Phylogenetic Relationship of Cowdria ruminantium, Agent of Heartwater, to Anaplasma marginale and Other Members of the Order Rickettsiales Determined on the Basis of 16S rRNA Sequence†

JOHN B. DAME,* SUMAN M. MAHAN, AND CHARLES A. YOWELL

Department of Infectious Diseases, University of Florida, Gainesville, Florida 32611-0633

The phylogenetic relationship between Cowdria ruminantium and representative members of the orders Rickettsiales and Chlamydiales has been examined on the basis of the sequence of the 16S rRNA. Phylogeny reconstruction by using both parsimony and distance methods supports the conclusion that C. ruminantium is closely related to the Rickettsiales and in particular to the family Anaplasmataceae. A signature of nine base substitutions delineated the family Anaplasmataceae with C. ruminantium and differentiated these two species from the 45 other members of the alpha group of Proteobacteria examined, and five of these base substitutions were unique among all members of the class Proteobacteria examined to date.

Heartwater, or cowdrosis, is a tick-borne disease of domestic and wild ruminants caused by the intracellular, rickettsialike pathogen, Cowdria ruminantium Cowdry 1925 (15). The disease is important in tropical Africa, where it is a major barrier to improved livestock production (13), and its presence on islands in the Caribbean Sea makes it a threat to domestic and wild ruminants caused by the intracellular, rickettsialike pathogen,

MATERIALS AND METHODS

Preparation of genomic DNA and amplified 16S rRNA gene. Genomic DNA was prepared (17) from the Crystal Springs isolate cultured in bovine endothelial cells derived from the aorta (28). C. ruminantium elementary bodies were collected from culture supernatants for extraction of DNA by centrifugation for 10 min at 2,000 x g to remove the remnants of bovine cells lysed by the infection followed by centrifugation at 4°C for 30 min at 50,000 x g to pellet the elementary bodies. DNA prepared from C. ruminantium purified in this fashion contains some bovine DNA, but 90 ng of DNA from this preparation was sufficient to amplify the 16S rRNA gene(s) in vitro via PCR (9) by using a primer set described by Wilson et al. (25). PCR was performed in a volume of 100 μl containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl2, 0.2 mM dideoxynucleoside triphosphates, 1 μM primer PC5, 1 μM primer POMOD, and 2.5 U of Taq polymerase. Thermal cycling was performed in a Coy Model 60 cycler in which the target DNA was denatured by incubation at 92°C for 1 min followed by 35 cycles of denaturation (92°C for 1 min), primer annealing (55°C for 3 min), and primer extension (72°C for 3 min). At the end of cycling, the reaction mixture was held at 72°C for 10 min and cooled to 4°C. A fragment of 1,463 bp, which ran as a single ethidium bromide-staining band on a 0.8% agarose gel, was amplified. The extreme 5’ and 3’ ends of the 16S rRNA gene and adjacent regions were amplified by inverse PCR (10) after the genomic DNA was cut with XbaI and then ligated. The PCR primers at positions 64 to 50 and 297 to 283, 257 to 245, 1337 to 1351, 1083 to 1097, and 1337 to 1351 (sequences shown in Table 1) were used under the conditions described above, and amplification produced a 1.8-kb fragment.

DNA sequence analysis and cloning. The 1,463-bp amplified DNA fragment was sequenced directly in both strands by the dyeodeoxy chain termination method (18) by using 32P-end-labeled oligonucleotide sequencing primers labeled with T4 polynucleotide kinase (17). Suitable primers for sequencing the 16S rRNA from members of the Rickettsiaceae were identified by aligning the sequences of 16S rRNAs of members of the Rickettsiaceae available in the GenBank data base with that of Escherichia coli (from the gamma group of the Proteobacteria) by using the multiple-sequence alignment program, CLUSTAL (Intelligenetics, Mountain View, Calif.). Regions of the 16S rRNA sequence which were sufficiently conserved between members of the Rickettsiaceae and E. coli were identified at positions 50 to 64, 297 to 311, 558 to 572, 787 to 806, 942 to 956, 1083 to 1097, and 1337 to 1351 (E. coli numbering). The corresponding and complementary oligonucleotide primers in Table 1 were prepared synthetically (14). In some cases, these primers were mixtures at one or two positions. The amplified DNA was ligated into the Smal site of pBluescript SK+ (Stratagene) after being prepared for cloning by polishing the ends with the Klenow fragment of DNA polymerase I (17). Regions of the sequence which were not determined with certainty by

