Host Distributions of Uncultivated Fecal *Bacteroidales* Bacteria Reveal Genetic Markers for Fecal Source Identification

Linda K. Dick,[1] Anne E. Bernhard,[1] Timothy J. Brodeur,[1] Jorge W. Santo Domingo,[2]
Joyce M. Simpson,[2] Sarah P. Walters,[1] and Katharine G. Field[1]*

*Department of Microbiology, Oregon State University, Corvallis, Oregon 97331,[1] and U.S. Environmental
Protection Agency, 26 West Martin Luther King Dr., Cincinnati, Ohio 45268[2]*

The purpose of this study was to examine host distribution patterns among fecal bacteria in the order *Bacteroidales*, with the goal of using endemic sequences as markers for fecal source identification in aquatic environments. We analyzed *Bacteroidales* 16S rRNA gene sequences from the feces of eight hosts: human, bovine, pig, horse, dog, cat, gull, and elk. Recovered sequences did not match database sequences, indicating high levels of uncultivated diversity. The analysis revealed both endemic and cosmopolitan distributions among the eight hosts. Ruminant, pig, and horse sequences tended to form host- or host group-specific clusters in a phylogenetic tree, while human, dog, cat, and gull sequences clustered together almost exclusively. Many of the human, dog, cat, and gull sequences fell within a large branch containing cultivated species from the genus *Bacteroides*. Most of the cultivated *Bacteroides* species had very close matches with multiple hosts and thus may not be useful targets for fecal source identification. A large branch containing cultivated members of the genus *Prevotella* included cloned sequences that were not closely related to cultivated *Prevotella* species. Most ruminant sequences formed clusters separate from the branches containing *Bacteroides* and *Prevotella* species. Host-specific sequences were identified for pigs and horses and were used to design PCR primers to identify pig and horse sources of fecal pollution in water. The primers successfully amplified fecal DNAs from their target hosts and did not amplify fecal DNAs from other species. Fecal bacteria endemic to the host species may result from evolution in different types of digestive systems.

The animal gastrointestinal (G.I.) tract maintains a rich microbial community in which many of the inhabitants are engaged in mutualistic association with one another and/or with the host (21, 37, 46). G.I. systems provide ideal models for investigating evolutionary relationships between hosts and their resident microbes. However, the microbial diversity and community structure of G.I. systems still are not well characterized. Analyses of 16S rRNA gene sequences from fecal DNAs suggest 60 to 80% of the human intestinal microflora remains uncultivated (20, 41). Studies of bacterial communities in other animal hosts reveal even greater uncultivated representation (8, 26, 32, 47). Molecular studies are beginning to identify differences in bacterial species composition in the G.I. tract dependent on diet and age as well as spatial, temporal, and host differences (14, 16, 17, 22, 27, 30).

Enteric bacteria may coevolve with their hosts (21, 37), and if so, host-specific clustering may be evident in phylogenetic analyses of gene sequences recovered from different species. Staley and Gosink (40) used the term "cosmopolitan" to describe bacteria with global distribution patterns, while endemism refers to the tendency to inhabit specific geographic regions or hosts. Hedlund and Staley (19) reported that *Simonsiella* strains isolated from the oral cavities of humans, cats, dogs, and sheep were endemic to the host species. However, bacteria in the G.I. tract often do not demonstrate host endemism. Multilocus enzyme electrophoresis analysis of

Escherichia coli isolates from 16 mammalian species in Australia demonstrated that only 6% of the genetic variation could be attributed to host species differences (16). The same study found host differences explained 20% of the variation in *Klebsiella pneumoniae*.

Patterns of microbial diversity in the G.I. tract have important implications for human and environmental health. These communities become a source of pathogens when released into aquatic environments as fecal pollution. Community members that exhibit host-specific distributions represent a valuable resource as potential markers for fecal pollution from specific sources.

Several studies have suggested that some species in the genus *Bacteroides* have host-specific distributions (2, 11, 25). Bernhard and Field used length heterogeneity PCR and terminal restriction fragment length polymorphism (T-RFLP) to identify human- and ruminant-specific 16S rRNA genetic markers for fecal anaerobes in the order *Bacteroidales* (3). They used the markers to design PCR primers for fecal source identification in aquatic environments (4). Characterization of additional markers should expand the applicability of this method to other animal hosts.

Members of the phylum *Bacteroidetes* (formerly *Cytophaga/Flavobacter/Bacteroides*) are phenotypically and ecologically diverse (5, 14). They are present in aquatic and terrestrial environments and are numerically important constituents of the animal oral cavity and intestinal microflora (8, 20, 26, 36, 41, 45). The cultivated fecal members of this phylum are in the order *Bacteroidales* and the genera *Bacteroides* and *Prevotella* (13).

