JOURNAL OF BACTERIOLOGY, Jan. 1991, p. 697–703
0021-9193/91/020697-07$02.00/0

Vol. 173, No. 2

16S Ribosomal DNA Amplification for Phylogenetic Study

WILLIAM G. WEISBURG,* SUSAN M. BARNS, DALE A. PELLETIER, AND DAVID J. LANE
GENE-TRAK Systems, 31 New York Avenue, Framingham, Massachusetts 01701

Received 16 April 1990/Accepted 7 November 1990

A set of oligonucleotide primers capable of initiating enzymatic amplification (polymerase chain reaction) on a phylogenetically and taxonomically wide range of bacteria is described along with methods for their use and examples. One pair of primers is capable of amplifying nearly full-length 16S ribosomal DNA (rDNA) from many bacterial genera; the additional primers are useful for various exceptional sequences. Methods for purification of amplified material, direct sequencing, cloning, sequencing, and transcription are outlined. An obligate intracellular parasite of bovine erythrocytes, *Anaplasma marginale*, is used as an example; its 16S rDNA was amplified, cloned, sequenced, and phylogenetically placed. Anaplasmas are related to the genera *Rickettsia* and *Ehrlichia*. In addition, 16S rDNAs from several species were readily amplified from material found in lyophilized ampoules from the American Type Culture Collection. By use of this method, the phylogenetic study of extremely fastidious or highly pathogenic bacterial species can be carried out without the need to culture them. In theory, any gene segment for which polymerase chain reaction primer design is possible can be derived from a readily obtainable lyophilized bacterial culture.

The comparison of rRNA sequences is a powerful tool for deducing phylogenetic and evolutionary relationships among bacteria, archaebacteria, and eucaryotic organisms. These sequences have been derived previously by methods including oligonucleotide cataloging (6), sequencing of clones, direct sequencing of RNA by using reverse transcriptase (11), and sequencing of material amplified by polymerase chain reaction (PCR) (3, 5, 15). The present study expands on the use of DNA amplification technology for the study of rRNA sequences within the eubacteria. Several primers are described, and novel methods for their use are presented.

Sogin and coworkers (15) described a primer pair for the enzymatic amplification of eucaryotic small-subunit rRNA gene (rDNA) sequences. Boettger (3) and Edwards and coworkers (5) examined the sequence of *Mycobacterium bovis* 16S rRNA PCR products without cloning. They also examined a limited set of organisms with their PCR primers: four gamma purple bacteria and two gram-positive species with high G+C content were experimentally amplified.

The present study describes primers for the amplification of most eubacterial 16S rRNAs, details preparation of samples, and presents our current sequencing procedures.

MATERIALS AND METHODS

Nucleic acid preparation. Genomic DNAs were prepared by using standard methods (13, 14). Preparation of DNA from lyophilized ampoules consisted of suspending the lyophilized cells in 0.2 ml of 10 mM Tris hydrochloride (pH 8.3)–2.5 mM $MgCl_2$–50 mM KCl. This was added to approximately 0.1 ml of 0.1-mm-diameter acid-washed glass beads, 10 μl of 20% sodium dodecyl sulfate, and approximately 200 μl of phenol, saturated with the above buffer. This sample was shaken in a 500-μl microcentrifuge tube, seated within a 2-ml tube, in a Mini-Bead Beater (Biospec Products) for 2 min. After phase separation (10 min at approximately 15,000 × *g*), the aqueous phase was ethanol precipitated. The pellet was suspended in 20 to 50 μl of TE buffer (10 mM Tris hydrochloride [pH 8.0], 1 mM EDTA), and between 2 and 5% of this resuspended DNA was put into the PCR amplification.

PCR amplification and purification of product. Approximately 1 to 3 μg of genomic DNA was amplified in a 100-μl reaction by using the Geneamp kit (U.S. Biochemicals, Cleveland, Ohio; presently, these kits are only available from Perkin-Elmer Cetus, Norwalk, Conn.). When the lyophilized ampoule DNA was amplified, 1 μl was routinely used. Conditions consisted of 25 to 35 cycles of 95°C (2 min), 42°C (30 s), and 72°C (4 min), plus one additional cycle with a final 20-min chain elongation. The temperature and salt conditions were not optimized for low inputs of template DNA. Between 25 and 35 cycles, the number of cycles did not seem to appreciably affect yield of product. All amplifications were performed in a Perkin-Elmer temperature controller, although a cursory examination of other instruments indicated that they may all perform adequately under these reaction conditions (data not shown). The amplification products were purified on Centricon 100 columns (Amicon), by following the specifications of the manufacturer, followed by ethanol precipitation.

Cloning. Cloning of PCR products was done by using standard methods (13). Amplified products were partially purified with spin columns (see above), cleaved at the appropriate restriction endonuclease sites within the linkers (see Table 1 and Fig. 1) and ligated into correspondingly restricted pGEM-3 or pGEM-4 vectors (Promega Corp., Madison, Wis.). We were able to successfully clone rDNAs that had been either agarose gel purified or not gel purified. Plasmid preparation was performed by using the alkaline lysis method (2, 13) followed by phenol extraction. For cloning of PCR products into plasmid vectors, approximately 20 μg of plasmid DNA restriction endonuclease cut with a combination of *Sal*I and *Bam*HI (rarely found internally within 16S) was dephosphorylated with calf intestinal alkaline phosphatase, gel purified, and stored frozen. This vector preparation should be enough for 20 or more cloning experiments. *Anaplasma marginale* rDNA was cloned into the *Sal*I-*Bam*HI site in this manner.

Sequencing methods. The sequencing method of plasmid-cloned material used the T7 polymerase (Sequenase; U. S. Biochemicals) kit, and standard procedures were followed

* Corresponding author.

