EVE 161:
Microbial Phylogenomics

Class #8:
Era II: rRNA Gene PCR for studying microbial diversity

UC Davis, Winter 2018
Instructor: Jonathan Eisen
Teaching Assistant: Cassie Ettinger

Midterm

- Midterm will be posted next week
- Answering a series of questions
- Hint - it might be about a paper we talked about on Day 1 …

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

JOURNAL OF BACTERIOLOGY, Sept. 1998, p. 4765–4774
0021-9193/98/$04.00+0

Vol. 180, No. 18

MINIREVIEW

Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity

PHILIP HUGENHOLTZ, BRETT M. GOEBEL,† AND NORMAN R. PACE*
Departments of Plant and Microbial Biology and Molecular and Cell Biology, University of California, Berkeley, Berkeley, California 94720-3102

Presenters?

Tree of Life
Eukaryotes
Plants
Alveolates
Heterokonts
Rhizaria
Opisthokonts
Discicristates
Excavates
Bacteria
Archaea
Crenarchaea
Euryarchaea
ROOT?

Questions?

Tree of Life
Eukaryotes
Plants
Land plants
Green algae
Red algae
Glaucophyte algae
Alveolates
Ciliates
Dinoflagellates
Apicomplexa
Heterokonts
Oomycetes
Diatoms
Phaeophytes
Chrysophytes
Opalinids
Haptophytes
Cryptophytes
Rhizaria
Foraminifera
Radiolarians
Cellular slime molds
Plasmodial slime molds
Amoebas
Mesomycetozoa
Choanoflagellates
Animals
Opisthokonts
Fungi
Acrasid slime molds
Vahlkampfiid amoebas
Euglenids
Discicristates
Core jakobids
Trypanosomes
Leishmanias
Parabasalids
Diplomonads
Oxymonads
Excavates
Bacteria
Acidobacteria
Verrucomicrobia
Chloroflexi (GNSB group)
Mitochondria
Spirochaetes
Planctomycetes
Chlamydia
Actinobacteria
Cyanobacteria
Plastids
Firmicutes
Mycoplasmas
Thermotogales
Aquificae
ROOT?
Archaea
Crenarchaea
Korarchaeota
Sulfolobus
Pyrobaculum
Nanoarchaeota
Halophiles
Methanosarcina
Thermoplasma
Pyrococcus
Euryarchaea

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.

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   CCCGCAGAAGAAGCACCGGCUAACUCCGUGCCAGCAGCCGCGGUAAUACGGAGGGUGCAAGCGUUAAUCGGAATTACTGGGCGUAAAGCGCACGCAGGCG
An.margi  CCUACAGAAGAAGUCCCGGCAAACUCCGUGCCAGCAGCCGCGGUAAUACGGAGGGGGCAAGCGUUGUUCGGAATTATTGGGCGUAAAGGGCAUGUAGGCG

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

CONS=90%  G GGUGAAAU CGUAGA AU    A GAA ACC  U GCGAAGGC       CUGG      A UGAC CU A G  CGAAAGCGUGGGGAGC AACAGGA
E. coli   GCGGUGAAAUGCGUAGAGAUCUGGAGGAATACCGGUGGCGAAGGCGGCCCCCUGGACGAAGACUGACGCUCAGGUGCGAAAGCGUGGGGAGCAAACAGGA
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   GAGUACGGCCGCAAGGUUAAAACUCAAAUGAAUUGACGGGGCCCGCACAAGCGGUGGAGCAUGUGGUUUAAUUCGAUGCAACGCGAAGAACCUUACCUG
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   UGGGGAUGACGUCAAGUCAUCAUGGCCCUUACGACCAGGGCUACACACGUGCUACAAUGGCGCAUACAAAGAAAGCGACCUCGCGAGAGCAAGCGGACC
An.margi  UGGGGAUGAUGUCAAGUCAGCACGGCCCUUAUGGGGUGGGCUACACACGUGCUACAAUGGCGACUACAAUAGGUUGCAACGUCGCAAGGCUGAGCUAAUC

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

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 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    UCAUAAAGUGCGUCGUAGUCCGGAUUGGAGUCUGCAACUCGACUCCAUGAAGUCGGAAUCGCUAGUAAUCGUGGAUCAGAAUGCCACGGUGAAUACGUUC
An.margi   CGU-AAAAGUCGUCUCAGUUCGGAUUGUCCUCUGUAACUCGAGGGCAUGAAGUCGGAAUCGCUAGUAAUCGUGGAUCAGCAUGCCACGGUGAAUACGUUC

