JOURNAL OF BACTERIOLOGY, Aug. 1986, p. 570–574
0021-9193/86/080570-05$02.00/0

Vol. 167, No. 2

Eubacterial Origin of Chlamydiae

W. G. WEISBURG,[1] T. P. HATCH,[2] AND C. R. WOESE[1]*

Department of Genetics and Development, University of Illinois, Urbana, Illinois 61801,[1] and Department of Microbiology and Immunology, University of Tennessee Center for the Health Sciences, Memphis, Tennessee 38163[2]

Received 7 February 1986/Accepted 24 April 1986

The sequence of the 16S rRNA gene from *Chlamydia psittaci* was determined. Comparison of this sequence with other 16S rRNA sequences showed the organism to be eubacterial. The organism represents a hitherto unrecognized major eubacterial group. However, this group may be peripherally related to the planctomyces and relatives. Although these two groups seem to have very little in common phenotypically (they have been studied in very different ways), cell walls in both cases contain no peptidoglycan.

Molecular sequencing is changing the way microbiologists look at their science. There now exists a microbial phylogenetic tree, which was nonexistent before the late 1970s (6, 28). The value of knowing the natural relationships among bacteria goes far beyond the formal structure of this tree, however. Phylogenetic categories are predictive; what is known about one organism can generally be applied to its relatives. Medical microbiology in particular should benefit from a phylogenetic perspective, especially in the study of pathogens difficult to cultivate. If such organisms can be related to nonpathogens, about which more is known or can readily be ascertained, study of them will be greatly facilitated. It has recently been shown, for example, that *Rochalimaea quintana*, one of the rickettsiae, is closely related to the group of eubacteria containing the plant-associated genera *Agrobacterium* and *Rhizobium* (25). Our knowledge of both of these latter genera is considerable and should help in understanding the rickettsiae. The present study investigates the phylogeny of the chlamydiae.

Chlamydiae are obligate intracellular parasites of eucaryotic cells that cause a variety of diseases in mammals and birds, e.g., one of the more common forms of human venereal disease (J. Moulder, ASM News **50**:353–362, 1984). The genus *Chlamydia* as now defined contains two species, *C. psittaci* and *C. trachomatis*. The two species are thought to be specifically related to one another; they share a common genus-specific antigen (Moulder, ASM News) and exhibit a low but significant level of DNA-DNA cross-hybridization (11; Moulder, ASM News). A number of chlamydialike organisms, of unknown relationship to the genus *Chlamydia*, have also been observed in various invertebrate hosts, such as clams and hydras (18, 20).

rRNA sequence comparisons provide a useful approach to phylogenetic relationship, especially in those cases in which classical taxonomic criteria are not informative (6). The approach is universally applicable (to self-replicating systems) and is generally superior to those based upon smaller molecules, e.g., cytochrome *c* or ferredoxin (6).

MATERIALS AND METHODS

C. psittaci 6BC was grown as previously described (8). Nucleic acids, RNA, and DNA were extracted by standard procedures (15). The 16S rRNA was isotopically labeled in vitro (14), and it served as a hybridization probe. The 16S rRNA gene, contained in an approximately 10-kilobase fragment produced by restriction endonuclease *Bam*HI, was cloned into bacteriophage lambda L47 (13). The gene was subcloned into the single-stranded phage M13mp8 and M13mp9 system (16) in two pieces, one a 2.0-kilobase *Eco*RI fragment, the other a 2.4-kilobase fragment produced by *Eco*RI and *Hin*dIII. (A site for restriction endonuclease *Eco*RI exists in the 16S rRNA sequence at approximately position 680.) Clones were produced in both orientations for sequencing.

Sequencing methods. The dideoxynucleotide chain termination method (21) was used on templates derived from phage M13 (16). Newly synthesized strands were labeled with [α-^{35}S]dATP (1). Two types of guanine sequencing reactions were routinely used, one normal, the other in which dGTP was replaced by dITP (2′, 3′-dideoxy GTP being used to terminate chain growth [17]). The 17-nucleotide M13 priming site (16), as well as specific priming sites within the rRNA genes (for which primers were synthesized, most by the University of Illinois DNA Synthesis Facility), were used. Specific 15- to 17-nucleotide-long primers, priming in the forward (i.e., same sequence as the rRNA) and reverse directions, have been synthesized for regions of the 16S rRNA molecule that are conserved among eubacteria (26, 28, 29). Those primers used in the present study cover positions 260F, 270R, 520F-R, 790F, 920R, 1100F-R, 1240F-R, and 1400F-R. The numbering refers to the *Escherichia coli* sequence; F is forward, R is reverse direction. An oligonucleotide that primed in the forward direction, about 80 bases upstream from the 5′ end of this particular molecule, was also used. The entire gene was sequenced in both directions except for the 3′-terminal approximately 150 bases (sequenced in the forward direction only). Sequences were aligned by the procedure of Woese et al. (26).