TABLE 1. Summary of primers for the PCR amplification of eubacterial 16S rDNA[a]

Primer	Sequence (5′ to 3′)	Designed for:
fD1	ccgaattcgtcgacaacAGAGTTTGATCCTGGCTCAG	Most eubacteria
fD2	ccgaattcgtcgacaacAGAGTTTGATCATGGCTCAG	Enterics and relatives
fD3	ccgaattcgtcgacaacAGAGTTTGATCCTGGCTTAG	*Borrelia* spirochetes
fD4	ccgaattcgtcgacaacAGAATTTGATCTTGGTTCAG	Chlamydiae
rD1	cccgggatccaagcttAAGGAGGTGATCCAGCC	Many eubacteria
rP1	cccgggatccaagcttACGGTTACCTTGTTACGACTT	Enterics (and most eubacteria)
rP2	cccgggatccaagcttACGGCTACCTTGTTACGACTT	Most eubacteria
rP3	cccgggatccaagcttACGGATACCTTGTTACGACTT	Fusobacteria (and most eubacteria)

[a] Primer abbreviations: f, forward; r, reverse; D, distal; P, proximal. All primer sequences are presented in 5′ to 3′ orientation. Linker sequences containing restriction sites for cloning are designated in lowercase letters. The "f" series of linkers all contain *Eco*RI and *Sal*I sites, and the "r" series all contain *Hin*dIII, *Bam*HI, and *Xma*I recognition sequences. Reverse primers produce sequences complimentary to the rRNA. Primers rP1, rP2, and rP3 are identical except for the 17th base from the 3′ end. Under most amplification conditions, they should be functionally equivalent. Primer rP2 has the sequence corresponding to the greatest diversity of bacteria.

```
CONS=90%                  AGAGUUUGAUC UGGCUCAG   GAACGCUGGCGG  - GC U A ACAUGCAAGUCG  CG
E. coli                AAAUUGAAGAGUUUGAUCAUGGCUCAGAUUGAACGCUGGCGGCA-GGCCUAACACAUGCAAGUCGAACGGUAACAGGAAGAAGCUUGC
An.margi                  agaguuugauccuggcucagAACGAACGCUGGCGGCA-AGCUUAACACAUGCAAGUCGAACGGACCGUAUACGCAGCUUGC
fD1       5'ccgaattcgtcgacaacAGAGTTTGATCCTGGCTCAG-3' (extend) ---->
fD2       5'ccgaattcgtcgacaacAGAGTTTGATCATGGCTCAG-3' (extend) ---->
fD3       5'ccgaattcgtcgacaacAGAGTTTGATCCTGGCTTAG-3' (extend) ---->
fD4       5'ccgaattcgtcgacaacAGAATTTGATCTTGGTTCAG-3' (extend) ---->

CONS=90%               G GGC  ACGGGUG GUAA    U    A  U CC          GG A A C     GAAA        UAAUACC  AU
E. coli   UUCUUUGCUGACGAGUGGCGGACGGGUGAGUAAUGUCUGGGAAACUGCCUGAUGGAGGGGGAUAACUACUGGAAACGGUAGCUAAUACCGCAUAACGUCG
An.margi  UGCGUGUAUGGUUAGUGGCAGACGGGUGAGUAAUGCAUAGGAAUCUACCUAGUAGUAUGGGAUAGCCACUAGAAAUGGUGGGUAAUACUGUAUAAUCCUG

CONS=90%       AAAG                              GA              AU AG U GUUGG   GGUAA GGC  ACCAAG C A GA    U
E. coli   CAAGACCAAAGAGGGGGACCUUCGGGCCUCUUGCCAUCGGAUGUGCCCAGAUGGGAUUAGCUAGUAGGUGGGGUAACGGCUCACCUAGGCGACGAUCCCU
An.margi  C-GGGGGAAAGA--------UUUA------UCGCUAUUAGAUGAGCCUAUGUCAGAUUAGCUAGUUGGUGGGGUAAUGGCCUACCAAGGCGGUGAUCUGU

CONS=90%  A C G  CUGAGAGG  GA C G CACA UGG ACUGAGACACGG CCA ACUCCUACGGGAGGCAGCAGU  GGAAU UU   CAAUGG  G AA  CU
E. coli   AGCUGGUCUGAGAGGAUGACCAGCCACACUGGAACUGAGACACGGUCCAGACUCCUACGGGAGGCAGCAGUGGGGAAUAUUGCACAAUGGGCGCAAGCCU
An.margi  AGCUGGUCUGAGAGGAUGAUCAGCCACACUGGAACUGAGACACGGUCCAGACUCCUACGGGAGGCAGCAGUGGGGAAUAUUGGACAAUGGGCGCAAGCCU

CONS=90%  GA   AGC A GCCGCGUG   GA GA G    U GG   GUAAA   CU U       GA G                                  UGAC  UA
E. coli   GAUGCAGCCAUGCCGCGUGUAUGAAGAAGGCCUUCGGGUUGUAAAGUACUUUCAGCGGGGAGGAAGGGAGUAAAGUUAAUACCUUUGCUCAUUGACGUUA
An.margi  GAUCCAGCUAUGCCGCGUGAGUGAGGAAGGCCUUAGGGUUGUAAAACUCUUUCAGUAGGGAAGAU----------------------------AAUGACGGUA

CONS=90%        A  AAGC  CGGCUAACU  GUGCCAGCAGCCGCGGUAAUAC  AGG  GC AGCGUU   CGGA U A UGGGCGUAAAG G   G AGG G
E. coli   CCCGCAGAAGAAGCACCGGCUAACUCCGUGCCAGCAGCCGCGGUAAUACGGAGGGUGCAAGCGUUAAUCGGAAUUACUGGGCGUAAAGCGCACGCAGGCG
An.margi  CCUACAGAAGAAGUCCCGGCAAACUCCGUGCCAGCAGCCGCGGUAAUACGGAGGGGGCAAGCGUUGUUCGGAAUUAUUGGGCGUAAAGGGCAUGUAGGCG