* Corresponding author.
† This paper is University of Florida Agricultural Experiment Stations Journal Series no. R-02162.
TABLE 1. PCR and sequencing primers

<table>
<thead>
<tr>
<th>Primer sequence</th>
<th>Position in 16S RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>in E. coli</td>
</tr>
<tr>
<td>Forward</td>
<td></td>
</tr>
<tr>
<td>5’AGAGTTGATGCTCTG3’</td>
<td>8-22</td>
</tr>
<tr>
<td>5’AAAGATACGATGCTG3’</td>
<td>50-64</td>
</tr>
<tr>
<td>5’GGAGGGA(C/T)GATCGAG3’</td>
<td>297-311</td>
</tr>
<tr>
<td>5’GAAT(C/T)ACTGGCCGAT3’</td>
<td>558-572</td>
</tr>
<tr>
<td>5’ATTAGATACGCTCGTGTGGCA3’</td>
<td>787-806</td>
</tr>
<tr>
<td>5’TGGAGA(C/T)ATG(C/T)GTGTT3’</td>
<td>942-956</td>
</tr>
<tr>
<td>5’TGGTTGCTTAATGCG3’</td>
<td>1083-1097</td>
</tr>
<tr>
<td>5’GAATCCTGAATATA3’</td>
<td>1337-1351</td>
</tr>
<tr>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>5’TACCTTGTTGACTC3’</td>
<td>1507-1492</td>
</tr>
<tr>
<td>5’ATTAGAATGACCTG3’</td>
<td>1351-1337</td>
</tr>
<tr>
<td>5’GGACTTAACCCAAC3’</td>
<td>1097-1083</td>
</tr>
<tr>
<td>5’ACC(C/A)GT(A/G)CTGAG3’</td>
<td>956-942</td>
</tr>
<tr>
<td>5’GGACATA(C/T/A)AGGGTATCTA3’</td>
<td>806-876</td>
</tr>
<tr>
<td>5’TACCGCGCGAT3’</td>
<td>572-558</td>
</tr>
<tr>
<td>5’GGACTAC(C/G)TGG3’</td>
<td>834-839</td>
</tr>
<tr>
<td>5’GACCTGGATGTGTT3’</td>
<td>64-50</td>
</tr>
</tbody>
</table>

a PO100; b P100; c PC5 (25).

direct sequencing of the PCR product were confirmed by
sequencing the cloned fragment. The 1.8-kb amplified frag-
mament was cloned as a blunt fragment in the EcoRV cloning
site of pBluescript SK+, and the ends were sequenced by
using T3 and T7 sequencing primers.

Phylogenetic analyses. Comparison of the C. ruminantium
sequence with the published sequences of Anaplasma mar-
ginale (M60313), Coxiella burnetii (M21291), Chlamydia
psittaci (M13769), Ehrlichia risticii (M21290), E. coli
(J01695), Rochalimaea quintana (M11927), Rickettsia
prowazekii (M21789), Rickettsia rickettsii (M21293), and
Wolbachia persica (M21292), by using Planctomycyes staleyi
(M34126) as the outgroup, was performed both by parsimony
analysis with PAUP (20) and by distance analysis to calcu-
late the Kimura two-parameter distance (6) by using the
DNADIST program (4). Then, the neighbor-joining method
(16) and tree construction were performed by using the
NEIGHBOR and DRAWTGRAM programs, respectively, in
the PHYLIP program package (4). Transversions were
weighted 2:1 versus transitions in parsimony analysis. All
sequences were aligned by using the CLUSTAL multiple-
sequence alignment program (available in the PC/Genie
package available from IntelliGenetics) and examined by
eye for regions with questionable homology.