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'
```

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 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

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).

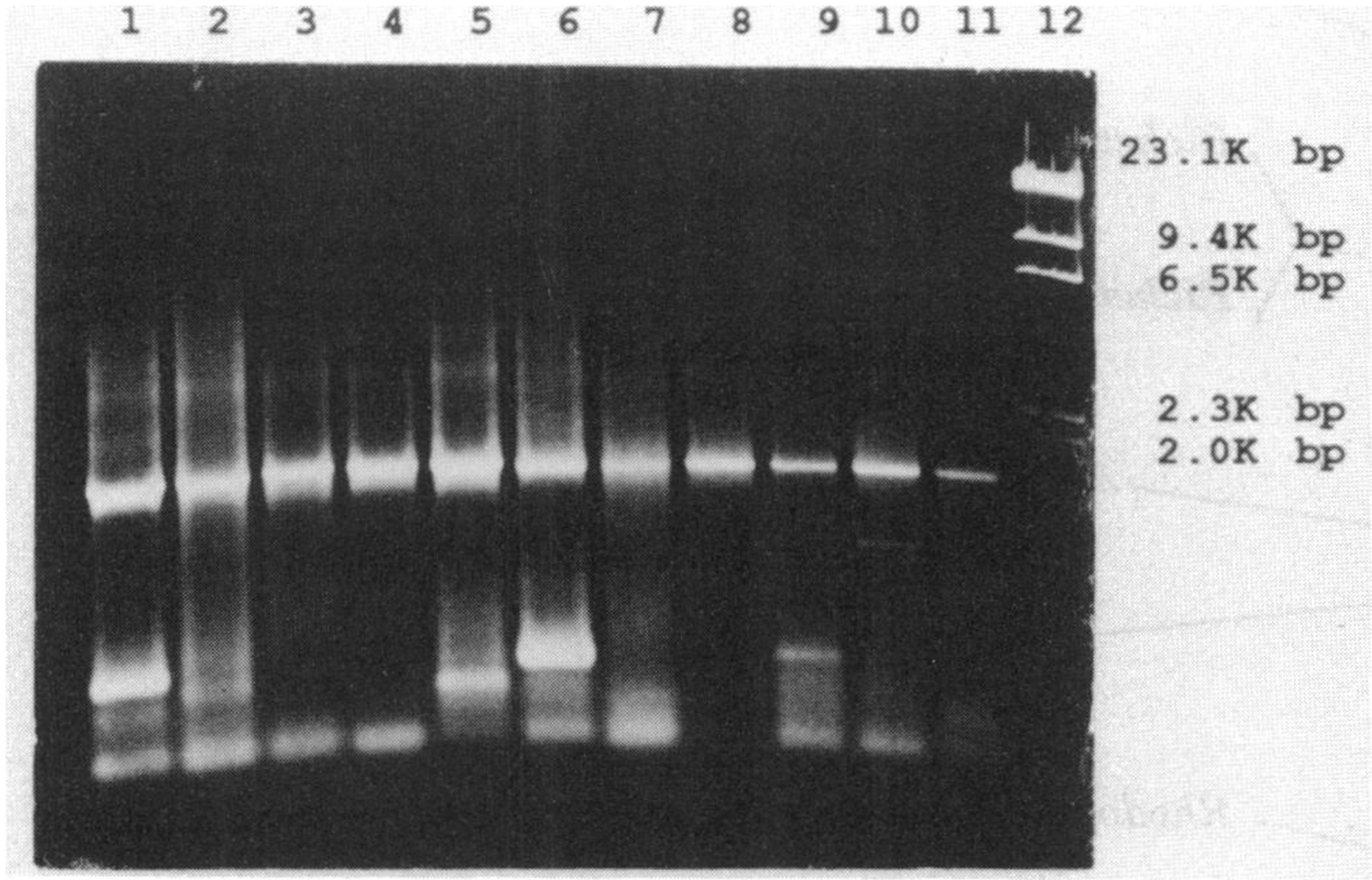


FIG. 2. Ethidium-bromide-stained 0.75% agarose gel displaying amplification products. Lanes: 1, *B. fragilis* DNA (fD1 + rP2); 2, *C. perfringens* DNA (fD1 + rD1); 3, *C. psittaci* DNA (fD4 + rD1); 4, *B. burgdorferi* DNA (fD3 + rD1); 5, lyophilized ampoule-derived *S. sanguis* DNA (fD1 + rP2); 6, lyophilized 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.

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.

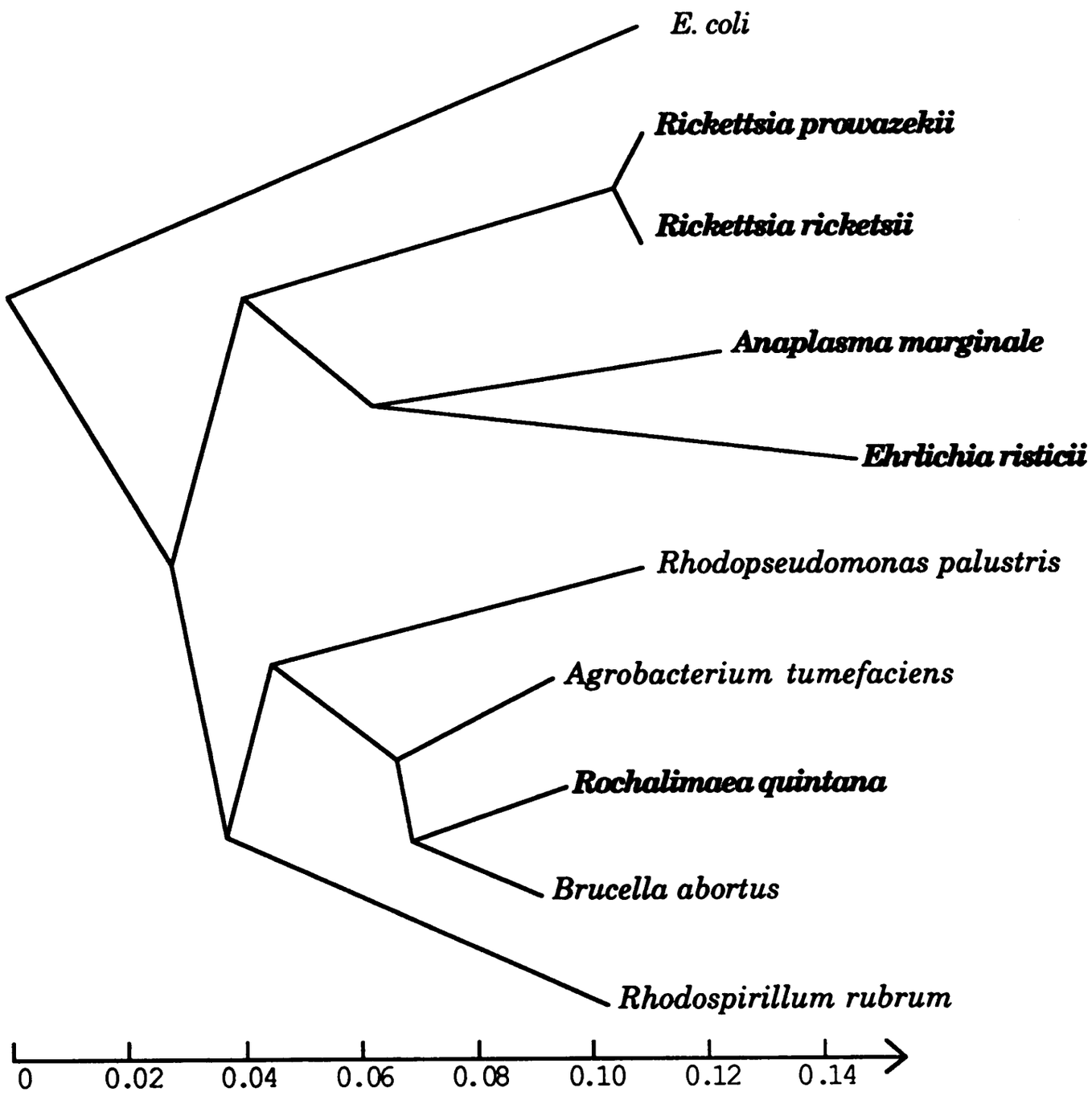


FIG. 3. Phylogenetic distance tree displaying the evolutionary origin of *A. marginale* within a lineage shared by the genera *Rickettsia* and *Ehrlichia*. All species belonging to the order *Rickettsiales* are shown in boldface type. *E. coli* is used as an outgroup sequence.

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.

JOURNAL OF BACTERIOLOGY, Sept. 1998, p. 4765–4774 Vol. 180, No. 18
0021-9193/98/$04.00+0

MINIREVIEW

Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity

PHILIP HUGENHOLTZ, BRETT M. GOEBEL,† AND NORMAN R. PACE*
Departments of Plant and Microbial Biology and Molecular and Cell Biology, University of California, Berkeley, Berkeley, California 94720-3102

Hugenholtz et al

INTRODUCTION

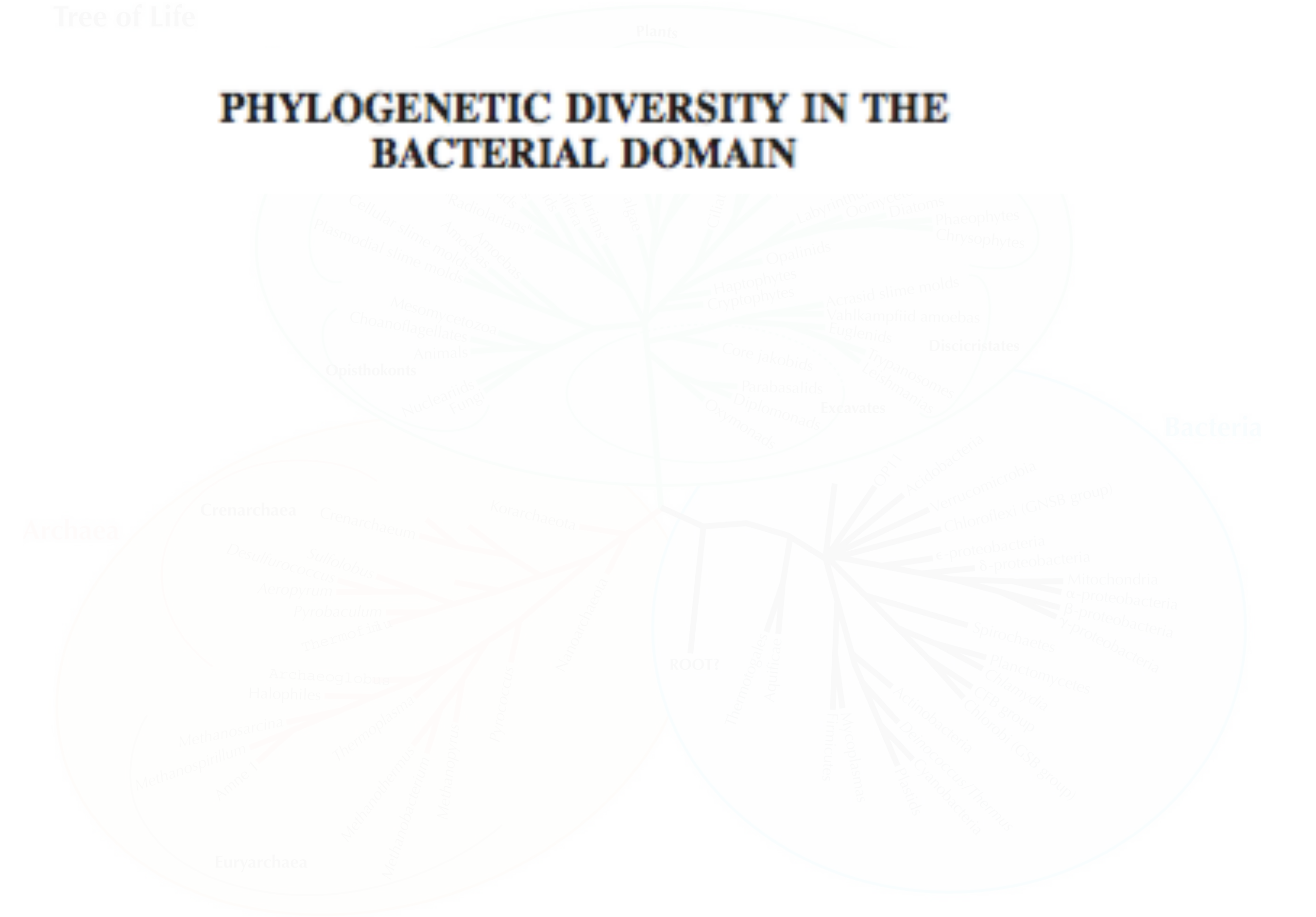
Our perspective on microbial diversity has improved enormously over the past few decades. In large part this has been due to molecular phylogenetic studies that objectively relate organisms. Phylogenetic trees based on gene sequences are maps with which to articulate the elusive concept of biodiversity. Thus, comparative analyses of small-subunit rRNA (16S or 18S rRNA) and other gene sequences show that life falls into three primary domains, *Bacteria*, *Eucarya*, and *Archaea* (51, 52). Based on rRNA trees, the main extent of Earth's biodiversity is microbial. Our knowledge of the extent and character of microbial diversity has been limited, however, by reliance on the study of cultivated microorganisms. It is estimated that >99% of microorganisms observable in nature typically are not cultivated by using standard techniques (1).