CONS=90%  G       AGU  G  GU AAA       GCU AAC                A AC       CU GA       AG GG       GAAUU    GUGUA
E. coli   GUUUGUUAAGUCAGAUGUGAAAUCCCCGGGCUCAACCUGGGAACUGCAUCUGAUACUGGCAAGCUUGAGUCUCGUAGAGGGGGGUAGAAUUCCAGGUGUA
An.margi  GUUUGGUAAGUUAAAGGUGAAAUACCAGGGCUUAACCCUGGGGCUGCUUUUAAUACUGCAGGACUAGAGUCCGGAAGAGGAUAGCGGAAUUCCUAGUGUA

CONS=90%  G GGUGAAAU CGUAGA AU    A GAA ACC  U GCGAAGGC      CUGG      A UGAC CU A G  CGAAAGCGUGGGGAGC AACAGGA
E. coli   GCGGUGAAAUGCGUAGAGAUCUGGAGGAAUACCGGUGGCGAAGGCGGCCCCCUGGACGAAGACUGACGCUCAGGUGCGAAAGCGUGGGGAGCAAACAGGA
An.margi  GAGGUGAAAUUCGUAGAUAUUAGGAGGAACACCAGUGGCGAAGGCGGCUGUCUGGUCCGGUACUGACGCUGAGGUGCGAAAGCGUGGGGAGCAAACAGGA

CONS=90%  UUAGAUACCCUGGUAGUCCACGC  UAAACGAUG     U    GU  G                G      AG  AACGC  UAA     CCGCCUGGG
E. coli   UUAGAUACCCUGGUAGUCCACGCCGUAAACGAUGUCGACUUGGAGGUUGUGCCCUUGAGGCGUGGCUUCCGGAGCUAACGCGUUAAGUCGACCGCCUGGG
An.margi  UUAGAUACCCUGGUAGUCCACGCUGUAAACGAUGAGUGCUGAAUGUGGGGGC-UUUU--GCCUCUGUGUUGUAgcUAACGCGUUAAGCACUCCGCCUGGG

CONS=90%   AGUACG  CGCAAG  U AAACUCAAA GAAUUGACGGGG CCCGCACAAGCGG GGAG AUGUGGUUUAAUUCGA G  ACGCG  GAACCUUACC
E. coli   GAGUACGGCCGCAAGGUUAAAACUCAAAUGAAUUGACGGGGGCCCGCACAAGCGGUGGAGCAUGUGGUUUAAUUCGAUGCAACGCGAAGAACCUUACCUG
An.margi  GACUACGGUCGCAAGACUAAAACUCAAAGGAAUUGACGGGGACnCGCACAAGCGGUGGAGCAUGUGGUUUAAUUCGAUGCAACGCGAAAAACCUUACCAC

CONS=90%     UUGACAU              -GA A      -        -          ACAGGUG UGCAUGG UGUCGUCAGCUCGUG CGUGAG U
E. coli   GUCUUGACAUCCACGGAA-GUUUUCA-GAGAUGAGAAU-GUGCCUUCG-GGAACCGUGAGACAGGUGCUGCAUGGCUGUCGUCAGCUCGUGUUGUGAAAU
An.margi  UUCUUGACAUGGAGGCUAGAUCCUUCUUAACAGAAGGGCG-CAGUUCGGCUGGGCCUCGCACAGGUGCUGCAUGGCUGUCGUCAGCUCGUGUCGUGAGAU

CONS=90%  GUUGGGUUAAGUCCCGCAACGAGCGCAACCC          GUU C A C          G G ACUC       ACUGCC   G  AA   GGAGGAAGG
E. coli   GUUGGGUUAAGUCCCGCAACGAGCGCAACCCUUAUCCUUUGUUGCCAGCGGUC-CGGCCGGGAACUCAAAGGAGACUGCCAGUGAUAAACUGGAGGAAGG
An.margi  GUUGGGUUAAGUCCCGCAACGAGCGCAACCCUCAUCCUUAGUUACCAGCGGGUAAUGCCGGGCACUUUAAGGAAACUGCCAGUGAUAAACUGGAGGAAGG

CONS=90%   G GGA GACGUCAA UC UCAUG CCCUUA G    GGGCUACACACGU CUACAAUGG     ACA  G G  GC A    G GA     AGC AA C
E. coli   UGGGGAUGACGUCAAGUCAUCAUGGCCCUUACGACCAGGGCUACACACGUGCUACAAUGGCGCAUACAAAGAGAAGCGACCUCGCGAGAGCAAGCGGACC
An.margi  UGGGGAUGAUGUCAAGUCAGCACGGCCCUUAUGGGGUGGGCUACACACGUGCUACAAUGGCGACUACAAUAGGUUGCAACGUCGCAAGGCUGAGCUAAUC

CONS=90%      AAA     UC  AGU CGGAU G    CUGCAACUCG    C UGAAG  GGA U GCUAGUAAUCG   AUCAG A     CGGUGAAUACGUUC
E. coli   UCAUAAAGUGCGUCGUAGUCCGGAUUGGAGUCUGCAACUCGACUCCAUGAAGUCGGAAUCGCUAGUAATCGGAAUCGUGGAUCAGAAUGCCACGGUGAAUACGUUC
An.margi  CGU-AAAAGUCGUCUCAGUUCGGAUUGUCCUCUGUAACUCGAGGGCAUGAAGUCGGAAUCGCUAGUAATCGUGGAUCAGCAUGCCACGGUGAAUACGUUC

CONS=90%  CGGG CUUGUACACACCGCCCGUCA   CA G  AG            AAG          AACC    GGA         C A  GU  G
E. coli   CCGGGCCUUGUACACACCGCCCGUCACACCAUGGGAGUGGGUUGCAAAAGAAGUAGGUAGCUUAACCUUCGGGAGGGCGCUUACCACUUUGUGAUUCAUG
An.margi  UCGGGUCUUGUACACACUGCCCGUCACGCCAUGGGAAUUGGCUUAACUCGAAGCUGGUGCGCCAACCGUAAGGAGGCAGCCAUUUAAGGUUGGGUCGGUG

CONS=90%  A UGGG   AAGUCGUAACAAGGUA CC UA   GAA  UG GG UGGAU ACCUCCUUU
E. coli   ACUGGGGUGAAGUCGUAACAAGGUAACCGUAGGGGAACCUGCGGUUGGAUCACCUCCUUA
An.margi  ACUGGGGUGAAGUCGUAACAAGGUAGCUGUAGGUGAACCUGCggcuggaucaccuccuu
rD1                          <---- (extend) 3'-CCGACCTAGTGGAGGAAttcgaacctagggccc5'
rP1       <---- 3'-TTCAGCATTGTTCCATTGGCAttcgaacctagggccc-5'
rP2       <---- 3'-TTCAGCATTGTTCCATCGGCAttcgaacctagggccc-5'
rP3       <---- 3'-TTCAGCATTGTTCCATAGGCAttcgaacctagggccc-5'
```

