2.81. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 47. Schmid, G. P., A. G. Steigerwalt, S. E. Johnson, A. G. Barbour, A. C. Steere, I. M. Robinson, and D. J. Brenner. 1984. DNA characterization of the spirochete that causes Lyme disease. *J. Clin. Microbiol.* **20**:155–158.
 48. Smibert, R. M. 1984. Genus III. *Treponema* Schaudinn 1905, 1728^{AL}, p. 49–57. In N. R. Krieg and J. G. Holt (ed.), *Bergey's manual of systematic bacteriology*, vol. 1. The Williams & Wilkins Co., Baltimore.
 49. Srivastava, A. K., and D. Schlessinger. 1989. *Escherichia coli* 16S rRNA 3'-end formation requires a distal transfer RNA sequence at a proper distance. *EMBO J.* **8**:3159–3166.
 50. Stanton, T. B., and E. Canale-Parola. 1979. Enumeration and selective isolation of rumen spirochetes. *Appl. Environ. Microbiol.* **38**:965–973.
 51. Stanton, T. B., N. S. Jensen, T. A. Casey, L. A. Tordoff, F. E. Dewhirst, and B. J. Paster. 1991. Reclassification of *Treponema hyodysenteriae* and *Treponema innocens* in a new genus, *Serpula*, gen. nov., as *Serpula hyodysenteriae* comb. nov. and *Serpula innocens* comb. nov. *Int. J. Syst. Bacteriol.* **41**:50–58.
 52. Turerk, C., P. Gauss, C. Thermes, D. R. Groebe, M. Gayle, N. Guild, G. Stormo, Y. D'Aubenton-Carafa, O. C. Uhlenbeck, I. Tinoco, Jr., E. N. Brody, and L. Gold. 1988. CUUCGG hairpins: extraordinarily stable RNA secondary structures associated with various biochemical processes. *Proc. Natl. Acad. Sci. USA* **85**:1364–1368.
 53. Veldcamp, H. 1960. Isolation and characteristics of *Treponema zuelzerae* nov. spec., an anaerobic, free-living spirochete. *Antonie van Leeuwenhoek J. Microbiol. Serol.* **26**:103–125.
 54. Wayne, L. G., D. J. Brenner, R. R. Colwell, P. A. D. Grimont, O. Kandler, M. I. Krichevsky, L. H. Moore, W. E. C. Moore, R. G. E. Murray, E. Stackebrandt, M. P. Starr, and H. G. Truper. 1987. Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. *Int. J. Syst. Bacteriol.* **37**:463–464.
 55. Weber, F. H., and E. P. Greenberg. 1981. Rifampin as a selective agent for the enumeration and isolation of spirochetes from salt marsh habitats. *Curr. Microbiol.* **5**:303–306.
 56. Weisburg, W. G., R. P. Hatch, and C. R. Woese. 1986. Eubacterial origin of chlamydiae. *J. Bacteriol.* **167**:570–574.
 57. Weisburg, W. G., Y. Oyaizu, H. Oyaizu, and C. R. Woese. 1985. Natural relationship between bacteroides and flavobacteria. *J. Bacteriol.* **164**:230–236.
 58. Woese, C. R. 1987. Bacterial evolution. *Microbiol. Rev.* **51**:221–271.
 59. Woese, C. R., E. Stackebrandt, T. J. Macke, and G. E. Fox. 1985. A phylogenetic definition of the major eubacterial taxa. *Syst. Appl. Microbiol.* **6**:143–151.
 60. Yasuda, A. H., A. G. Steigerwalt, K. R. Sulzer, A. F. Kaufmann, F. Rogers, and D. J. Brenner. 1987. Deoxyribonucleic acid relatedness between serogroups and serovars in the family *Leptospiraceae* with proposal for seven new *Leptospira* species. *Int. J. Syst. Bacteriol.* **37**:407–415.