Recombinant DNA and molecular phylogenetic methods have recently provided means for identifying the types of organisms that occur in microbial communities without the need for cultivation (see references 1, 20, and 35 for reviews). Results from application of these methods to a number of diverse environments confirm that our view of microbial diversity was limited and point to a wealth of novel and environmentally important diversity yet to be studied (34). It is the aim of this review to collate, compare, and incorporate the results of the environmental sequence-based studies into the context of known bacterial diversity. We discuss the sequence data at the taxonomic level of the phylogenetic division because divisions constitute first-order clades for describing the breadth of bacterial diversity. Although we have yet to determine even the outlines of the bacterial tree, common threads are beginning to emerge that revise our current views of bacterial diversity and distribution in the environment.

Hugenholtz et al

- Although we have yet to determine even the outlines of the bacterial tree, common threads are beginning to emerge that revise our current views of bacterial diversity and distribution in the environment.

PHYLOGENETIC DIVERSITY IN THE BACTERIAL DOMAIN



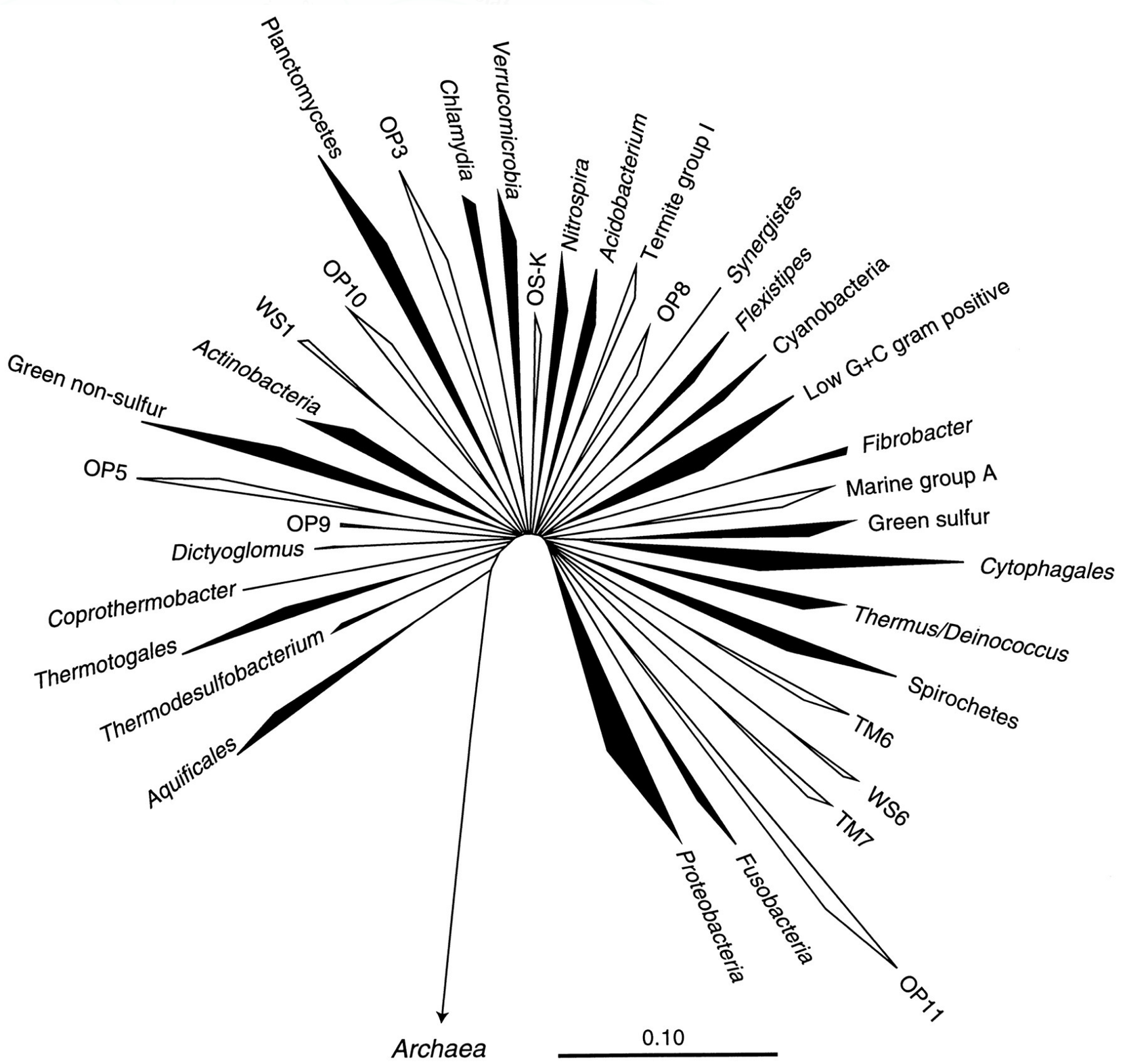
Hugenholtz et al

- These relatedness groups have variously been called “kingdoms,” “phyla,” and “divisions”; we use the latter term.
- For the purposes of this review we define a bacterial division purely on phylogenetic grounds as a lineage consisting of two or more 16S rRNA sequences that are reproducibly monophyletic and unaffiliated with all other division-level relatedness groups that constitute the bacterial domain
- We judge reproducibility by the use of multiple tree-building algorithms, bootstrap analysis, and varying the composition and size of data sets used for phylogenetic analyses.

- Division-level nomenclature has not even been consistent between studies, so some divi- sions are identified by more than one name. For instance, green sulfur bacteria is synonymous with Chlorobiaceae; high- G C gram-positive bacteria is synonymous with Actinobacteria and Actinomycetales. Indeed, it probably is premature to standardize taxonomic rankings for bacterial divisions at this point when our picture of microbial diversity is likely still incomplete and the topology of the bacterial tree is still unresolved.