TABLE 2. Primer combinations that have been proven to produce an approximately 1,500-bp fragment

Species	Primer pair
Neisseria gonorrhoeae	fD1 + rD1
Coxiella burnetii	fD1 + rD1
Anaplasma marginale	fD1 + rD1
Neisseria meningitidis	fD1 + rD1
Bacteroides fragilis	fD1 + rP2
Borrelia burgdorferi	fD3 + rD1
Borrelia hermsii	fD3 + rD1
Clostridium perfringens	fD1 + rD1
Mycoplasma pneumoniae	fD1 + rP1
Mycoplasma hominis	fD1 + rP1
Mycoplasma genitalium	fD1 + rP1
Ureaplasma urealyticum	fD1 + rP1
Campylobacter jejuni	fD1 + rP1
Shigella flexneri	fD2 + rP1
Shigella sonnei	fD2 + rP1
Chlamydia psittaci	fD4 + rD1
Chlamydia trachomatis	fD4 + rD1
Chlamydia pneumoniae	fD4 + rD1
Mycobacterium bovis	fD1 + rD1
Legionella pneumophila	fD1 + rD1

(1; see Sequenase package insert), including the alkaline denaturation method recommended by the manufacturer.

For sequencing of amplified material directly, four identical 100-μl amplification reactions were performed on each sample, with the resultant material being pooled and purified (see above). A 500-ng amount of template (amplification product) was combined with 10 ng of primer, 2 μl of Sequenase buffer, and water to 10 μl. This sample was held at 98°C for 7 min and cooled to room temperature for 1 min, and then the labeling reaction was performed at either room temperature or 37°C for 5 min. (These are slight modifications of the procedures outlined in reference 19.) Chain elongation was terminated with sample loading buffer, and sequencing was performed on buffer-gradient gels (1).

Sequencing primers. Sequencing primers useful for conserved regions within 16S rDNA genes, both forward and reverse, have been described previously (11, 20). Forward primers used in this study spanned positions (*Escherichia coli* numbers) (4) as follows: 339 to 357, 785 to 805, 907 to 926, 1391 to 1406. Reverse primers included 357 to 342, 536 to 519, 802 to 785, 926 to 907, 1115 to 1100, 1406 to 1392, and 1513 to 1494.

RESULTS AND DISCUSSION

Primers for 16S rDNA gene amplification. The primers used for amplification of bacterial 16S rDNA are displayed in Table 1 as well as in an aligned format in Fig. 1. Their empirical and theoretical utility are described in Tables 2 and 3. Their implied hybridization is measured against the collection of complete 16S rRNA sequences available to us at the present time. It would be possible to introduce nucleotide ambiguities during the DNA synthesis of these primers to obtain a smaller set of primers that would work on virtually all bacteria, although it is likely that this would cause the appearance of spurious amplification products. The primer pair designated fD1 and rD1 is capable of amplifying a wide variety of bacterial taxa. Replacing rD1 with rP1 extends the diversity of species even further, but from the perspective of amplifying the maximum number of nucleotides of 16S rDNA, rD1 is preferable; it is closer to the 3′ end.

Amplification of 16S rDNAs from several different bacterial taxa by using different combinations of primers is shown in Fig. 2. The PCR conditions used in this study were not rigorously optimized for either specificity or low-input target. In general, these conditions seem to give a dependable yield of full-sized product. Additional products of varying intensities were produced from some of the DNAs. As shown in Fig. 2, the smaller-sized band produced from *Bacteroides fragilis* DNA, derived from both culture and lyophilized ampoule, is the most abundant spurious band within the bacterial strains examined. The occurrence of the additional PCR product in *B. fragilis* did not interfere with the ability to clone the 16S rDNA. Of note is the fact that the only reaction shown in which the rP2 primer was used was the *B. fragilis* amplification; in another experiment (not shown), the same spurious band was produced when primer rP1 was substituted for rP2.