* Corresponding author. Mailing address: Department of Microbiology, 220 Nash Hall, Oregon State University, Corvallis, OR 97331. Phone: (541) 737-1837. Fax: (541) 737-0496. E-mail: kate.field@orst.edu.
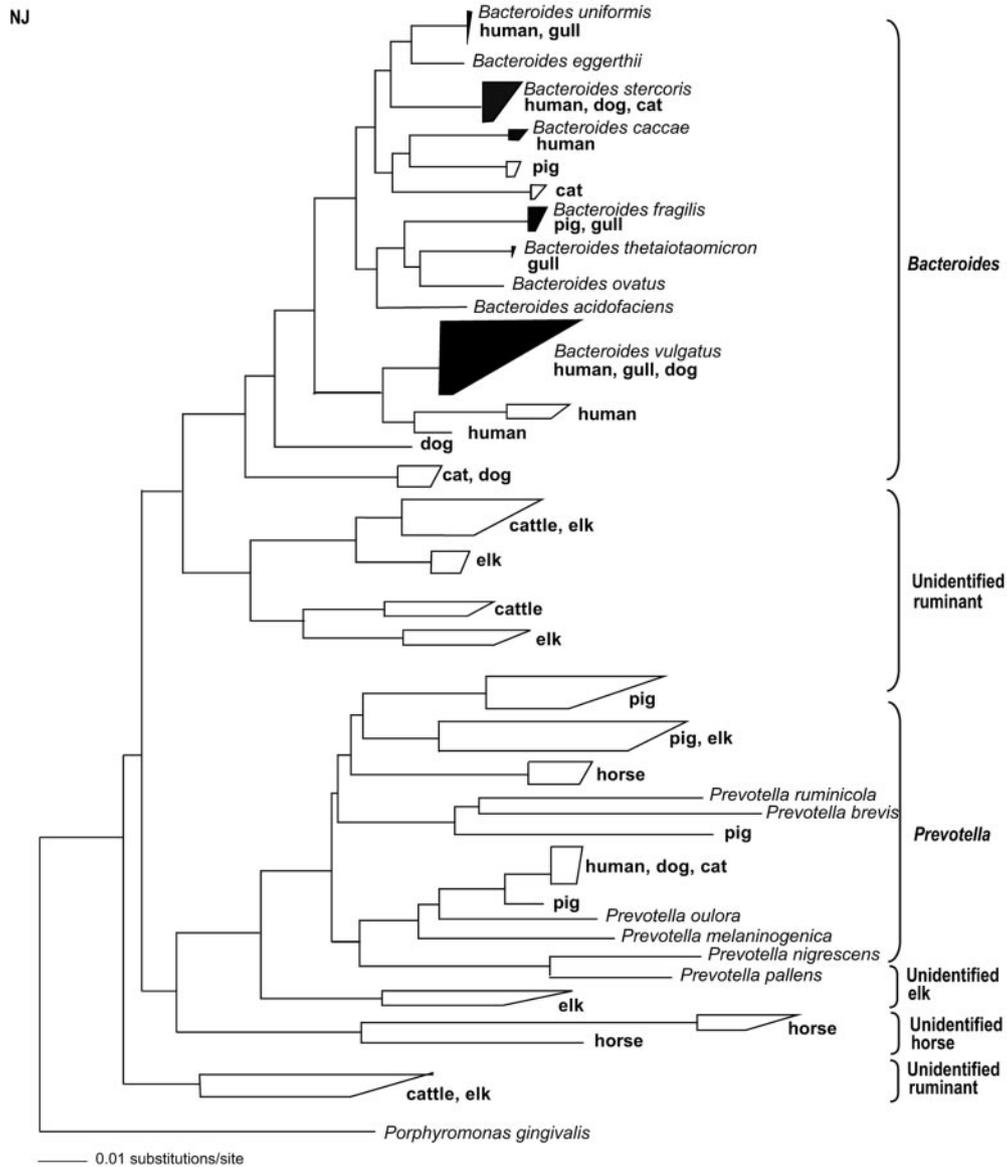
FIG. 1. Schematic representation of host distributions based on a neighbor-joining tree inferred from 16S rRNA gene sequences from *Bacteroidales* bacteria. Host clades containing cultivated species are depicted as solid wedges. Open wedges represent clusters with no cultivated representatives. *Prevotella brevis* and *Prevotella ruminicola* are of ruminant origin; all other cultivated species in the tree are of human origin. Sequences in bold, this study.
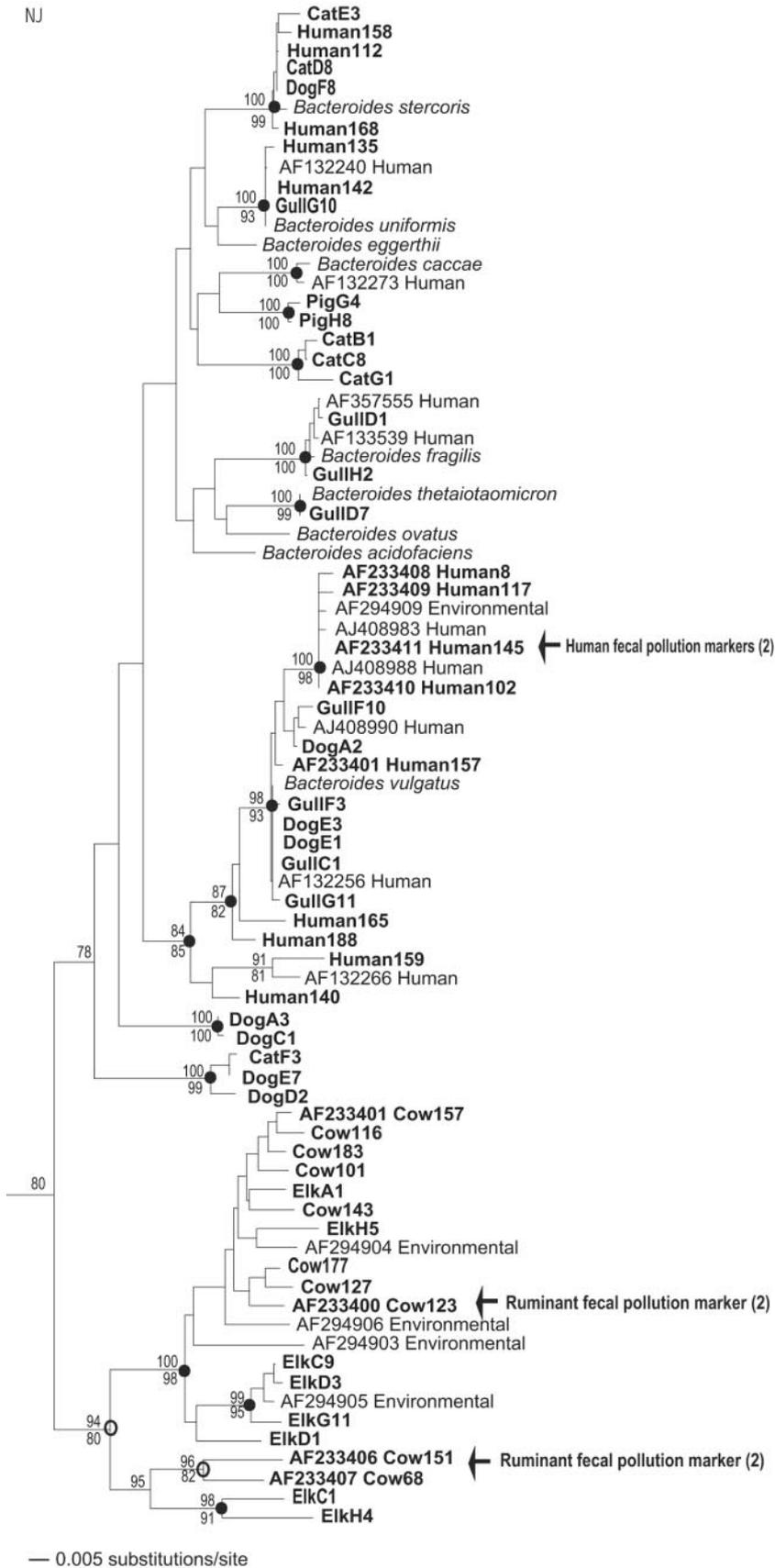
Culture and culture-independent methods have revealed that *Bacteroides* species account for 20 to 52% of the human fecal flora (9, 12, 20, 22, 38, 41). Analyses of 16S rRNA gene clone libraries suggest slightly less abundance in nonhuman hosts. Leser and colleagues found 11.2% of the phylotypes in a pig fecal clone library were related to *Bacteroides* or *Prevotella* (26). Representatives of the *Bacteroidales* accounted for 18% of the recovered sequences in a clone library of horse fecal DNAs (8). Analyses of feces-derived 16S rRNA gene sequences have also found large contributions from the *Bacteroidales* in cattle (3, 45).