- Figure 1 represents the division-level diversity of the bacterial domain as inferred from representatives of the approximately 8,000 bacterial 16S rRNA gene sequences currently available. Although 36 divisions are shown in Fig. 1, several other division-level lineages are indicated by single environmental sequences (9, 21, 37), suggesting that the number of bacterial divisions may be well over 40.

FIG. 1.
Evolutionary distance tree of the bacterial domain showing currently recognized divisions and putative (candidate) divisions. The tree was constructed using the ARB software package (with the Lane mask and Olsen rate-corrected neighbor-joining options) and a sequence database modified from the March 1997 ARB database release (**43**). Division-level groupings of two or more sequences are depicted as wedges. The depth of the wedge reflects the branching depth of the representatives selected for a particular division. Divisions which have cultivated representatives are shown in black; divisions represented only by environmental sequences are shown in outline. The scale bar indicates 0.1 change per nucleotide. The aligned, unmasked data sets used for this figure and Fig. **3** through **6**are available from **http://crab2.berkeley.edu/pacelab/176.htm**.



- Indeed, 13 of the 36 divisions shown in Fig. 1 are characterized only by environmental sequences (shown outlined) and so are termed “candidate divisions” new bacterial divisions
- One of these candidate divisions, OP11, is now sufficiently well represented by environmental sequences to conclude that it constitutes a major bacterial group (see below).
- Phylogenetic studies so far have not re- solved branching orders of the divisions; bacterial diversity is seen as a fan-like radiation of division-level groups (Fig. 1). The exception to this, however, is the Aquificales division, which branches most deeply in the bacterial tree in most analyses.

TABLE 1. Summary of 16S rRNA-based clonal analyses of diversity of uncultivated bacteria[a]