***A. marginale* 16S rRNA.** To test the utility of these methods for phylogenetic study, *A. marginale*, a member of the *Rickettsiales*, was used as an example. *A. marginale* is an obligate intracellular, erythrocytic, arthropod-transmitted parasite whose pathogenic range includes several ruminant species (16). Previous attempts at sequencing directly from the RNA isolated from Renografin-purified *A. marginale* cells (supplied by J. Samuel, Pullman, Wash.) were unsuccessful (data not shown). Attempts at cloning into lambda phage vectors also were without success. A DNA preparation (by standard phenol methods) (13, 14) from this same gradient-purified material was used for amplification. By using the generally applicable primer pair, fD1 + rD1, amplification of the 16S rDNA of *A. marginale* was enabled. The 16S rRNA sequence determined for *A. marginale* is shown in Fig. 1, aligned with *E. coli* and with the primers.

Approximately 250 nucleotides of sequence were initially determined directly from PCR-amplified material by using a primer located at position 519 (*E. coli* numbers) (4), reading toward the 5′ end of the molecule. After cloning the amplification product into a pGEM vector, that same sequence was again determined to validate the accuracy of the direct sequencing, without disagreement. Aligning the sequence with the accepted secondary structure (8, 25) indicated a deleted helix within the 455 to 480 region (*E. coli* numbers) (4) characteristic of a few bacterial groups (8, 25). The complete *A. marginale* 16S rRNA sequence was determined from the cloned amplification product. The structure of the helices between positions 180 and 220 (*E. coli* numbers) (4) strongly suggested that this was a member of the alpha subdivision of the purple bacteria. Similarity and evolutionary distance data (Table 4) prove *A. marginale* to be related to other *Rickettsiales* within the alpha subdivision of the purple bacteria (19, 21), specifically related to the genera *Rickettsia* and *Ehrlichia* (19). A phylogenetic tree displaying this three-way relationship is shown in Fig. 3.

FIG. 1. Sequence alignment of the amplification primers with 16S rRNAs of *E. coli*, *A. marginale*, and a eubacterial consensus sequence. The consensus shows positions that are greater than 90% conserved for a phylogenetically diverse collection of approximately 85 bacterial sequences.

TABLE 3. Theoretical specificity of amplification primers for 16S rDNA

Primer	Phylogenetic grouping and genera which should amplify with indicated primer[a]
fD1	Gram-positive bacteria and relatives *Bacillus*, *Clostridium*, *Staphylococcus*, *Listeria*, *Lactobacillus*, *Streptococcus*, *Mycoplasma*, *Spiroplasma*, *Ureaplasma*, *Acholeplasma*, *Erysipelothrix*, *Fusobacterium*, *Arthrobacter*, *Mycobacterium*, *Streptomyces* Purple bacteria and relatives (proteobacteria) *Rochalimaea*, *Brucella*, *Rhodopseudomonas*, *Agrobacterium*, *Rhodospirillum*, *Pseudomonas*, *Neisseria*, *Caulobacter*, *Myxococcus*, *Campylobacter*, *Rickettsia*, *Ehrlichia* Cyanobacteria *Anacystis* (*Synechococcus*) *Bacteroides*/flavobacteria *Bacteroides*, *Flavobacterium* *Deinococcus* and relatives *Deinococcus*, *Thermus* Spirochetes *Treponema*, *Spirochaeta* *Planctomyces* and relatives *Planctomyces* *Chlorobium*-green sulfur bacteria *Chlorobium* *Thermotoga* *Thermotoga*
fD2	Enteric members of gamma subdivision of proteobacteria *Escherichia*, *Shigella*, *Salmonella*, *Serratia*, *Erwinia*, and *Citrobacter*, etc. (all the enterics); *Oceanospirillum*, *Haemophilus*, *Actinobacillus*, *Vibrio*, *Pasteurella*
fD3	Spirochetes of the genus *Borrelia*
fD4	Genus *Chlamydia*
rD1	Purple bacteria and relatives (proteobacteria) *Pseudomonas*, *Neisseria*, *Rochalimaea*, *Agrobacterium*, *Myxococcus*, *Desulfovibrio* Gram-positive bacteria and relatives *Bacillus*, *Staphylococcus*, *Arthrobacter*, *Streptomyces*, *Mycobacterium*, *Heliobacterium* *Cyanobacteria* *Anacystis* (*Synechococcus*) Spirochetes *Treponema*, *Leptospira* Planctomyces *Planctomyces* *Chlorobium* *Chlorobium* *Thermotoga* *Thermotoga* (plus selected archaebacteria)
rP1, rP2, or rP3 (probably all functionally equivalent)	Should prime all bacteria, plus plant mitochondria, chloroplasts, archaebacteria, and *Dictyostelium*, but not yeasts or vertebrates

[a] Primers are considered applicable if there is a perfect match for approximately 15 bases at the 3′ end of the primer. The list is definitive only in the sense that the taxa mentioned represent the sequences available to the authors. The absence of a genus from the list does not imply that the primer will not work. Because the majority of the available rRNA sequences are derived from direct sequencing of rRNAs with reverse transcriptase, there is far less information available about the 3′ end of the 16S. In some cases, the indicated genus is represented by numerous species; in other cases the indicated genus is represented by only one. The sequence alignment from which these data were derived is unpublished (12, 24). Phylogenetic groupings are those of Woese (23).

A. marginale has historically been a difficult organism to study (16) because of the lack of a culture system. Its phylogenetic placement, which was made possible by amplification with broad-specificity 16S rDNA PCR primers, has implications which should shed light on the biology of this fastidious pathogen. Many properties found in the genera *Rickettsia* or *Ehrlichia*, such as the utilization of host ATP (22), may also be characteristic of the genus *Anaplasma*. Obviously, the sequence of the 16S rRNA of *A. marginale* provides basic information useful for designing probes or PCR primers specific for the detection of this veterinary pathogen.