RESULTS AND DISCUSSION

The *C. psittaci* 16S rRNA sequence was aligned with that of *E. coli* (2) (Fig. 1). A sequence homology matrix derived from seven aligned eubacterial sequences, the two in Fig. 1 plus those of *Planctomyces staleyi* (H. Oyaizu and C. R. Woese, unpublished data), *Flavobacterium heparinum* (24), *R. quintana* (25), *Desulfovibrio desulfuricans* (19), *Bacillus subtilis* (7), and *Anacystis nidulans* (23), is shown in Table 1. The sequence of *Methanococcus vannielii* 16S rRNA (9), an

* Corresponding author.

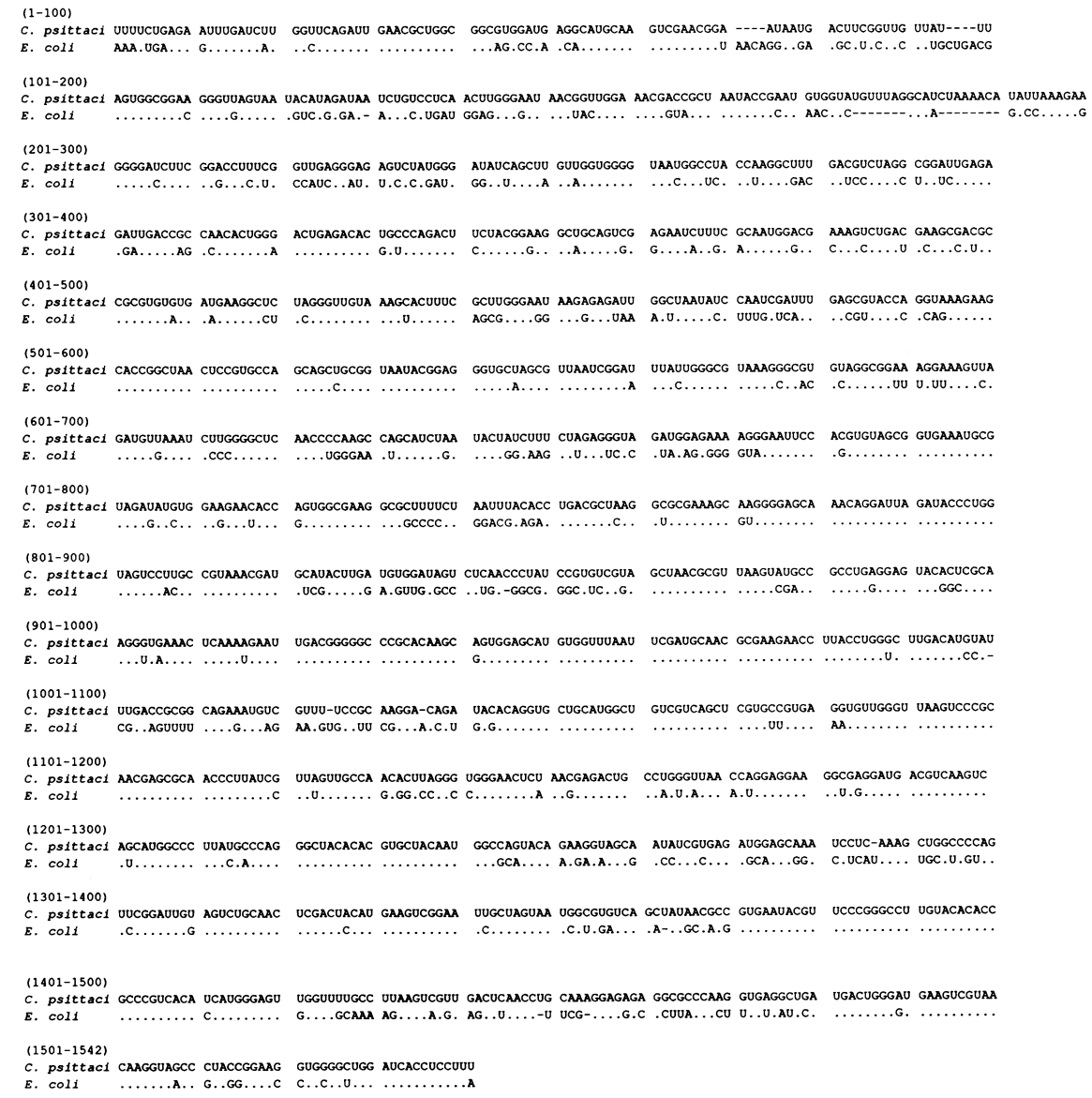

FIG. 1. Sequence of *C. psittaci* 16S rRNA, shown aligned with that of *E. coli*, the standard eubacterial rRNA sequence. *E. coli* numbering is used (2). Dots (in the *E. coli* sequence) mean composition identical to that of the *C. psittaci* sequence. Dashes signify no nucleotide at the given position.

archaebacterium, is included as an outgroup. (This last sequence is typical of the archaebacteria in that any statements based upon it in the present context hold for other individual or consensus archaebacterial sequences as well).

A phylogenetic tree derived with a distance matrix algorithm from the evolutionary distances in Table 1 is shown in Fig. 2 (4).

The chlamydial rRNA sequence is much closer to eubacterial sequences than to the archaebacterial sequence (Table 1). Therefore, the organism is a eubacterium. This is not overly surprising in view of the fact that the organism is known to have a gram-negative type cell envelope; two of the three antigenic domains of the common chlamydial lipopolysaccharide cross-react with typical bacterial cell envelopes (3).