Habitat type	No. of studies	No. of sequences[c]	Bacterial division[b]: *Proteobacteria*[d] α[c]	β	γ	δ[c]	ε[c]	Cytophagales	*Actinobacteria*	Low-G+C gram positives[e]	*Acidobacterium*	*Verrucomicrobia*	Spirochetes	Nitrospira	CNS	OP11	Planctomycetes	Green sulfur	TM7	TM6	*Thermus*/*Deinococcus*	Cyanobacteria[f]	Synergistes	OP8	Termite group 1	OS-K	Chlamydia	OP3	OP10	WS1	OP5	Marine group A	*Fibrobacter*	*Flexibacter*	*Dictyoglomus*	*Thermotogales*	*Thermodesulfo-bacteriales*	*Aquificales*
Geothermal	10	212	○	○		○		○			○		○	○	○		○	○			○				○										○		○
Soil	16	743	●	○	○	○		○	●	○	●	●			○		○																				
Marine	23	687	●	○	●	○		○	○	○		○					○					○															
Freshwater	4	107	●	○					○		○																										
Wastewater	5	430	●	●	●	○		○	●	○		○			○		○				○																
Pollutant associated	7	202	○	○	○	○	○	○	○	○	○																										
Acid metal leaching	2	21	●	○	●			○	○	○	○			●																							
Subsurface	6	229	○	●	●	○	○		○	●				○	○	○																○					
Symbionts and commensals	10	280			●			○		○							○																				
Disease associated	3	7		○	○			○		○							○																				
Totals	86	2,938																																			

[a] An expanded version of this table detailing individual studies is available at http://crn.b2.berkeley.edu/~pacelab/136.htm.

[b] Incidence of division-level representatives in studies of particular habitat types ranked from most represented to least represented divisions: >75% (●), 25 to 75% (○), or <25% (no symbol) of studies have representatives of division.

[c] Excluding organelles.

[d] *Proteobacteria* are presented at the subdivision level due to the extensive sequence representation of this division.

[e] Cannot establish as a monophyletic group in all analyses.

ERRATUM

Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity

PHILIP HUGENHOLTZ, BRETT M. GOEBEL, AND NORMAN R. PACE

Departments of Plant and Microbial Biology and Molecular and Cell Biology, University of California, Berkeley, Berkeley, California 94720-3102

Volume 180, no. 18, p. 4765–4774, 1998. Page 4767, column 1: Table 1 should appear as shown below.

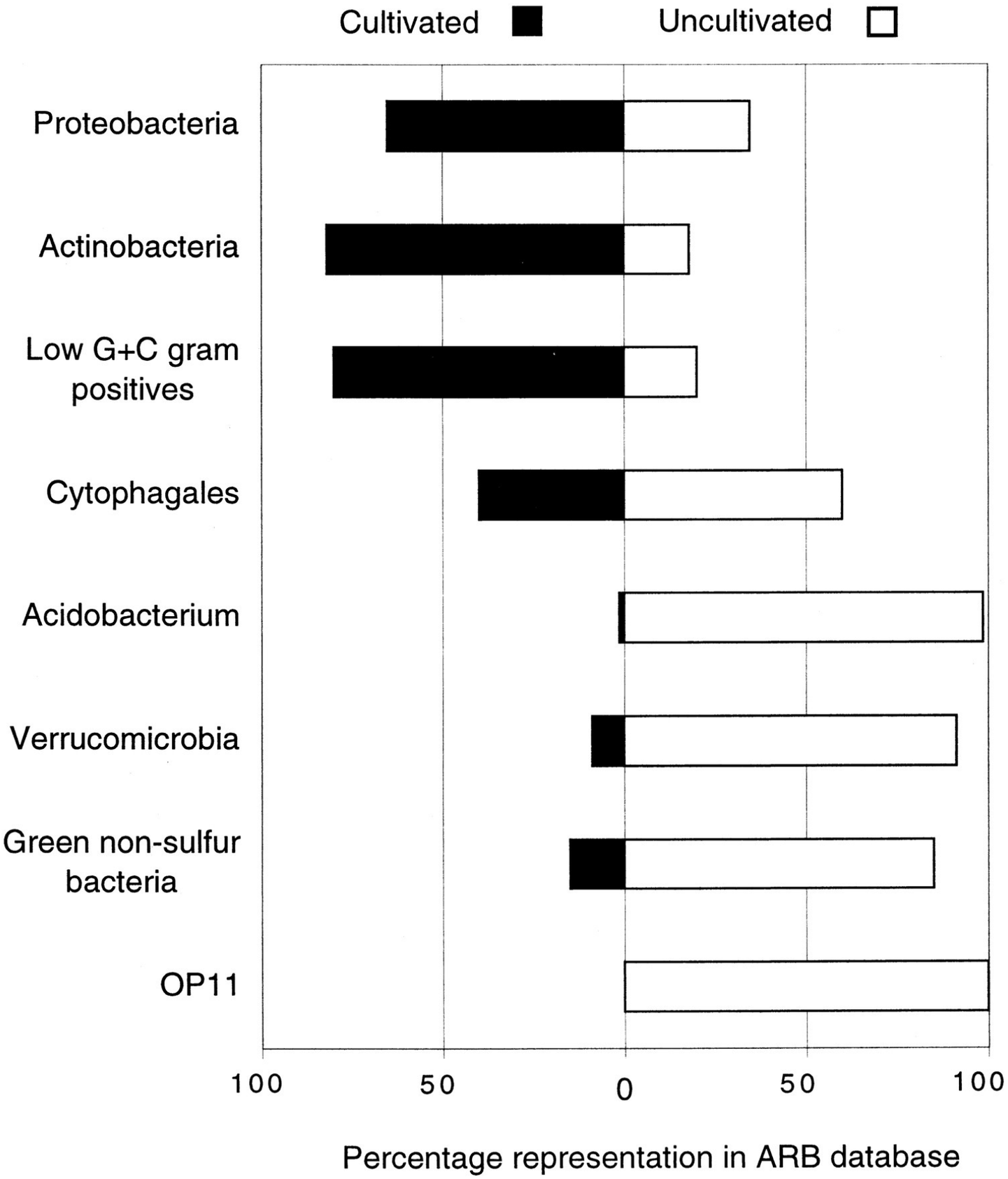
TABLE 1 Summary of 16S rRNA-based clonal analyses of diversity of uncultivated bacteria[a]