As shown with *A. marginale*, a potential advantage of this technology in the study of bacterial phylogeny is the ability to work with relatively small numbers of cells. This ability is of greatest use when the target species is either extremely fastidious or highly pathogenic. In the case of *A. marginale*, obtaining sufficient material for phylogenetic study would have required the sacrifice of several cows; this is the only established culture system. It is also highly desirable to keep culture volumes small when dealing with highly infectious organisms such as respiratory pathogens.

16S rDNA amplified from ATCC lyophilized ampoules. One obvious source of either fastidious or pathogenic organisms

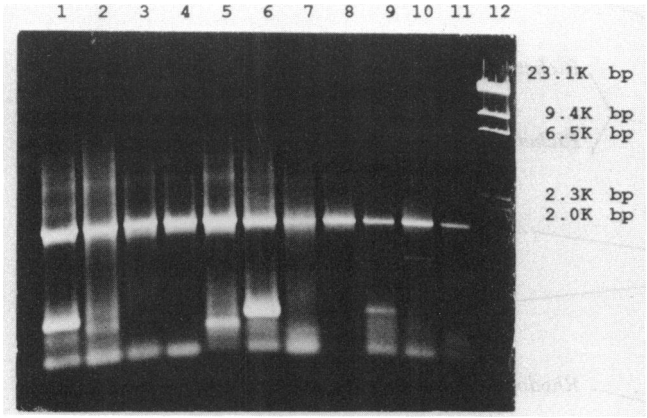

FIG. 2. Ethidium-bromide-stained 0.75% agarose gel displaying amplification products. Lanes: 1, *B. fragilis* DNA (fD1 + rP2); 2, *C. perfringens* DNA (fD1 + rD1); 3, *C. psittaci* DNA (fD4 + rD1); 4, *B. burgdorferi* DNA (fD3 + rD1); 5, lyophilized ampoule-derived *S. sanguis* DNA (fD1 + rP2); 6, lyophilized ampoule-derived *B. fragilis* DNA (fD1 + rP2); 7, lyophilized ampoule-derived *C. perfringens* DNA (fD1 + rD1); 8, lyophilized ampoule-derived *Y. enterocolitica* DNA (fD2 + rP1); 9, lyophilized ampoule-derived *P. magnus* DNA (fD1 + rD1); 10, lyophilized ampoule-derived *M. smegmatis* DNA (fD1 + rP1); 11, lyophilized ampoule-derived *M. phlei* DNA (fD1 + rD1); 12, *Hin*dIII digest of lambda phage. Labeled bands include 23,130, 9,416, 6,557, 2,322, and 2,027 bp.

which are in viable condition is a culture collection such as the American Type Culture Collection (ATCC). According to the *Manual of Methods for General Bacteriology* (7), a standard double-vial lyophilized ampoule contains 2×10^5 cells. Depending on the quantitation of the input cells (for example, counting of cells in terms of viable CFU) and the growth phase of the preserved cells, there could be several-fold more genomes present in the lyophilized ampoule. Factoring in the multicistronic nature of rRNA operons (seven copies in *E. coli*) (10) contributes to even greater target numbers available for gene amplification.

Lyophilized ampoules of various types from diverse taxa were selected to examine their potential to yield DNA which would initiate 16S rDNA amplification. DNAs from both ATCC double-vial and single-vial preparations were purified as described in Materials and Methods. The following strains were amplified with the indicated primer pair (Fig. 2): *Streptococcus sanguis* ATCC 10556 (fD1 + rD1), *B. fragilis* ATCC 25285 (fD1 + rP2), *Clostridium perfringens* ATCC 13124 (fD1 + rD1), *Peptostreptococcus magnus* ATCC 29328 (fD1 + rD1), *Yersinia enterocolitica* ATCC 23715 (fD2 + rP1), *Mycobacterium smegmatis* ATCC 14468 (fD1 + rP1), and *Mycobacterium phlei* ATCC 11758 (fD1 + rP1). The last four species listed above were also amplified with a forward primer similar to fD1; its rDNA-like sequence was that of fD1, but the linker was replaced with a T7 promoter sequence (see discussion of transcription below). The fact that *Y. enterocolitica* amplified with this primer suggests that, under the right conditions, one forward and one or two reverse primers may amplify almost all of the eubacterial 16S rDNAs. A similar interchangeability of primers was found with *Campylobacter jejuni* rDNA (data not shown).

The lyophilized ampoules of *S. sanguis* actually had an expiration date 8 months prior to these experiments. Methods even simpler than the single phenol extraction or bead-beating method were tested with *S. sanguis*-lyophilized ampoules. The simplest method involved resuspending a lyophilized ampoule in 100 μl of the buffer (described above) containing 0.45% each of Tween 20 and Nonidet P-40 detergents. This sample was boiled for 5 min, and 1 μl was used to prime an amplification reaction. Although successful, this method was rejected from further exploration because it is not likely to work for every bacterium, unlike the bead-beating method.