Some eubacterial sequences, including that of *C. psittaci*, are also shown to be more distant than others from the outgroup sequence (Table 1); this implies that their lineages are evolving relatively rapidly (27). Determining exact phylogenetic relationships for rapidly evolving lines is difficult; all methods of analysis now in use tend to place them as deeper branchings in phylogenetic trees than they actually are (5). Therefore, the precise relationship of *C. psittaci* to the other eubacterial lineages cannot be determined at this stage. (Neither the exact branching order nor the exact position of the root in Fig. 2 should be considered certain.) However, what can be said of *C. psittaci* is that the organism has no known close (specific) relatives. Certainly none appear in Table 1 (or in Fig. 2). The organism is not related to the rickettsiae (with which the chlamydiae are often conveniently grouped taxonomically). In fact, the *C. psittaci* sequence shows no specific relationship to any of the more

TABLE 1. Percent similarity and evolutionary distance among various eubacterial sequences[a]

Organism	% Similarity to:								
	C. psittaci	*P. staleyi*	*F. heparinum*	*E. coli*	*R. quintana*	*D. desulfuricans*	*B. subtilis*	*A. nidulans*	*M. vannielii*
C. psittaci		51.7	52.5	56.8	55.4	60.1	58.5	58.7	33.7
P. staleyi	26.0		51.1	54.1	53.2	56.8	55.9	54.2	33.9
F. heparinum	25.4	28.0		56.1	56.0	55.5	56.9	54.8	32.8
E. coli	22.2	24.7	23.8		64.6	66.3	63.9	61.5	33.5
R. quintana	24.2	25.8	24.0	18.5		66.5	65.3	62.6	34.6
D. desulfuricans	20.5	23.5	23.2	15.5	17.7		66.5	63.2	36.3
B. subtilis	21.5	24.1	23.2	17.3	16.7	15.0		66.7	38.6
A. nidulans	21.8	24.8	24.6	19.2	20.6	17.4	15.4		37.7