The objectives of this study were to establish the extent to which host-specific distributions (endemism) occur in the *Bacteroidales*, thus providing potential new host-specific markers

for fecal pollution. Understanding the evolutionary relationships and host distributions of enteric microorganisms is critical to evaluating their practical application as fecal source indicators in aquatic ecosystems. We report the results of our 16S rRNA gene sequence analysis of *Bacteroidales* from the feces of eight hosts: human, bovine, elk, pig, dog, cat, gull, and horse.

MATERIALS AND METHODS

Sample collection and DNA extraction. We collected individual fecal samples from 20 dogs, 20 cats, 10 elk, 20 pigs, and 10 gulls. Samples came from Oregon State University animal research facilities, local animal shelters, farms, pet owners, hunters, and colleagues. Ten samples each were collected from human and bovine sources in a previous study (3), and sequences from that analysis were

NJ

CatE3
Human158
Human112
CatD8
DogF8
100 Bacteroides stercoris
99 Human168
Human135
AF132240 Human
Human142
100 GullG10
93 Bacteroides uniformis
Bacteroides eggerthii
100 Bacteroides caccae
100 AF132273 Human
100 PigG4
100 PigH8
CatB1
100 CatC8
100 CatG1
AF357555 Human
GullD1
AF133539 Human
100 Bacteroides fragilis
100 GullH2
100 Bacteroides thetaiotaomicron
99 GullD7
Bacteroides ovatus
Bacteroides acidofaciens
AF233408 Human8
AF233409 Human117
AF294909 Environmental
AJ408983 Human
AF233411 Human145   ← Human fecal pollution markers (2)
100 AJ408988 Human
98 AF233410 Human102
GullF10
AJ408990 Human
DogA2
AF233401 Human157
Bacteroides vulgatus
98 GullF3
93 DogE3
DogE1
GullC1
AF132256 Human
87 GullG11
82 Human165
84 Human188
85 91 Human159
78 81 AF132266 Human
Human140
100 DogA3
100 DogC1
CatF3
100 DogE7
99 DogD2
AF233401 Cow157
Cow116
Cow183
Cow101
ElkA1
Cow143
ElkH5
80 AF294904 Environmental
Cow177
Cow127
AF233400 Cow123   ← Ruminant fecal pollution marker (2)
AF294906 Environmental
AF294903 Environmental
100 ElkC9
98 ElkD3
99 AF294905 Environmental
95 ElkG11
ElkD1
94 96 AF233406 Cow151   ← Ruminant fecal pollution marker (2)
80 82 AF233407 Cow68
95 ElkC1
98 ElkH4
91

— 0.005 substitutions/site

used in this study. Horse sequences were recovered from a manure pile beside a creek in Cincinnatti, Ohio, also in a previous study (39). Samples were collected in sterile containers and stored at −80°C. Approximately 300 mg feces (wet weight) from each sample were used in separate DNA extractions. The FastDNA kit for Soils (Q-Biogene, Carlsbad, CA) was used by following the manufacturer's protocol, with an additional wash using the SEWS-M reagent to reduce PCR inhibition due to phenolic compounds in feces.

PCR and clone library construction. We amplified 700-bp partial 16S rRNA gene sequences from individual fecal extracts with the *Bacteroidales*-specific PCR primers Bac32F (AACGCTAGCTACAGGCTT) (4) and Bac708R (CAATCG GAGTTCTTCGTG) (4). Each 50-μl PCR mixture contained 1× *Taq* polymerase buffer, 10 μM concentrations of each primer, 200 μM concentrations of each deoxynucleoside triphosphate, 1.25 U of *Taq* polymerase, 0.06% bovine serum albumin, and 1.5 mM MgCl$_2$. Reactions were carried out for 30 cycles of 95°C for 1 min, 53°C for 45 s, and 72°C for 1 min. To obtain equal representation from individual hosts, separate PCR products from each sample were pooled in equal amounts based on ImageQuant analysis (Molecular Dynamics, Inc., Sunnyvale, CA). The pooled, host-specific amplicons were gel purified with the QiaQuick gel purification kit (QIAGEN, Inc., Valencia, CA) and cloned into TOPO TA vectors (Invitrogen, Inc., Carlsbad, CA) according to the manufacturer's protocol. The vectors were transformed into competent *E. coli* cells (One Shot TOP 10; Invitrogen, Corp., Carlsbad, CA). Ninety-six transformants were randomly selected from each host-specific library, inoculated into 100 μl LB-ampicillin broth in 96-well culture plates, and incubated overnight at 37°C. Replica plates were made from each original plate.

T-RFLP analysis. DNAs from the replica plates were amplified using 6-FAM-labeled Bac32F and Bac708R. Ninety-six clones from each host species library were digested with the restriction enzyme HaeIII, chosen based on previous work (3). The products (20 fmols) were separated electrophoretically on an ABI 377 automated sequencer (PE Applied Biosystems, Foster City, CA), and fragment sizes were estimated using GeneScan software (PE Applied Biosystems). Unique restriction patterns within each host-specific library were identified, and clones representing each pattern were chosen for sequencing. Duplicate clones were sequenced for dominant restriction patterns. Sequencing was bidirectional with the T7 promoter and M13R primers (Davis Sequencing, Davis, CA).