Fidelity of amplification. To examine whether there were potential artifacts due to PCR amplification, portions of several sequences were examined and compared with sequences which had previously been determined by other methods. A comparison of a clone derived from PCR with one derived by conventional means should be the most informative. The following numbers of bases were determined from the clones indicated (all in pGEM vectors). (i) An amplification clone-derived *Borrelia burgdorferi* sequence revealed one discrepancy within 1,517 bases as compared with an RNA sequence (unpublished data), resolved in favor of the clone sequence, on the basis of sequence composition of related species at this position. In addition, seven unresolved nucleotide assignments (Ns) from within the sequenceable region of the rRNA sequence were determined from the clone, and several other bases were determined from the 3′ end of the rRNA. (Approxi-

TABLE 4. Percentage similarity and evolutionary distance (9) for nine bacteria belonging to the alpha subdivision of the purple bacteria (23), plus *E. coli* (a gamma bacterium) as an outgroup[a]

Bacterium	% Similarity/evolutionary distance (×100)[b]									
	E. coli	*R. palustris*	*R. rubrum*	*A. marginale*	*E. risticii*	*R. prowazekii*	*R. rickettsii*	*R. quintana*	*B. abortus*	*A. tumefaciens*
Escherichia coli	—	81.0	84.0	81.2	78.6	80.5	80.3	81.2	81.9	81.2
Rhodopseudomonas palustris	21.8	—	88.5	83.3	82.1	85.4	85.3	89.3	90.0	89.3
Rhodospirillum rubrum	18.0	12.4	—	84.9	84.9	84.9	84.9	87.9	88.9	88.4
Anaplasma marginale	21.5	18.8	16.8	—	86.9	86.0	86.2	85.6	85.3	85.8
Ehrlichia risticii	25.2	20.4	20.0	14.3	—	84.6	84.6	83.2	83.8	84.2
Rickettsia prowazekii	22.7	16.2	16.8	15.4	17.2	—	99.0	87.0	87.1	86.9
Rickettsia rickettsii	22.7	16.3	16.8	15.2	17.1	0.9	—	87.1	87.1	86.9
Rochalimaea quintana	21.7	11.5	13.1	16.0	18.9	14.2	14.1	—	95.3	94.7
Brucella abortus	20.7	10.7	11.9	16.3	18.2	14.1	14.1	4.8	—	94.9
Agrobacterium tumefaciens	21.7	11.5	12.6	15.7	17.6	14.3	14.3	5.5	5.2	—

[a] A mask was used which eliminated a small number of positions from consideration within the alignment; all positions in which base composition was not at least 50% conserved were eliminated. All of the sequences represented may be obtained from Genbank except *R. palustris* which was used courtesy of C. R. Woese.

[b] Numbers below the diagonal indicate evolutionary distance.

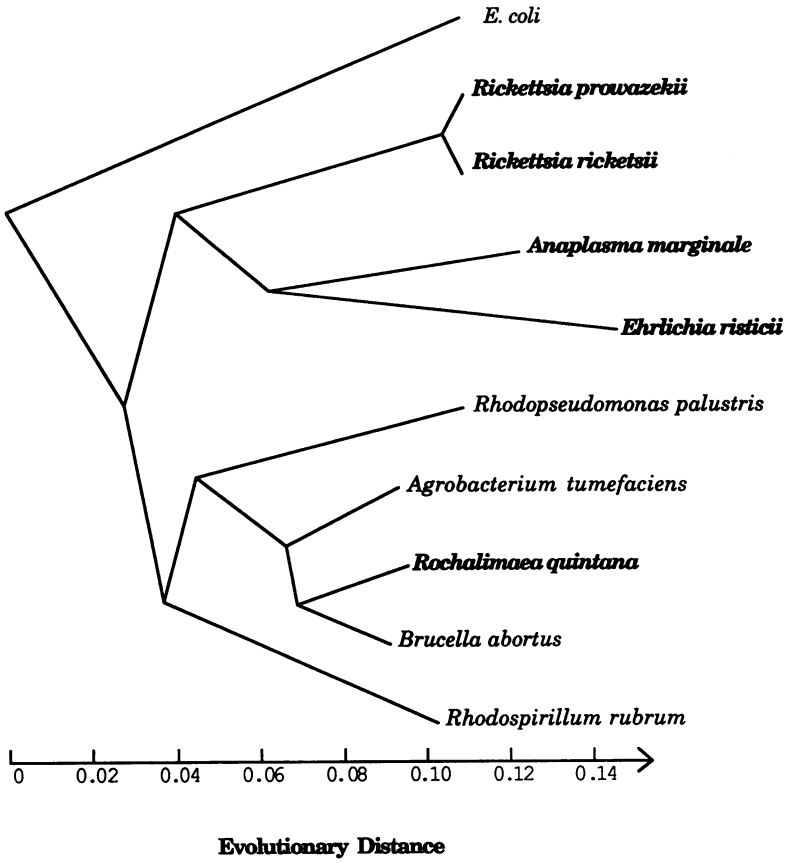

FIG. 3. Phylogenetic distance tree displaying the evolutionary origin of *A. marginale* within a lineage shared by the genera *Rickettsia* and *Ehrlichia*. All species belonging to the order *Rickettsiales* are shown in boldface type. *E. coli* is used as an outgroup sequence.

mately 60 nucleotides at the 3′ end of 16S rRNA cannot be sequenced directly from the rRNA.) (ii) A total of 226 bases were determined from a *Neisseria gonorrhoeae* PCR clone, with no disagreement with the RNA sequence or a published rDNA sequence (17). (iii) The sequence from the *Neisseria meningitidis* clone showed no disagreement with 250 bases of RNA-derived sequence. One N from the RNA sequence was now determinable from the clone. (iv) In 212 bases of a *Chlamydia trachomatis* sequence, there was no disagreement with the conventionally cloned sequence (unpublished data). (v) In a *C. jejuni* PCR-derived clone, 328 bases were inspected, and there was one discrepancy with the RNA-derived sequence (an A-to-G transition) which may be the result of minor cistron-to-cistron variation. Five N's were resolved.