Nucleotide sequence accession number. The nucleotide
sequence described in this paper has been submitted to the
EMBL data bank and assigned accession no. X61659.

RESULTS AND DISCUSSION

The sequence of the C. ruminantium 16S rRNA sequence
is shown in Fig. 1. To compare C. ruminantium to A. mar-
ginale, the 16S sequence of A. marginale was also ampli-
fied and sequenced as described above. During analy-
ysis of our data, the sequence of the 16S rRNA of A.
marginale (M60313) was published (22). The sequence
as reported (22) was identical to the sequence obtained in
our laboratory for the Florida isolate (7) for the region compa-
rable to bases 23 through 1491 (E. coli 16S rRNA num-
bering).

Regions comparable to bases 1 to 22, 69 to 102, 182 to 210,
454 to 478, 834 to 856, 998 to 1044, 1243 to 1294, and 1449 to
1542 of the E. coli sequence for all taxa were excluded from
parsimony analysis. For these regions, it was not possible to
confidently determine whether the bases aligned in these
regions represented homologous nucleotide positions. The
sequences of the regions at the ends of the gene (bases 1 to
22 and 1492 to 1542) were excluded, since the C. ruminanti-
tum gene sequence was available from only a single strand
in these regions. The 25 bases (454 to 478) were present in C.
burnettii, W. persica, E. coli, C. psittaci, and P. staleyi
but were absent from the other taxa. A total of 1,221 character

FIG. 1. Sequence of the structural rRNA encoding the C. ruminantium 16S rRNA (413 A’s, 298 C’s, 434 G’s, and 361 T’s).
positions were examined, and 318 were phylogenetically informative by parsimony analysis.

The most parsimonious tree was determined for the 11-taxon data set by using the branch and bound algorithm and by weighting transversions 2:1 versus transitions. This tree with a length of 1,619 steps and a consistency index of 0.493 is shown in Fig. 2A. Forty-four trees were within 31 additional steps of the most parsimonious tree. In all 44 trees, C. ruminantium was grouped with A. marginale and C. psittaci branching was placed at the deepest position on the tree at a large phylogenetic distance from C. ruminantium and the members of the Rickettsiales. Differences in the positioning of the other Proteobacteria relative to each other occurred among these longer trees. For example, the next most parsimonious tree, which was three steps longer, differed only by placement of the nodes separating E. coli, W. persica, and C. burnetii. The exact phylogenetic relationships of E. risticii and R. quintana to the other species in the alpha group of Proteobacteria and E. coli, C. burnetii, and W. persica among the gamma group of Proteobacteria are less certain (see below). The robustness of the essential features of the most parsimonious tree was examined by data resampling (bootstrap) analysis of 100 repetitions (20). This method defined the 95% consensus tree shown in Fig. 2B. The clustering of C. ruminantium with A. marginale and of R. rickettsii with R. prowazekii, together with the deep branching of C. psittaci, was a fully supported feature of the tree. Thus, by parsimony analysis, C. ruminantium and A. marginale are monophyletic within the alpha group of Proteobacteria (R. rickettsii, R. prowazekii, R. quintana, E. risticii, A. marginale, and C. ruminantium), and this group forms a clade separate from the other three major bacterial groups, the gamma group of Proteobacteria (E. coli, C. burnetii, and W. persica), the Chlamydiaceae (C. psittaci), and the outgroup (P. staleyi), represented in the analysis.