Phylogenetic analysis. Elk, pig, dog, cat, and gull sequences were aligned with human, bovine, and horse sequences from previous studies (3, 4, 39, 45) using the ARB software program (28). The shorter, cloned sequences were added to a tree of full-length sequences using the parsimony insertion tool in ARB. We removed potential chimeras based on both the CHECK_CHIMERA program (7) and on trees inferred from each half of the sequence data (24). Sequences with different placement in the trees made from the two halves were removed from the analysis. Phylogenetic analysis used PAUP∗ 4.0 beta 6 (42), with ambiguous regions of alignment removed. Trees were inferred from 526 sequence positions using three tree-building methods: neighbor joining with a Kimura-2 parameter correction, maximum parsimony with a heuristic search, and maximum likelihood with a heuristic search. Bootstrap values were obtained from a consensus of 1,000 neighbor-joining and parsimony trees.

Marker identification and PCR primer design. Host-specific sequences were identified for pig and horse samples, and PCR primers were designed using the Primer Design and Probe Match functions in the ARB software program.

Primer specificity and sensitivity trials. Primers were tested for cross-reactivity against host pools of fecal DNAs representing 10 to 30 individual hosts from each species, with the exception of sheep (5 hosts) and deer (4 hosts). The pools were normalized to 3 ng/μl by Picogreen assay (Molecular Probes, Inc., Eugene, OR). Primer specificity was optimized by manipulation of annealing temperature, Mg^{2+} concentration, and cycle number. Marker distributions within target host species were tested using pig and horse feces from hosts not used in the clone library analysis and from different geographic regions.

The sensitivity of each primer set was estimated using serial dilutions of plasmid DNAs of known copy number. A theoretical detection limit was determined for the primer sets in pure water, creek water (Beaver Creek, Alsea, OR), and seawater (5

miles off the central Oregon coast). Three nanograms of total genomic DNA from one of the two natural water sources was added to each PCR mixture.

Nucleotide sequence accession numbers. The partial 16S rRNA gene sequences have been deposited in the GenBank database under accession numbers AY695667 to AY695716 and AY859646 to AY859674.

RESULTS

Restriction and sequence analysis. We distinguished between 10 and 20 restriction patterns in each of the host clone libraries analyzed with T-RFLP. Ten to 24 clones per host species were selected for sequencing. Thirty-seven sequences (about 28%) were identified as potential chimeras and removed from the analysis. The final analysis contained 97 cloned sequences from the eight host species. There was a high degree of branching order agreement among the 3 tree-building methods, although bootstrap support varied.

Host distributions. Figure 1 is a diagrammatic representation of the host distributions from a phylogenetic analysis of the sequences. Members of the *Bacteroidales* were not strictly monophyletic with respect to host species; however, several host-specific clades could be identified. Ruminant, pig, and horse bacterial sequences tended to form unique clusters apart from other hosts, while human, dog, cat, and gull bacterial sequences clustered together almost exclusively.

***Bacteroides*-related sequences.** A branch of the tree containing cultivated species from the genus *Bacteroides* was supported in 78% of the neighbor-joining bootstrap resamplings (Fig. 2). It included all the fecal *Bacteroides* species identified in a phylogenetic analysis of the genera *Bacteroides*, *Prevotella*, and *Porphyromonas* by Paster and colleagues (31). Within this branch, a major well-supported gene cluster contained *Bacteroides vulgatus* and fecal sequences from human, dog, and gull sources. The remaining fecal *Bacteroides* species and fecal sequences from human, cat, dog, pig, and gull sources clustered together with low bootstrap support.

Many of the human, dog, cat, and gull bacterial sequences were closely related to known *Bacteroides* species (97 to 99% identity). *B. vulgatus*, *B. uniformis*, and *B. stercoris* clusters included fecal sequences from two or more of these four host species. Some of the cloned sequences from different hosts appeared identical or almost identical to each other and to sequences from cultivated species. For example, a gull sequence was >99% identical to *B. thetaiotaomicron*; 694 of 695 bases matched the published *B. thetaiotaomicron* sequence. Similarly, cat, dog, and human bacterial sequences were 98% identical to *B. stercoris*, and gull, dog, and human sequences were 99% identical to *B. vulgatus*. Since sequence artifacts may be introduced during PCR due to *Taq* polymerase errors, these small differences may not be significant.