[a] Upper right section, percent similarity among the sequences. Similarity is calculated on the basis of those positions in the alignment that are represented in, but that are not of constant composition in, all eubacterial sequences (including the outgroup sequence). (Removing positions of constant composition facilitates comparisons for the reader; it has no effect on rank order of values but accentuates differences among them.) Lower left section, evolutionary distances among the species, i.e., the percent difference among the sequences corrected for multiple changes by the method of Jukes and Cantor (10). The calculation in this case uses only those positions in the alignment in which four or more of the eubacterial sequences have the same composition. (Restricting the alignment in this way selectively eliminates some of the more rapidly varying positions, which tend to distort the analysis.) The additional sequences included here but not shown in Fig. 1 are *D. desulfuricans* (19), *R. quintana* (25), *P. staleyi* (H. Oyaizu and C. R. Woese, unpublished data), *F. heparinum* (24), *B. subtilis* (7), *A. nidulans* (23), and the archaebacterium *M. vannielii* (9).

than 400 partial eubacterial 16S rRNA sequences (i.e., oligonucleotide catalogs) determined thus far (6, 28; unpublished analysis).

Nevertheless, when the data are analyzed not in terms of evolutionary distances but in terms of derived characters, i.e., positions that have a composition unique to and characteristic of particular groupings, a remote but specific relationship between the chlamydiae and the planctomyces (22) is suggested. For example, scoring the alignment used to construct Table 1 for positions in which the *C. psittaci* sequence has a common composition with some other sequence, but only with that sequence, shows about 25 such positions with the planctomyces sequence, a number that is significantly higher than the chlamydial sequence shows with any other eubacterial sequence.

The possible relationship between *C. psittaci* and the planctomyces is most convincingly seen in terms of the higher order structure of the molecule. That portion of the 16S rRNA structure in which the resemblance between the two sequences is most pronounced is shown in Fig. 3; those features strongly characteristic of the putative relationship are indicated by numbered arrows. The salient points are as follows. (The statements here are based upon a collection of about 15 eubacterial 16S sequences, some unpublished, and

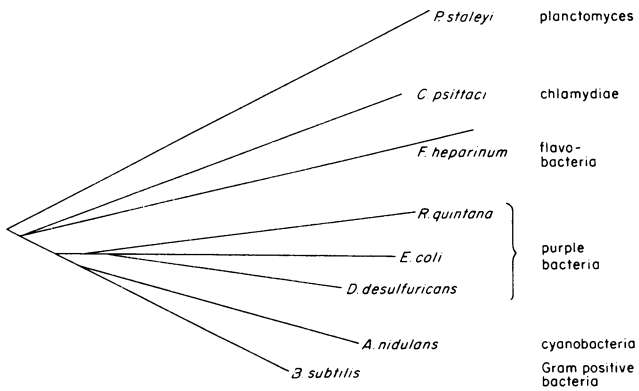

FIG. 2. Phylogenetic tree constructed from the evolutionary distances shown in Table 1 with the treeing algorithm of DeSoete (4). Lengths on the figure are drawn to scale (4) but are in arbitrary units. The root has been imposed by using an archaebacterial sequence.

over 400 oligonucleotide catalogs [28]). (i) Except for the G-A seen in *C. psittaci* and *P. staleyi*, eubacterial sequences contain pyrimidines at both positions 47 and 48, a generalization drawn from more than 400 examples (28). Moreover, these two are the only sequences in which it is not possible to form a canonical base pair between positions 47 and 394. (ii) The guanine at position 52 is very rare among eubacteria; as judged by the catalogs, the position shows a pyrimidine or (rarely) an adenine in every member of every eubacterial division, except for the *Deinococcus* group, the species of which also show a guanine here (28). The neighboring guanine at position 53 is seen elsewhere among eubacteria only in the chloroflexus group (28); however, no other sequences (including three from the latter group) show the G-U pair involving positions 53 and 358. (iii) The indicated adenine residue (position 361) is found elsewhere among eubacteria only in the spirochetes (unpublished analysis). (iv) The uracil residue here (position 353) is not found in any other eubacterial sequences. (v) The failure to form a canonical base pair between positions 322 and 331 is unique to the *C. psittaci* and *P. staleyi* sequences. (vi) This G-C pair (positions 290 and 310) is found elsewhere among eubacterial sequences only in the bacteroides-flavobacteria group. The residue at position 310 is detectable in eubacterial 16S rRNA catalogs, and, except for these cases (in which it is cytosine), it is always a guanine (28). (vii) A U-G pair involving these positions (242 and 284) is confined to these two (eubacterial) sequences. (viii) The adenine residue at position 110 has been encountered only in the two sequences of Fig. 3B.