In summary, there appears to be no reason for more than the usual concern, because of *Taq* polymerase fidelity of replication, about sequence accuracy in this method. Several positions that had previously been scored as unknown or ambiguous are now determinable by taking advantage of bidirectional gene sequencing and the absence of artifacts due to rRNA secondary structure and posttranscriptionally modified nucleotides. As always, a cautious approach should be taken by checking any sequence anomalies and confirming known secondary structural constraints on sequences.

Although we have spent considerable time optimizing and testing the direct sequencing of PCR-amplified rDNA, we highly recommend cloning the fragments if a near-perfect sequence is desired.

Transcription of amplified material. One additional application of primers similar to those described herein is the in vitro transcription of PCR-amplified material. We have transcribed virtually full-length pseudo-rRNA from both cloned material and directly from amplified product (not shown). Cloned rDNAs, in pGEM vectors, were transcribed by using either T7 or SP6 polymerases, depending on clone orientation. In addition, PCR primers similar to fD1, in which the cloning linker was replaced with a T7 promoter sequence (5′-TTAATACGACTCACTATAGGG-followed by the 16S rDNA-like sequence), readily amplified and transcribed full-length pseudo-rRNA. This material can be used for sequencing, hybridization studies (data not shown), or potentially for reconstitution experiments.

Conclusion. The amplification by PCR of a taxonomically diverse collection of eubacterial 16S rDNA genes is possible with a small number of primers. These products can readily be cloned for sequencing or they can be sequenced directly. The ability to determine rRNA sequences from ATCC lyophilized ampoules, without culture, enables the study of fastidious or pathogenic species without employing tricky or expensive microbiological methods. While this should not be a routine substitute for growing bacteria, picking individual colonies, and confirming their phenotypic and biochemical identities, it will enable experiments to be performed that were not previously possible.

ACKNOWLEDGMENTS

We greatly appreciate the work of J. Samuel, who provided *A. marginale* cells purified from bovine erythrocytes. Carl Woese graciously provided unpublished rRNA sequences for analysis.

REFERENCES

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. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1518.
3. **Boettger, E. C.** 1989. Rapid determination of bacterial ribosomal RNA sequences by direct sequencing of enzymatically amplified DNA. FEMS Microbiol. Lett. **65**:171–176.
4. **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.
5. **Edwards, U., T. Rogall, H. Bloecker, M. Emde, and E. Boettger.** 1989. Isolation and direct complete nucleotide determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. Nucleic Acids Res. **17**:7843–7853.
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. **Gherna, R. L.** 1981. Preservation, p. 208–217. *In* P. Gerhardt, R. G. E. Murray, R. N. Costilow, E. W. Nester, W. A. Wood, N. R. Krieg, and G. B. Phillips (ed.), Manual of methods for general bacteriology. American Society for Microbiology, Washington, D.C.
8. **Gutell, R. R., B. Weiser, C. R. Woese, and H. F. Noller.** 1985. Comparative anatomy of 16S-like ribosomal RNA. Prog. Nucleic Acid Res. Mol. Biol. **32**:155–216.
9. **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.
10. **Kenerley, M. E., E. A. Morgan, L. Post, L. Lindahl, and M. Nomura.** 1977. Characterization of hybrid plasmids carrying individual ribosomal ribonucleic acid transcription units of *Escherichia coli*. J. Bacteriol. **132**:931–949.
11. **Lane, D., 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 analysis. Proc. Natl. Acad. Sci. USA **82**:6955–6959.
12. **Lane, D. J., and W. G. Weisburg.** Unpublished data.
13. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. **Marmur, J.** 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. J. Mol. Biol. **3**:208–218.
15. **Medlin, L., H. J. Elwood, S. Stickel, and M. L. Sogin.** 1988. The characterization of enzymatically amplified eukaryotic 16S-like rRNA coding regions. Gene **71**:491–499.
16. **Ristic, M., and J. P. Kreier.** 1984. Family III. *Anaplasmataceae* Philip 1957, 980^{AL}, p. 719–729. *In* N. R. Krieg and J. H. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 1. The Williams & Wilkins Co., Baltimore.
17. **Rossau, R., L. Heyndrickx, and H. V. Heuverswyn.** 1988. Nucleotide sequence of a 16S ribosomal RNA gene from *Neisseria gonorrhoeae*. Nucleic Acids Res. **16**:6227.
18. **Thein, S. L.** 1989. A simplified method of direct sequencing of PCR amplified DNA with Sequenase T7 DNA polymerase, p. 8. *In* U.S. Biochemical technical literature, vol. 16, no. 1. Cleveland, Ohio.
19. **Weisburg, W. G., M. E. Dobson, J. E. Samuel, G. A. Dasch, L. P. Mallavia, O. Baca, L. Mandelco, J. E. Sechrest, E. Weiss, and C. R. Woese.** 1989. Phylogenetic diversity of the rickettsiae. J. Bacteriol. **171**:4202–4206.
20. **Weisburg, W. G., Y. Oyaizu, H. Oyaizu, and C. R. Woese.** 1985. Natural relationship between bacteroides and flavobacteria. J. Bacteriol. **164**:230–236.
21. **Weisburg, W. G., C. R. Woese, M. E. Dobson, and E. Weiss.** 1985. A common origin of ricketsiae and certain plant pathogens. Science **230**:556–558.
22. **Weiss, E., J. C. Williams, G. A. Dasch, and Y.-H. Kang.** 1989. Energy metabolism of monocytic *Ehrlichia*. Proc. Natl. Acad. Sci. USA **86**:1674–1678.
23. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. **51**:221–271.
24. **Woese, C. R.** Unpublished data.
25. **Woese, C. R., R. Gutell, R. Gupta, and H. R. Noller.** 1983. Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. Microbiol. Rev. **47**:621–669.