Comparison of *Bacteroides-Prevotella* 16S rRNA Genetic Markers for Fecal Samples from Different Animal Species

Lisa R. Fogarty and Mary A. Voytek*

U.S. Geological Survey, Water Resources Discipline, National Research Program, Reston, Virginia 20192

To effectively manage surface and ground waters it is necessary to improve our ability to detect and identify sources of fecal contamination. We evaluated the use of the anaerobic bacterial group *Bacteroides-Prevotella* as a potential fecal indicator. Terminal restriction length polymorphism (T-RFLP) of the 16S rRNA genes from this group was used to determine differences in populations and to identify any unique populations in chickens, cows, deer, dogs, geese, horses, humans, pigs, and seagulls. The group appears to be a good potential fecal indicator in all groups tested except for avians. Cluster analysis of *Bacteroides-Prevotella* community T-RFLP profiles indicates that *Bacteroides-Prevotella* populations from samples of the same host species are much more similar to each other than to samples from different source species. We were unable to identify unique peaks that were exclusive to any source species; however, for most host species, at least one T-RFLP peak was identified to be more commonly found in that species, and a combination of peaks could be used to identify the source. T-RFLP profiles obtained from water spiked with known-source feces contained the expected diagnostic peaks from the source. These results indicate that the approach of identifying *Bacteroides-Prevotella* molecular markers associated with host species might be useful in identifying sources of fecal contamination in the environment.

Fecal contamination continues to threaten the quality of receiving waters despite efforts made to minimize inputs from sewage and septic systems, feedlots, and other point sources. During the years 1999 to 2000, the Centers for Disease Control and Prevention reported 39 water-borne disease outbreaks associated with drinking water and 59 outbreaks associated with recreational waters (8). It is estimated that nearly 40,000 kilometers of streams and coastal waters in the United States are contaminated with concentrations of fecal bacteria that exceed recommended criteria for the protection of human health (33, 34). In order to effectively manage these waters, managers must be able to detect fecal pollution and identify the potential source(s) in a reliable and timely manner.

Current methods to detect sources of fecal pollution rely upon cultivation of fecal-indicator bacteria (fecal coliforms, *Escherichia coli*, or enterococci) or viruses (coliphage). Sources of the contamination are inferred based on biochemical characterization or molecular fingerprinting of the cultivated indicator; this is often referred to as microbial source tracking. The most commonly used microbial source tracking methods include antibiotic resistance analysis (15, 16, 36), repetitive PCR (9), pulsed-field gel electrophoresis (25), and ribotyping (7, 25).

Currently used fecal indicator bacteria constitute a small percentage of the fecal flora, and cultivation methods may bias the populations by selection, thus misrepresenting the source populations. In addition, many commonly used methods require a comprehensive library of isolate profiles from known sources to which isolates from an unknown source can then be compared. A minimal number of isolates needed to perform statistically sound studies has not been determined and may vary by system or source (21, 36). A recent study by Stoeckel et al. (30) comparing several library-based methods demonstrated the abilities and limitations of all methods considered and highlighted the critical importance of the library size.

Several studies have proposed using members of the *Bacteroides* genus as alternative indicators of fecal pollution (1, 2, 3, 13, 20). Members of the *Bacteroides* group are strict anaerobes and represent a much larger proportion of the fecal population than the fecal coliform or enterococcus groups (28, 29, 31, 37). The difficulty in growing these strict anaerobes limited their utility as fecal indicators. However, the recent developments of PCR primers to target the genus *Bacteroides* (2, 20) have made detection and identification without cultivation possible. Studies by Bernhard and Field (2, 3) identified specific *Bacteroides-Prevotella* 16S rRNA gene markers for human and cow fecal material by using length heterogeneity PCR and terminal restriction fragment length polymorphism (T-RFLP). Bernhard and Field (3) suggested that these markers could be used as an alternative to cultivation-dependent microbial source tracking methods.

The goals of this study were to compare *Bacteroides-Prevotella* populations from nine host species collected at multiple geographical locations and to determine if unique populations could be identified for each host species that could be used to develop markers for fecal source tracking.