The common unique features seen in the *C. psittaci* and planctomyces sequences do not seem to represent just one or a few interdependent (higher-order structural) changes; for the most part they probably represent independent occurrences. If so, they are too many and otherwise too rare not to indicate a specific relationship between the two sequences. However, this relationship should not be considered reliably established until it emerges from a more extensive collection of sequences covering this part of the eubacterial tree. The only phenotypic characteristic the two groups are known to have in common is a cell wall that lacks peptidoglycan, an otherwise universal eubacterial characteristic (12; Moulder, ASM News).

If a specific relationship between the chlamydiae and the planctomyces exists, the two are sufficiently distinct that *C. psittaci* should still be given the status of a separate eubacte-

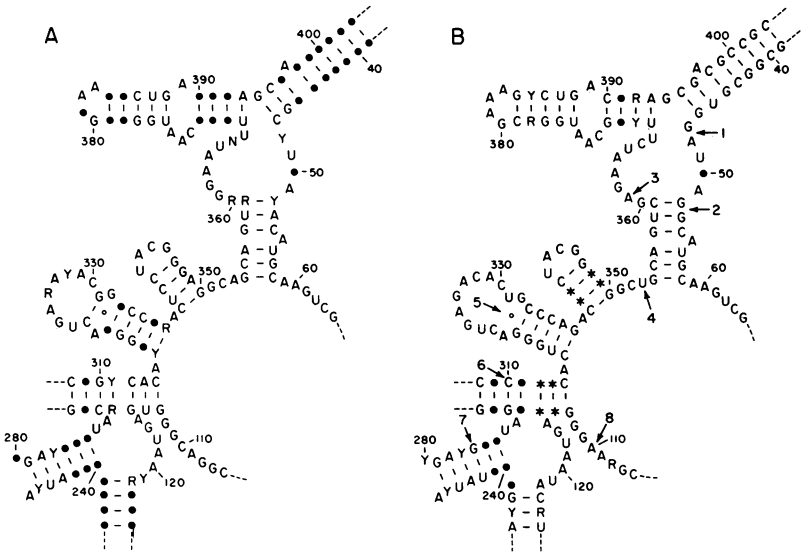

FIG. 3. Comparison of higher-order structure for a region of the 16S rRNA from a consensus of all eubacteria (26, 28) (A), except for *P. staleyi* and *C. psittaci* (B). The regions indicated by numbered arrows are those specifically discussed in the text. R, Purine; Y, pyrimidine; ●, no consensus at the position; *, position in which one or both sequences in panel B differ from the eubacterial consensus and from each other as well.

rial division. The fact that the *C. psittaci* sequence is so unique should make the detection of the organism (and its relatives) with probes based upon rRNA a relatively easy matter.

ADDENDUM IN PROOF

We have recently determined greater than 98% of the *Chlamydia trachomatis* 16S rRNA gene sequence. There was less than 5% difference between this sequence and its counterpart from *C. psittaci*.

ACKNOWLEDGMENTS

We thank R. R. Gutell for his help in preparing Fig. 1.

These studies were supported by grant NSG 7044 from the National Aeronautics and Space Administration and by Public Health Service grants AI 19750 and AI 22910 from the National Institutes of Health.