Habitat type	No. of studies	No. of sequences[c]	Bacterial divisions[b]: Proteobacteria[d] α[c]	β	γ	δ[d]	ε[e]	*Cytophagales*	*Actinobacteria*	Low-G+C gram positives[e]	*Acidobacterium*	*Verrucomicrobia*	Spirochetes	*Nitrospira*	GNS	OP11	*Planctomycetes*	Green sulfur	TM7	TM6	*Thermus/Deinococcus*	Cyanobacteria[c]	*Synergistes*	OP8	Termite group I	OS-K	*Chlamydia*	OP3	OP10	WS1	OP5	Marine group A	*Fibrobacter*	*Flexistipes*	*Dictyoglomus*	*Thermotogales*	*Thermodesulfobacterium*	*Aquificales*
Geothermal	10	212	○	○	√	○	√	○	√		○	√	○	○	○	√	○	○			○	√		√		○		√	√		√				√	○	√	○
Soil	16	743	●	○	○	○		○	●	○	●	●	√	√	○	√	○	√	√	√							√			√								
Marine	23	687	●	○	●	○	√	○	○	○	√	○			√	√	○	√	√			○				√						√						
Freshwater	4	107	●	○	√			√	○		○	√				√																						
Wastewater	5	430	●	●	●	○	√	○	●	○	√	○	√	√	○		○		√	√			√				√							√				
Pollutant associated	7	202	○	○	○	○	○	○	○	○	○	√	√	√	√	√								√	√				√	√	√							
Acid metal leaching	2	2	●	○	●			○	○	○	○			●																								
Subsurface	6	229	○	●	●	○	○	√	○	●			√	○	○	○	√	√			√							○										
Symbionts and commensals	10	280	√	√	●	√		○	√	○			○										√		√								√					
Disease associated	3	7		○	○			○		○			○																									
Totals	86	2,918																																				

[a] An expanded version of this table detailing individual studies is available at http://crab2.berkeley.edu/~pacelab/176.htm.
[b] Incidence of division-level representatives in studies of particular habitat types ranked from most represented to least represented divisions: >75% (●), 25 to 75% (○) or <25% (√) of studies have representatives of division. No symbol indicates division not detected.
[c] Excluding organelles.
[d] *Proteobacteria* are presented at the subdivision level due to the extensive sequence representation of this division.
[e] Cannot establish as a monophyletic gorup in all analyses.

FIG. 2.

Relative representation in selected cosmopolitan bacterial divisions of 16S rRNA sequences from cultivated and uncultivated organisms. Results were compiled from 5,224 and 2,918 sequences from cultivated and uncultivated organisms, respectively.