Two pig bacterial sequences also clustered with the genus

FIG. 2. Phylogenetic relationships of partial *Bacteroidales* 16S rRNA gene sequences (526 positions) with cultivated members of the genus *Bacteroides*. Sequences from cultivated species are shown in italics. Environmental sequences were recovered from Tillamook Bay on the Oregon coast (3). Trees were inferred using three tree-building algorithms: neighbor joining with a Kimura-2 parameter correction, maximum parsimony, and maximum likelihood. Values at the nodes were obtained by bootstrap analysis based on 1,000 resamplings of both neighbor-joining and parsimony trees (above and below the line, respectively). Bootstrap values over 70% are shown. The closed and open circles at the nodes represent branching orders observed in all treeing methods and those observed in two of the three methods, respectively. The scale bar represents 0.5% estimated sequence divergence. Sequences in bold, this study.
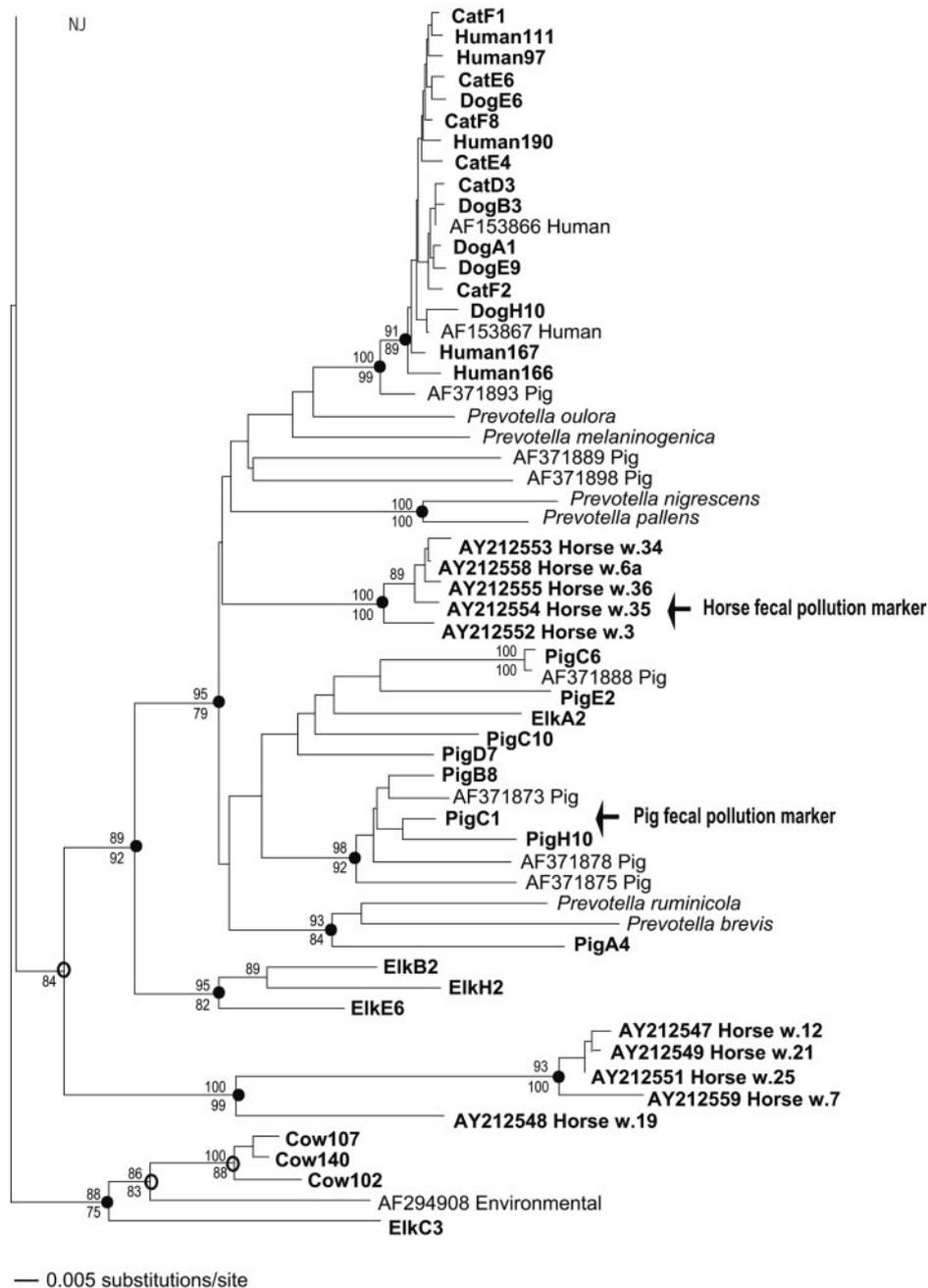
FIG. 3. Phylogenetic relationships of partial 16S rRNA gene sequences from fecal members of the *Bacteroidales* with cultivated members of the genus *Prevotella*. See the legend to Fig. 2 for explanation.

Bacteroides but had only 93% sequence identity with the closest cultivated species.

Ruminant sequences. We recovered no ruminant bacterial sequences in the branch that included the cultivated *Bacteroides* species. Two phylogenetic studies of the rumen also did not recover sequences from the genus *Bacteroides* (33, 44); however, isolates very closely related to *B. thetaiotaomicron* and *B. uniformis* have been recovered from sheep rumen (GenBank accession numbers AF139524 and AF139525) (personal communication, K. Gregg, Rumen Biotech, Murdoch University, Murdoch, Australia). Inter-

estingly, one of the sheep rumen isolates was 100% identical to a gull sequence in our library.

Most ruminant (cow and elk) bacterial sequences clustered together, separate from the sequences from other hosts, and they were distantly related to cultivated relatives (87 to 91%). The largest ruminant cluster formed a sister group to the branch containing the genus *Bacteroides* (Fig. 2). A second small cluster, containing only sequences from elk, formed a sister group to the branch containing the genus *Prevotella* (Fig. 3). The placement of a third small cluster containing elk and cow sequences could not be resolved (Fig. 3).
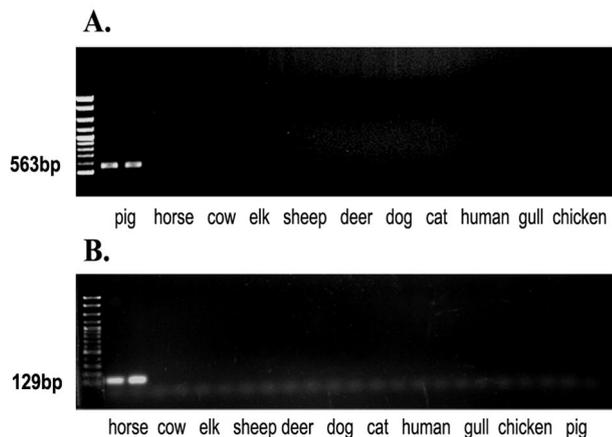
A.

B.

FIG. 4. Host DNA pools amplified with host-specific *Bacteroidales* PCR primers. Pools consisted of 4 to 30 individual host fecal DNAs, with each pool normalized to 3 ng/μl total DNA (Picogreen assay). (A) Primer PF163F distinguished pig fecal DNA from other host sources. (B) Primer HoF597F distinguished horse fecal DNA from other host sources.

Prevotella-related sequences. A branch containing sequences from the genus *Prevotella* was supported in 95% of the boot-strap resamplings (Fig. 3). None of the cloned sequences was closely related to any cultivated *Prevotella* species. A tight gene cluster of human, cat, and dog bacterial sequences exhibited 97 to 99% intraclade identity and an average 92% identity with the oral bacterium *Prevotella oulorum*. The range of pig sequence identities with the closest cultivated species was 88 to 93%. All but one of the pig bacterial sequences were most closely related to oral *Prevotella* species.

Horse bacterial sequences produced two unique clusters, one within the cluster containing *Prevotella* species and the second related to it.

Design of new host-specific markers. Host-specific sequences were identified from the 16S rRNA gene sequences for pig and horse fecal *Bacteroidales* bacteria. They were used to design PCR primers to identify pig (PF163F, GCGGATTA ATACCGTATGA) and horse (HoF597F, CCAGCCGTAAA ATAGTCGG) sources of fecal pollution in water. Both primers were paired with the *Bacteroidales*-specific reverse primer Bac708R (4). The primers were highly specific using fecal DNA pools from target and nontarget host species (Fig. 4). The new primers also amplified fecal DNAs from pig and horse sources not used to construct the clone libraries and from geographically distant sources. The host-specific markers were present in all 19 individual pig samples and 9 of 10 individual horse fecal samples tested (data not shown).