LITERATURE CITED

1. **Biggin, M. D., T. J. Gibson, and G. F. Hong.** 1983. Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. USA **80**:3963–3965.
2. **Brosius, J., J. L. Palmer, J. P. Kennedy, and H. F. Noller.** 1978. Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*. Proc. Natl. Acad. Sci. USA **75**: 4801–4805.
3. **Caldwell, H. D., and P. J. Hitchcock.** 1984. Monoclonal antibody against a genus-specific antigen of *Chlamydia* species: location of the epitope on chlamydial lipopolysaccharide. Infect. Immun. **44**:306–314.
4. **DeSoete, G.** 1983. A least squares algorithm for fitting additive trees to proximity data. Psychometrika **48**:621–626.
5. **Felsenstein, J.** 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. **27**:401–410.
6. **Fox, G. E., E. Stackebrandt, R. B. Hespell, J. Gibson, J. Maniloff, T. A. Dyer, R. S. Wolfe, W. E. Balch, R. Tanner, L. Magrum, L. B. Zablen, R. Blakemore, R. Gupta, L. Bonen, B. J. Lewis, D. A. Stahl, K. R. Luehrsen, K. N. Chen, and C. R. Woese.** 1980. The phylogeny of prokaryotes. Science **209**: 457–463.
7. **Green, C. J., G. C. Stewart, M. A. Hollis, B. S. Vold, and K. F. Bott.** 1985. Nucleotide sequence of *Bacillus subtilis* ribosomal RNA operon, *rrnB*. Gene **37**:261–266.
8. **Hatch, T. P.** 1975. Competition between *Chlamydia psittaci* and L cells for host isoleucine pools: a limiting factor in chlamydial multiplication. Infect. Immun. **12**:211–220.
9. **Jarsch, M., and A. Böck.** 1985. Sequence of the 16S ribosomal RNA gene from *Methanococcus vannielii*: evolutionary implications. Syst. Appl. Microbiol. **6**:54–59.
10. **Jukes, T. H., and C. R. Cantor.** 1969. Evolution of protein molecules, p. 21–132. *In* H. N. Munro (ed.), Mammalian protein metabolism. Academic Press, Inc., New York.
11. **Kingsbury, D. T., and E. Weiss.** 1968. Lack of deoxyribonucleic acid homology between species of the genus *Chlamydia*. J. Bacteriol. **96**:1421–1423.
12. **König, E., H. Schlesner, and P. Hirsch.** 1984. Cell wall studies on budding bacteria of the *Planctomyces/Pasteuria* group and on a *Prosthecomicrobium* sp. Arch. Microbiol. **138**:200–205.
13. **Loenen, W. A. M., and W. J. Brammar.** 1980. A bacteriophage lambda vector for cloning large DNA fragments made with several restriction enzymes. Gene **20**:249–259.
14. **Maizels, N.** 1976. Dictyostelium 17S, 25S, and 5S rDNAs lie within a 38,000 base pair repeated unit. Cell **9**;431–438.
15. **Marmur, J.** 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. J. Mol. Biol. **3**:208–218.
16. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
17. **Mills, D. R., and F. R. Kramer.** 1979. Structure-independent nucleotide sequence analysis. Proc. Natl. Acad. Sci. USA **76**:2232–2235.
18. **O'Brien, T. L., and R. MacLeod.** 1983. Description of chlamydia-like micro-organisms in *Hydra viridis* and initial characterization of the relation between micro-organisms and host. Cytobios **36**:141–151.
19. **Oyaizu, H., and C. R. Woese.** 1985. Phylogenetic relationships among the sulfate respiring bacteria, myxobacteria and purple bacteria. Syst. Appl. Microbiol. **6**:257–263.
20. **Page, L. A., and R. C. Cutlip.** 1982. Morphological and immunological confirmation of the presence of chlamydiae in the gut tissue of the Chesapeake bay clam, *Mercenaria mercenaria*. Curr. Microbiol. **7**:297–300.
21. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci.

USA **74**:5463–5467.

22. **Stackebrandt, E., W. Ludwig, W. Schubert, F. Klink, H. Schlesner, T. Roggentin, and P. Hirsch.** 1984. Molecular genetic evidence for early evolutionary origin of budding peptidoglycan-less eubacteria. Nature (London) **307**:735–737.
23. **Tomioka, N., and M. Sugiura.** 1983. The complete nucleotide sequence of a 16S ribosomal RNA gene from a blue-green alga, *Anacystis nidulans*. Mol. Gen. Genet. **191**:46–50.
24. **Weisburg, W. G., Y. Oyaizu, H. Oyaizu, and C. R. Woese.** 1985. Natural relationship between bacteroides and flavobacteria. J. Bacteriol. **164**:230–236.
25. **Weisburg, W. G., C. R. Woese, M. E. Dobson, and E. Weiss.** 1985. A common origin of rickettsiae and certain plant pathogens. Science **230**:556–558.
26. **Woese, C. R., R. Gutell, R. Gupta, and H. F. Noller.** 1983. Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. Microbiol. Rev. **47**:621–669.
27. **Woese, C. R., E. Stackebrandt, and W. Ludwig.** 1985. The relationship of tempo and mode in bacterial evolution. J. Mol. Evol. **21**:305–316.
28. **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.
29. **Yang, D., Y. Oyaizu, H. Oyaizu, G. J. Olsen, and C. R. Woese.** 1985. Mitochondrial origins. Proc. Natl. Acad. Sci. USA **82**:4443–4447.