- The database of environmental rRNA sequences is compromised in resolving some phylogenetic issues by a large number of relatively short sequences. More than half of the sequences collated in Table 1 are less than 500 nucleotides (nt) long, which represents only one-third of the total length of 16S rRNA. This is due to an unfortunate trend in many environmental studies of sequencing only a portion of the gene in the belief that a few hundred bases of sequence data is sufficient for phylogenetic purposes. Indeed, 500 nt is sufficient for placement if some longer sequence is closely related (90% identity in homologous nucleotides) to the query sequence. In the case of novel sequences, 85% identical to known sequences, however, 500 nt is usually insufficient comparative information to place the sequence accurately in a phylogenetic tree and can even be misleading

Acidobacterium

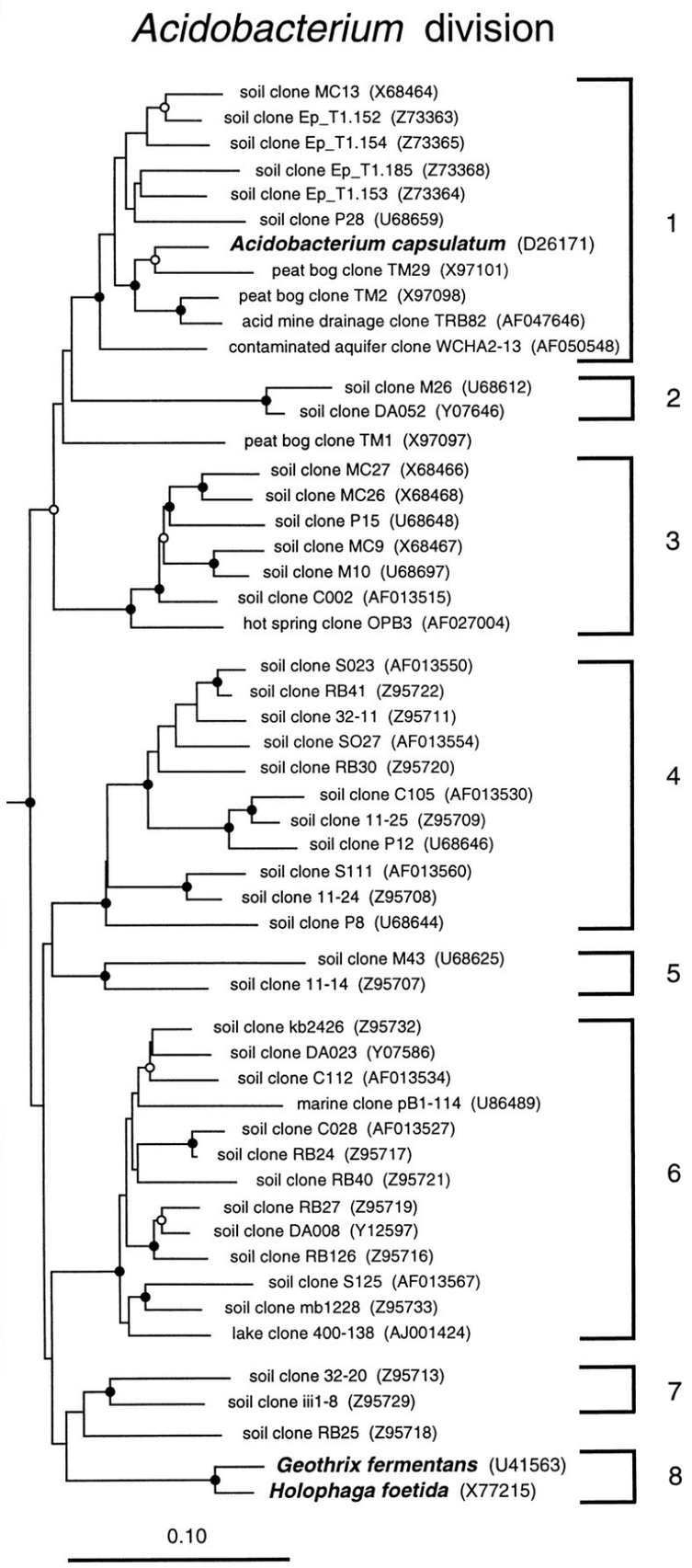


FIG. 3.
Phylogenetic dendrogram of the *Acidobacterium* division. Names of cultivated organisms are shown in bold. The habitat source of each environmental sequence is indicated before the clone name. GenBank accession numbers are listed parenthetically. Subdivisions (see the text) are indicated by brackets at the right of the tree. Construction of the tree was as described for Fig. 1. The robustness of the topology presented was estimated by bootstrap resampling of independent distance, parsimony, and rate-corrected maximum-likelihood analyses as previously described (2). Distance and parsimony analyses were conducted using test version 4.0d61 of PAUP*, written by David L. Swofford. Branch points supported (bootstrap values of >75%) by most or all phylogenetic analyses are indicated by filled circles; open circles indicate branch points marginally supported (bootstrap values of 50 to 74%) by most or all analyses. Branch points without circles are not resolved (bootstrap values of <50%) as specific groups in different analyses. The scale bar indicates 0.1 change per nucleotide.

Verrucomicrobia

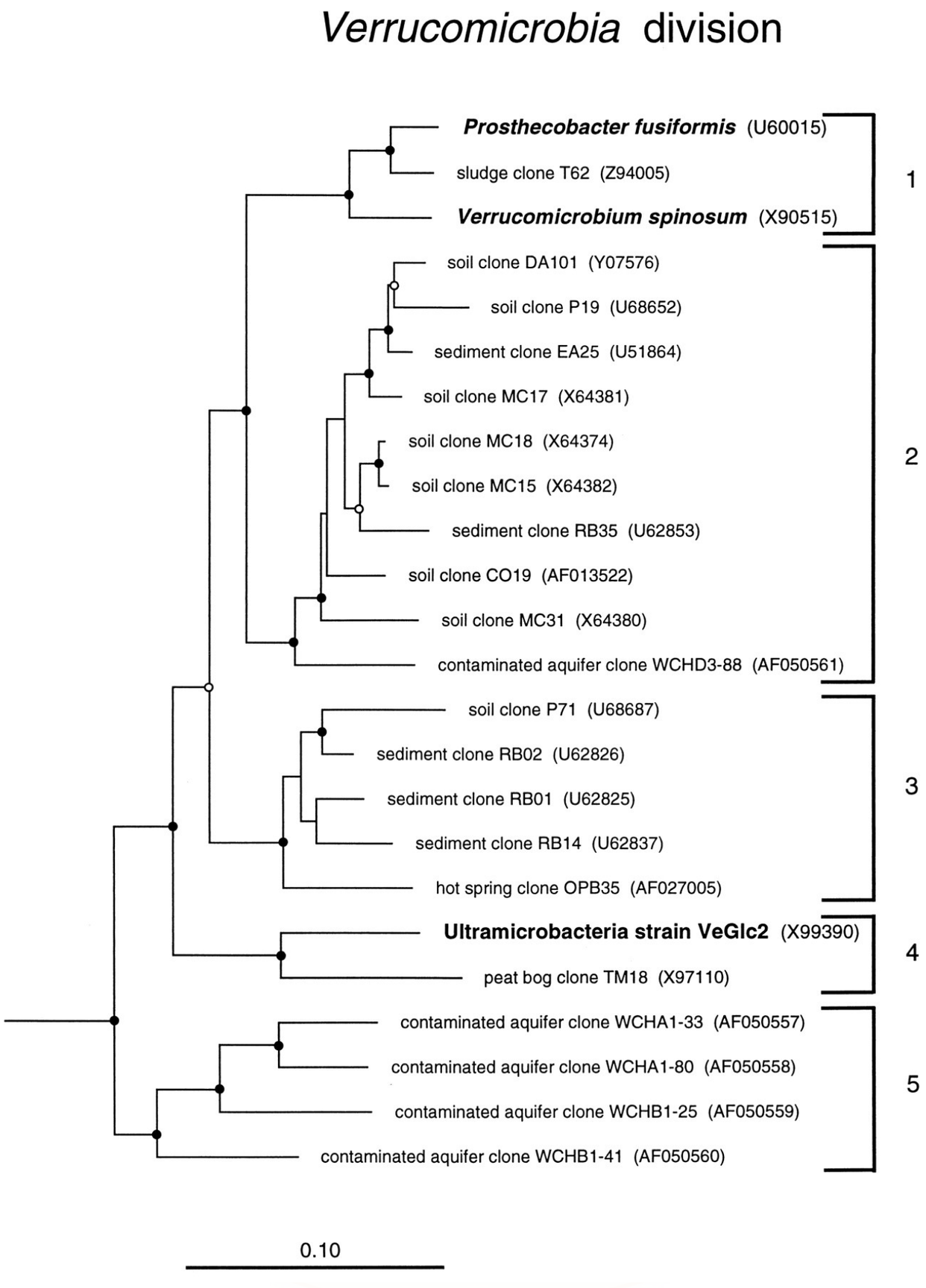


FIG. 4. Phylogenetic dendrogram of the *Verrucomicrobia* division. Names of cultivated organisms are shown in bold. The habitat source of each environmental sequence is indicated before the clone name. GenBank accession numbers are listed parenthetically. Subdivisions (see the text) are indicated by brackets at the right of the tree. Tree construction and support for branch points was as described for Fig. **1** and **3**, respectively. The scale bar indicates 0.1 change per nucleotide.