The sequences, when examined by distance methods, imply the same groupings as those obtained by parsimony analysis. In the group of 11 taxa examined, 1,221 base positions were aligned as above, and a distance matrix was prepared by using the Kimura two-parameter (6) model of nucleotide substitution (Table 2) with the DNADIST program (4). When the sequences were plotted with P. staleyi as the outgroup via the neighbor-joining method (NEIGHBOR in PHYLIP [4]) (Fig. 3), C. ruminantium was clustered with A. marginale among the alpha group of Proteobacteria (R.

![FIG. 2. (A) Most parsimonious tree relating C. ruminantium to 10 other bacterial taxa; (B) 95% consensus tree from parsimony analysis of 100 bootstrap replications from an 11-taxon data set used for panel A.](image)

### Table 2. Evolutionary distances adjusted by using Kimura two-parameter correction

<table>
<thead>
<tr>
<th>Organism</th>
<th>A. marginale</th>
<th>C. burnetii</th>
<th>C. psittaci</th>
<th>C. ruminantium</th>
<th>E. coli</th>
<th>E. risticii</th>
<th>P. staleyi</th>
<th>R. prowazekii</th>
<th>R. rickettsii</th>
<th>R. quintana</th>
<th>W. persica</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. marginale</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. burnetii</td>
<td>0.1760</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. psittaci</td>
<td>0.2737</td>
<td>0.2553</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. ruminantium</td>
<td>0.0517</td>
<td>0.1875</td>
<td>0.2896</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td>0.1910</td>
<td>0.1349</td>
<td>0.2567</td>
<td>0.2060</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. risticii</td>
<td>0.1323</td>
<td>0.1840</td>
<td>0.3080</td>
<td>0.1450</td>
<td>0.2150</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. staleyi</td>
<td>0.2786</td>
<td>0.2520</td>
<td>0.2998</td>
<td>0.2868</td>
<td>0.2882</td>
<td>0.2987</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. prowazekii</td>
<td>0.1363</td>
<td>0.1737</td>
<td>0.2813</td>
<td>0.1373</td>
<td>0.2082</td>
<td>0.1488</td>
<td>0.2930</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. rickettsii</td>
<td>0.1342</td>
<td>0.1696</td>
<td>0.2838</td>
<td>0.1352</td>
<td>0.2098</td>
<td>0.1500</td>
<td>0.2963</td>
<td>0.0041</td>
<td>0.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. quintana</td>
<td>0.1433</td>
<td>0.1612</td>
<td>0.2706</td>
<td>0.1535</td>
<td>0.1860</td>
<td>0.1745</td>
<td>0.2799</td>
<td>0.1384</td>
<td>0.1373</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>W. persica</td>
<td>0.1914</td>
<td>0.1409</td>
<td>0.2597</td>
<td>0.1904</td>
<td>0.1524</td>
<td>0.2140</td>
<td>0.2724</td>
<td>0.1957</td>
<td>0.1959</td>
<td>0.1705</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
**TABLE 3. Sequence signature linking C. ruminantium and A. marginale**

<table>
<thead>
<tr>
<th>Position</th>
<th>Base common to C. ruminantium and A. marginale</th>
<th>Base at position indicated for other Proteobacteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
<td>A</td>
<td>G or U, U, A, G, or U</td>
</tr>
<tr>
<td>396</td>
<td>U</td>
<td>A, C, or G, A, or C</td>
</tr>
<tr>
<td>508</td>
<td>U</td>
<td>A, C, or G, U</td>
</tr>
<tr>
<td>611</td>
<td>A</td>
<td>C, G, or U</td>
</tr>
<tr>
<td>988</td>
<td>U</td>
<td>A, C, or G</td>
</tr>
<tr>
<td>1192</td>
<td>U</td>
<td>C</td>
</tr>
<tr>
<td>1205</td>
<td>C</td>
<td>C, U</td>
</tr>
<tr>
<td>1427</td>
<td>A</td>
<td>C, G, or U</td>
</tr>
<tr>
<td>1510</td>
<td>U</td>
<td>C, C</td>
</tr>
</tbody>
</table>