The theoretical limit of detection for both HoF597 and PF163 was 100 copies, and detection was not reduced by the presence of creek water or seawater DNA in PCR mixtures (Fig. 5).

DISCUSSION

A notable feature of this and other clone library analyses from G.I. environments is the number of sequences that do not cluster with any known species. Sequences from nonruminant
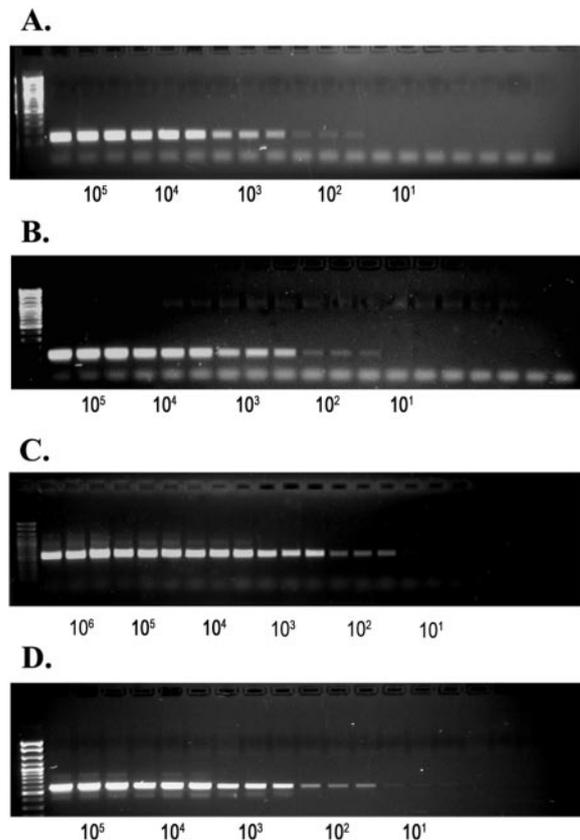
A.

B.

C.

D.

FIG. 5. Theoretical sensitivity (detectable copy number) of host-specific primer sets using serial dilutions of template inserted into plasmids. One hundred template copies were detected with primer HoF597F in pure water (A) and in creek water (B). Primer PF163F detected 100 copies in pure water (C) and in seawater (D). Three nanograms of total DNA from seawater or creek water extracts was added to each reaction mixture.

feces fell within the genera *Bacteroides* and *Prevotella*, while most sequences from ruminant hosts did not cluster with any known species. Both the range of sequence identity with the closest known species (87 to 91%) and the interclade identity range (81 to 94%) suggest that taxonomic diversity exists among ruminant sequences. Representatives of these groups must be isolated before phenotypic characterization or classification can be determined.

Within the genus *Bacteroides*, cloned sequences were closely related to known species, reflecting the greater number of cultivated representatives in this genus. This in turn reflects a greater emphasis placed on the study of the human fecal flora relative to other animal hosts. Many cultivated *Bacteroides* species, particularly *B. vulgatus*, *B. uniformis*, *B. thetaiotaomicron*, and *B. stercoris*, may not provide useful targets for fecal source discrimination assays because of the many similar or identical sequences from nonhuman hosts.

Phylogenetic resolution may be limited by the use of partial 16S rRNA gene sequences in a comparative analysis (23). However, since the primers used here were designed for aquatic fecal source identification, they were constrained by the need to exclude amplification of closely related aquatic bacteria such as *Cytophaga* spp. (3). Bootstrap values com-

pared favorably with those in the earlier study by Paster and colleagues (31), in which full-length sequences were used.

Twenty-eight percent of the cloned sequences were identified as potential chimeras, fewer than the 32% identified by Wang and Wang (43) in a study analyzing chimera frequency in 16S rRNA gene sequences from a mixed genomic population. They used PCR cycling parameters similar to those used in this study (30 cycles). Robison-Cox and colleagues (35) proposed comparing trees inferred from half sequences to identify chimeric sequences. This method and the CHECK_CHIMERA program reliably identify chimeras formed between distantly related sequences (23, 35); we eliminated such chimeras from our analyses. Any chimeras remaining in our analyses were formed between sequences too closely related to perturb placement of the chimera halves in trees; these would not affect our conclusions on host distributions. In addition, the probability of formation of identical chimeras in different gene libraries is low (10). Thus, recovery of clusters containing closely related or identical sequences from different libraries and hosts provides evidence that these clusters are not of chimeric origin.

Several evolutionary hypotheses could explain the existence of both cosmopolitan and endemic host distributions of fecal bacteria. Endemic distributions could occur among host species with limited physical contact and, thus, no horizontal transmission of fecal bacteria. Several studies have provided evidence for endemism when geographic barriers inhibit dispersal of bacterial populations (18, 40). If fecal bacteria within a host species had diversified in the time since host species diverged, each host species might contain unique types. In this case, we would expect to see types endemic to distantly related hosts such as humans and gulls. Instead, we see that humans and gulls share closely related *Bacteroidales* sequences.

A striking example of endemic distribution of fecal *Bacteroidales* is provided by ruminants, which share unusual clades and do not have the common groups shared by other host species. The unique ruminant digestive system may provide a different way for these organisms to make a living than those inhabiting nonruminant hosts. Populations of fecal bacteria endemic to host species may be the result of evolution in different types of digestive systems (29). Alternatively, the evolution of different G.I. systems may have been influenced by the microbial populations themselves (34, 37).

Cosmopolitan distributions could occur with frequent horizontal transmission of fecal bacteria among hosts with similar digestive systems. Sequences from multiple hosts were nearly identical to *B. vulgatus*, *B. thetaiotaomicron*, *B. uniformis*, *B. fragilis*, and *B. stercoris*, suggesting that these *Bacteroides* species are cosmopolitan with respect to host species. The "bushy tips" at branch termini within the genus *Bacteroides* also suggest a cosmopolitan population structure, in which adaptations to new host environments represent species "ecotypes" (1, 6, 15). Humans share proximity with domestic pets, making frequent occurrence of horizontal transmission of fecal bacteria likely. Gulls inhabit beaches, picnic areas, and landfills, where contact with human and domestic pet excrement may occur.