Green non sulfur

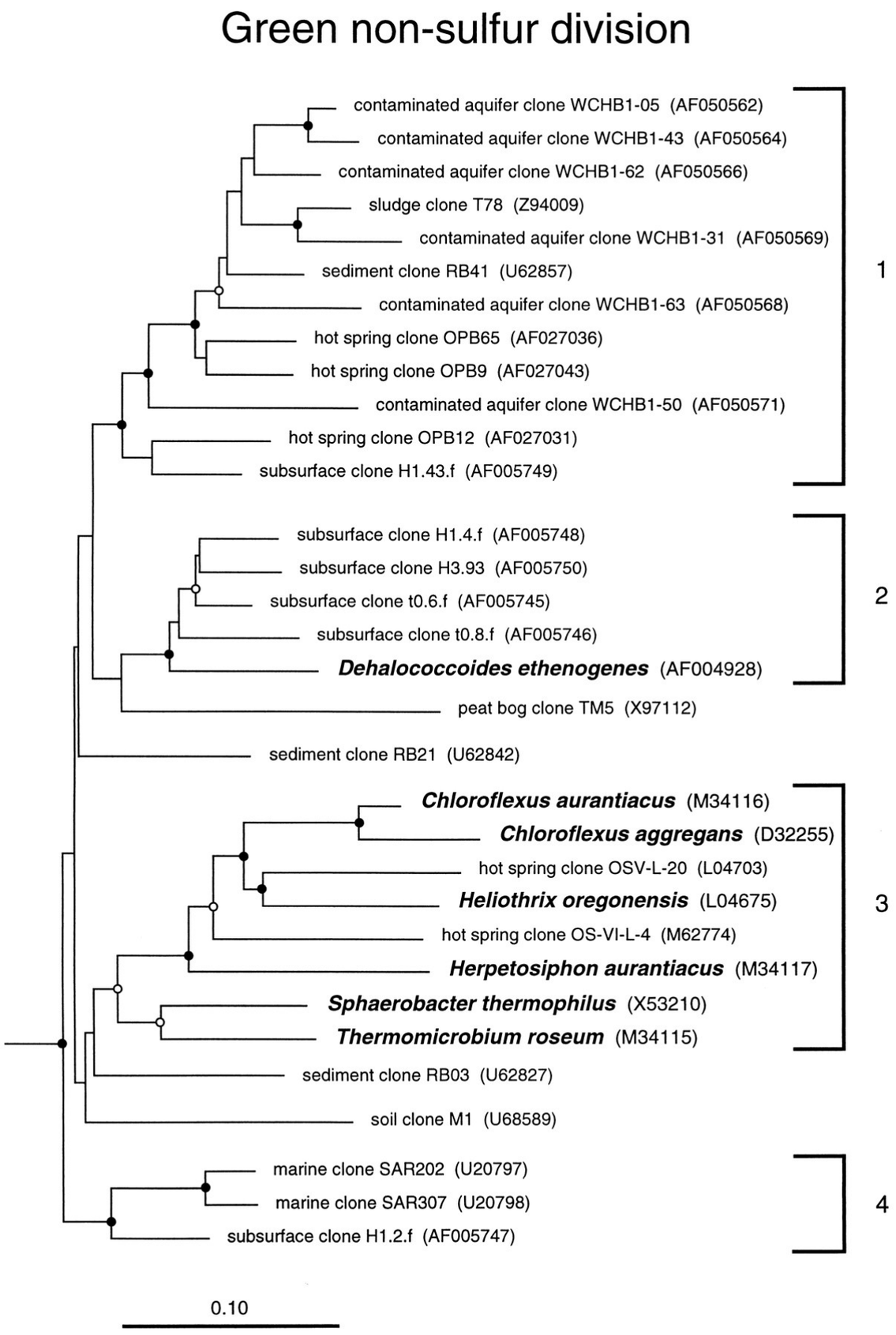


FIG. 5. Phylogenetic dendrogram of the GNS division. Names of cultivated organisms are shown in bold. The habitat source of each environmental sequence is indicated before the clone name. GenBank accession numbers are listed parenthetically. Subdivisions (see the text) are indicated by brackets at the right of the tree. Tree construction and support for branch points was as described for Fig. **1** and **3**, respectively. The scale bar indicates 0.1 change per nucleotide.

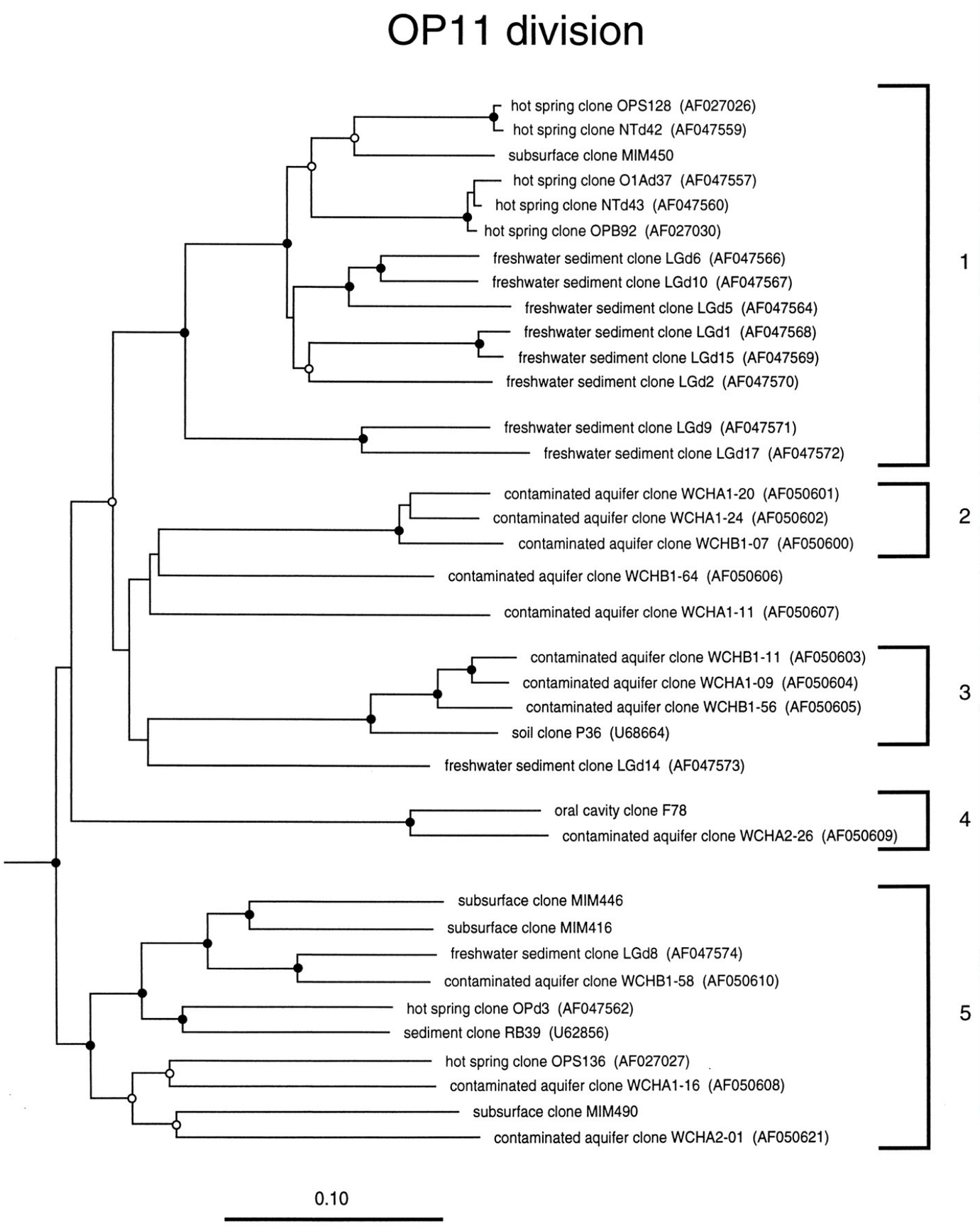


FIG. 6.
Phylogenetic dendrogram of the OP11 division. The habitat source of each environmental sequence is indicated before the clone name. GenBank accession numbers are listed parenthetically. Subdivisions (see the text) are indicated by brackets at the right of the tree. Tree construction and support for branch points was as described for Fig. **1** and **3**, respectively. The four MIM clones and F78 clone are unreleased sequences generously made available to us by Pascale Durand (**10**) and Floyd Dewhirst (**8**). The scale bar indicates 0.1 change per nucleotide.

Conclusions

- novelties are known as well, for instance, endospore formation by the low-G C gram-positive bacteria or axial filaments (endoflagella) in the spirochetes. Some biochemical properties evidently have transferred laterally among the divisions. For example, the two types of photosynthetic complexes, photosystem I (PSI) and PSII, are each distributed sporadically among the divisions, consistent with lateral transfer (3). Lateral transfer may also have resulted in combinatorial novelty among the divisions; PSI and PSII, for instance, apparently came together in the cyanobacteria to create oxygenic photosynthesis, with profound consequences to the biosphere (3).
- Many more such division-specific qualities and cooperations should become evident at the molecular level as comparative genomics gives us a sharper phylogenetic picture of bacterial diversity.

- PCR and microbial community surveys possible issues

- Where could this go “wrong”?