*Position (E. coli numbering) at which C. ruminantium and A. marginale have a common sequence which differs from those of all other members of the alpha group of Proteobacteria currently in the RNA Database Project (45 species). Sequence differences from the beta and gamma groups of Proteobacteria reflect examination of 24 and 36 species, respectively.*

Parsimony and distance analyses confirm the current taxonomic placement of *C. ruminantium* among the *Rickettsiales* (15), rather than among the *Chlamydiales* as had been suggested previously (8, 19, 21). Previously, phylogenetic placement of *C. ruminantium* was based on qualitative phenotypic characters such as transmission by tick vectors, development in intracellular vacuoles, and similar host cell-type preference, etc. Conflicting conclusions have been drawn from analysis of different sets of these characters (19), since such characters are not necessarily good phylogenetic characters because of the potential for convergent evolution. Sequence data of the 16S rRNA gene refute the suggestion that *C. ruminantium* may be related phylogenetically to the *Chlamydiales* and fully support the close phylogenetic relationship between *C. ruminantium* and the *Rickettsiales*. Further, our parsimony and distance analyses of the 16S rRNA data confirm the conclusion of Weisburg et al. (24) that the *Chlamydia* form a eubacterial branch distinct from the *Rickettsiales*.

Our data strongly indicate a closer phylogenetic relationship of the genus *Cowdria* to the *Anaplasmataceae* rather than to the *Ehrlichiae* (family *Rickettsiaceae*) (15) within the rickettsial sublineage in the *Proteobacteria* (11, 23, 26, 27). The phylogenetic positions of other species sharing the property of intimate association with eukaryotic cells, namely, *Brucella abortus*, *Bartonella bacilliformis*, and *cat scratch disease bacillus*, were recently evaluated from 16S rRNA sequence data (2, 3, 11). These organisms grouped among the *Proteobacteria* but within the alpha-2 subgroup (2, 3, 11), represented by *R. quintana* in our studies (Fig. 2 and 3), which is distinct from the rickettsial sublineage (26a). Proper placement of *C. ruminantium* among the *Anaplasmataceae*, or determining whether it should form a tribe of its own, will require measuring the relatedness of *Anaplasma* to other genera in the *Anaplasmataceae*, including *Aegyptianella*, *Eperythrozoon*, and *Haemobartonella*. Sequences of the 16S rRNAs from species representative of these genera are not currently available, however.

The obligate, exclusive parasitism of erythrocytes among the *Anaplasmataceae* is currently a primary character linking this family. *C. ruminantium* does not share this characteristic, since it does not infect erythrocytes. Rather, it infects vascular endothelial cells, macrophages, and neutrophils. A study of the molecular phylogeny of the genera...
presently placed in the *Anaplasmataceae* would address whether the adaptation to obligate parasitism of the erythrocytes and/or plasma of vertebrates is a primitive character or whether it is a convergent life-style which has arisen independently multiple times among this group of organisms. The more distant relationship of *Anaplasma* to the other members of the *Anaplasmataceae* (rickettsial sublineage) suggests that the latter is the case. On the other hand, *C. ruminantium*, like all species of *Anaplasma*, is parasitic in ruminants exclusively. This raises the possibility that *C. ruminantium* and the four *Anaplasma* species described have arisen as descendants of a common ancestral parasite of ruminants.

ACKNOWLEDGMENTS

We thank C. Yunker and A. F. Barbet for providing *C. ruminantium* from culture for the preparation of DNA and C. J. Nairn and M. Miyamoto for comments on the manuscript. This research was supported by USAID Cooperative Agreement AFR-0435-A-00-9084-00 and USDA CBAG grant 91-34135-6178.

REFERENCES

26a.Woese, C. R. Personal communication.