Prevotella-related clones exhibited more varied host distribution patterns than *Bacteroides*-related clones. A cluster of human and domestic pet *Prevotella* sequences was not closely related to any known species but again showed a cosmopolitan host distribution. *Prevotella*-related pig sequences clustered together, separate from other host species, as did the horse sequences, suggesting a more endemic distribution pattern. Pig-only and horse-only clades were used to identify unique sequences for these hosts, and PCR primers designed from these clades did not amplify DNA from nontarget hosts. However, this approach was not always successful, most likely because coverage of clone libraries was low. For example, a primer designed from a clade that appeared to be elk-specific amplified cattle fecal sequences (data not shown), probably because cattle clone library representation was lacking in that group. The overlap in domestic pet and human sequences also precluded the development of a unique dog or cat marker. The representation issue when using clone library sequences has led us to look for supplemental methods, including subtractive hybridization, to identify host-specific fecal markers in new hosts (8a).

The 16S rRNA gene sequence analysis of fecal *Bacteroidales* revealed both endemic and cosmopolitan distributions among the eight hosts used in the study. The evolution of host-enteric bacterial interactions is complex and will certainly not be understood in the context of a single gene. Multiple interactions and sources of gene flow result in evolutionary trajectories that may not be predictable. However, the evolutionary consequences of these host distribution patterns have potential for practical application in the field of fecal source identification.

ACKNOWLEDGMENTS

REFERENCES

1. **Acinas, S. G., V. Klepac-Ceraj, D. E. Hunt, C. Pharino, I. Ceraj, D. L. Distel, and M. F. Polz.** 2004. Fine-scale phylogenetic architecture of a complex bacterial community. Nature **430:**551–554.
2. **Allsop, K., and J. D. Stickler.** 1985. An assessment of *Bacteroides fragilis* group organisms as indicators of human faecal pollution. J. Appl. Bacteriol. **58:**95–99.
3. **Bernhard, A. E., and K. G. Field.** 2000. Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from fecal anaerobes. Appl. Environ. Microbiol. **66:**1587–1594.
4. **Bernhard, A. E., and K. G. Field.** 2000. A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Appl. Environ. Microbiol. **66:**4571–4574.
5. **Boone, D. R., and R. W. Castenholz (ed.).** 2001. Bergey's manual of systematic bacteriology, 2nd ed., vol. 1. Springer-Verlag, New York, N.Y.
6. **Cohan, F. M.** 2001. Bacterial species and speciation. Syst. Biol. **50:**513–524.
7. **Cole, J. R., B. Chai, T. Marsh, R. Farris, Q. Wang, S. Kulam, S. Chandra, D. McGarrell, T. Schmidt, G. Garrity, and J. Tiedje.** 2003. The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. Nucleic Acids Res. **31:**442–443.
8. **Daly, K., C. S. Stewart, H. J. Flint, and S. P. Shirazi-Beechey.** 2001. Bacterial diversity within the equine large intestine as revealed by molecular analysis of cloned 16S rRNA genes. FEMS Microbiol. Ecol. **38:**141–151.
8a.**Dick, L. K., M. T. Simonich, and K. G. Field.** 2005. Microplate subtractive hybridization to enrich for *Bacteroidales* genetic markers for fecal source identification. Appl. Environ. Microbiol. **71:**3179–3183.
9. **Duerden, B. I.** 1980. The isolation and identification of *Bacteroides* spp. from normal human fecal flora. J. Med. Microbiol. **13:**69–78.
10. **Field, K., D. Gordon, T. Wright, M. Rappé, E. Urbach, K. Vergin, and S. Giovannoni.** 1997. Diversity and depth-specific distribution of SAR11 cluster rRNA genes from marine planktonic bacteria. Appl. Environ. Microbiol. **63:**63–70.
11. **Fiksdal, L., J. S. Make, S. J. LaCroix, and J. T. Staley.** 1985. Survival and

detection of *Bacteroides* spp., prospective indicator bacteria. Appl. Environ. Microbiol. **49:**148–150.

12. **Franks, A. H., H. J. M. Harmsen, G. C. Raangs, G. J. Jansen, F. Schut, and G. W. Welling.** 1998. Variations of bacterial populations in human feces measured by fluorescent in situ hybridization with group-specific 16S rRNA-targeted oligonucleotide probes. Appl. Environ. Microbiol. **64:**3336–3345.

13. **Garrity, G. M., K. L. Johnson, J. Bell, and D. B. Searles.** May 2002, posting date. Taxonomic outline of the procaryotes. Bergey's manual of systematic bacteriology, 2nd ed, release 5.0. Springer-Verlag, New York, N.Y. [Online.] http://dx.doi.org/10.1007/bergeysoutline200301.

14. **Gherna, R., and C. R. Woese.** 1992. A partial phylogenetic analysis of the "*Flavobacter-Bacteroides*" phylum: basis for taxonomic restructuring. Syst. Appl. Microbiol. **15:**513–521.

15. **Giovannoni, S.** 2004. Oceans of bacteria. Nature **430:**515–516.

16. **Gordon, D. M.** 2001. Geographical structure and host specificity in bacteria and the implications for tracing the source of coliform contamination. Microbiology **147:**1079–1085.

17. **Gordon, D. M., and J. Lee.** 1999. The genetic structure of enteric bacteria from Australian mammals. Microbiology **145:**2673–2682.

18. **Hebert, P. D. N., and C. C. Wilson.** 1994. Provincialism in plankton: endemism and allopatric speciation in Australian *Daphnia*. Evolution **48:**1333–1349.

19. **Hedlund, B. P., and J. T. Staley.** 2002. Phylogeny of the genus *Simonsiella* and other members of the *Neisseriaceae*. Int. J. Syst. Evol. Microbiol. **52:**1377–1382.

20. **Hold, G. L., S. E. Pryde, V. J. Russell, E. Furrie, and H. J. Flint.** 2002. Assessment of microbial diversity in human colonic samples by 16S rDNA sequence analysis. FEMS Microbiol. Ecol. **39:**33–39.

21. **Hooper, L. V., and J. I. Gordon.** 2001. Commensal host-bacterial relationships in the gut. Science **292:**1115–1118.

22. **Hopkins, M. J., R. Sharp, and G. T. Macfarlane.** 2001. Age and disease related changes in intestinal bacterial populations assessed by cell culture, 16S rRNA abundance, and community cellular fatty acid profiles. Gut **48:**198–205.

23. **Hugenholtz, P., B. Goebel, and N. R. Pace.** 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J. Bacteriol. **180:**4766–4774.

24. **Hugenholtz, P., and T. Huber.** 2003. Chimeric 16S rDNA sequences of diverse origin are accumulating in the public databases. Int. J. Syst. Evol. Microbiol. **53:**289–293.

25. **Kreader, C. A.** 1995. Design and evaluation of *Bacteroides* DNA probes for the specific detection of human fecal pollution. Appl. Environ. Microbiol. **61:**1171–1179.

26. **Leser, T. D., J. Z. Amenuvor, T. K. Jensen, R. H. Lindecrona, M. Boye, and K. Moller.** 2001. Culture-independent analysis of gut bacteria: the pig gastrointestinal tract microbiota revisited. Appl. Environ. Microbiol. **68:**673–690.

27. **Leser, T. D., R. H. Lindecrona, T. K. Jensen, B. B. Jensen, and K. Moller.** 2000. Changes in bacterial community structure in the colon of pigs fed different experimental diets and after infection with *Brachyspira hyodysenteriae*. Appl. Environ. Microbiol. **66:**3290–3296.

28. **Ludwig, W., O. Strunk, R. Westram, L. Richter, H. Heier, Yadhukumar, A. Buchner, T. Lai, S. Steppi, G. Jobb, W. Förster, L. Brettske, S. Gerber, A. W. Ginhart, O. Gross, S. Grumann, S. Hermann, R. Jost, A. König, T. Liss, R. Lüβmann, M. May, B. Nonhoff, B. Reichel, R. Strehlow, A. Stamatakis, N. Stuckmann, A. Vilbig, M. Lenke, T. Ludwig, A. Bode, and K. H. Schleifer.** 2004. ARB: a software environment for sequence data. Nucleic Acids Res. **32:**1363–1371.

29. **McBee, R. H.** 1971. Significance of intestinal microflora in herbivory. Annu. Rev. Ecol. Syst. **2:**165–176.

30. **McCartney, A., W. Wenzhi, and G. Tannock.** 1996. Molecular analysis of the composition of the bifidobacterial and lactobacillus microflora of humans. Appl. Environ. Microbiol. **62:**4608–4613.

31. **Paster, B. J., F. E. Dewhirst, I. Olsen, and G. J. Fraser.** 1994. Phylogeny of *Bacteroides*, *Prevotella*, and *Porphyromonas* spp. and related bacteria. J. Bacteriol. **176:**725–732.

32. **Pryde, S. E., A. J. Richardson, C. S. Stewart, and H. J. Flint.** 1999. Molecular analysis of the microbial diversity present in the colonic wall, colonic lumen, and cecal lumen of a pig. Appl. Environ. Microbiol. **65:**5372–5377.

33. **Ramšak, A., M. Peterka, K. Tajima, J. C. Martin, J. Wood, M. E. A. Johnston, R. I. Aminov, H. J. Flint, and G. Avguštin.** 2000. Unravelling the genetic diversity of ruminal bacteria belonging to the CFB phylum. FEMS Microbiol. Ecol. **33:**69–79.

34. **Rawls, J. F., B. S. Samuel, and J. I. Gordon.** 2004. Gnotobiotic zebrafish reveal evolutionarily conserved responses to the gut microbiota. Proc. Natl. Acad. Sci. USA **101:**4596–4601.

35. **Robison-Cox, J., M. Bateson, and D. Ward.** 1995. Evaluation of nearest-neighbor methods for detection of chimeric small-subunit rRNA sequences. Appl. Environ. Microbiol. **61:**1240–1245.

36. **Salyers, A. A.** 1984. *Bacteroides* of the human lower intestinal tract. Annu. Rev. Microbiol. **38:**293–313.

37. **Salyers, A. A., and J. A. Shipman.** 2002. Getting in touch with your prokaryotic self: mammal-microbe interactions, p. 315–341. *In* J. T. Staley and A. L. Reysenbach (ed.), Biodiversity of microbial life. Wiley-Liss, Inc., New York, N.Y.

38. **Sghir, A., G. Gramet, A. Suau, V. Rochet, P. Pochart, and J. Dore.** 2000. Quantification of bacterial groups within human fecal flora by oligonucleotide probe hybridization. Appl. Environ. Microbiol. **66:**2263–2266.

39. **Simpson, J. M., J. W. Santo Domingo, and D. J. Reasoner.** 2003. Assessment of equine fecal contamination: the search for alternative bacterial source-tracking targets. FEMS Microbiol. Ecol. **1593:**1–11.

40. **Staley, J. T., and J. J. Gosink.** 1999. Poles apart: biodiversity and biogeography of sea ice bacteria. Annu. Rev. Microbiol. **53:**189–215.

41. **Suau, A. R. B., M. Sutren, J. J. Godon, G. R. Gibson, M. D. Collins, and J. Dore.** 1999. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. Appl. Environ. Microbiol. **65:**4799–4807.

42. **Swofford, D. L.** 2001. Phylogenetic analysis using parsimony (∗and other methods), version 4.0b6. Sinauer Associates, Sunderland, Mass.

43. **Wang, G. C. Y., and Y. Wang.** 1997. Frequency of formation of chimeric molecules as a consequence of PCR coamplification of 16S rRNA genes from mixed bacterial genomes. Appl. Environ. Microbiol. **63:**4645–4650.

44. **Whitford, M. F., R. J. Forster, C. E. Beard, J. Gong, and R. M. Teather.** 1998. Phylogenetic analysis of rumen bacteria by comparative sequence analysis of cloned 16S rRNA genes. Anaerobe **4:**153–163.

45. **Wood, J., K. P. Scott, C. J. Newbold, and H. J. Flint.** 1998. Estimation of the relative abundance of different *Bacteroides* and *Prevotella* ribotypes in gut samples by restriction enzyme profiling of PCR-amplified 16S rRNA gene sequences. Appl. Environ. Microbiol. **64:**3683–3689.

46. **Xu, J., and J. I. Gordon.** 2003. Honor thy symbionts. Proc. Natl. Acad. Sci. USA **100:**10452–10459.

47. **Zhu, X. Y., T. Zhong, Y. Pandya, and R. D. Joerger.** 2002. 16S rRNA-based analysis of microbiota from the cecum of broiler chickens. Appl. Environ. Microbiol. **68:**124–137.