MATERIALS AND METHODS

Sample collection. Freshly excreted fecal material was collected from individual animals using sterile techniques from the following host species: chicken, cattle, horses, deer, geese, seagulls, dogs, pigs, and humans. Table 1 lists the number of samples analyzed for each host species and the locations where samples were collected. For domesticated animals, samples were collected from

* Corresponding author. Mailing address: U.S. Geological Survey, Water Resources Discipline, National Research Program, MS430, 12201 Sunrise Valley Dr., Reston, VA 20192. Phone: (703) 648-6894. Fax: (703) 648-5484. E-mail: mavoytek@usgs.gov.

TABLE 1. Number of T-RFLP profiles obtained for each fecal source and sampling locations

Fecal source	No. of fecal samples tested[a]	Location of collection (no. of samples)
Cow	48	Virginia (10), West Virginia (25), Indiana (13)
Dog	20	Virginia (9), West Virginia (11)
Deer	21	Virginia (10), West Virginia (11)
Geese	36 (14)	Virginia (36)
Human	20	West Virginia (20)
Pig	19	Virginia (4), West Virginia (15)
Horse	23	Virginia (10), West Virginia (13)
Chicken	22 (5)	Virginia (10), West Virginia (13)
Seagull	14 (0)	Virginia (14)

[a] Number of samples that were amplifiable with the *Bacteroides/Prevotella* primers shown in parentheses.

at least four different locations and approximately four to five animals were sampled per location. All samples from Virginia and Indiana were transported on ice to the laboratory within 6 hours of collection and held at −20°C for later analysis. Samples collected in West Virginia were collected in conjunction with a separate study (the Ohio and West Virginia District offices of the U.S. Geological Survey) and are designated WV. A 1-ml slurry (0.1 g of fecal material in 3 ml of phosphate buffer water, pH 7.2) of the West Virginia samples was preserved as 40% glycerol stocks and frozen at −80°C. A comparison of replicate samples preserved by both methods demonstrated that the preservation method did not influence the T-RFLP profiles (results not shown).

DNA extraction. About 2 g of fecal material were homogenized in 10 ml of QIamp DNA stool mini kit ASL buffer (QIAGEN, Valencia, CA). DNA was extracted from 2 ml of the homogenate following the manufacturer's protocols for pathogen detection. A study on different DNA extraction kits and methods by McOrist et al. (23) determined that the QIamp DNA stool mini kit was the most effective extraction method for their application. For this application, the QIamp DNA Stool Mini kit was compared with UltraClean Soil DNA kit (Mo Bio Laboratories, Inc., Solana Beach, CA) and FastDNA Spin kit for soil (Bio 101, Carlsbad, CA). Using replicate samples, the QIamp DNA stool mini kit consistently resulted in higher DNA recovery and higher DNA purity. The West Virginia samples (frozen glycerol stock fecal slurries) required different treatment. Glycerol stocks were thawed and added to 5 ml of ASL buffer (QIAGEN). DNA was extracted from this entire slurry following the same protocol as above except adding proportionally more of each reagent. To evaluate the homogeneity of the fecal samples and reproducibility of our results, triplicate DNA extractions were performed on a select group of samples from each source species.

T-RFLP. Extracted DNA was amplified with primers described in Field and Bernhard (2), forward primers HF183 AND CF128, labeled with flourophores 6-carboxyfluorescein and Bac708R (Table 2). The PCR protocol used was optimized to ensure specificity and inclusiveness and to ensure the highest product yield. The PCR mixture components consisted of 1X PCR buffer, 1.2 mg/ml bovine serum albumin, 0.2 mM deoxynucleoside triphosphates, 0.2 μM of each primer, 1.25 U of *Taq*, approximately 20 ng of template DNA (determined by spectrophotometer absorbance at 260 nm), and sterile tissue culture water (Sigma-Aldrich, St. Louis, MO) to bring the volume up to 50 μl.

A touchdown DNA amplification was carried out in a PE Biosystems Gene Amp 9700 (Perkin Elmer-Cetus, Norwalk, CT) with the following conditions: 94° C for 3 min; 10 cycles of 94°C for 30 seconds, 63°C for 30 seconds (decreasing 1°C each cycle), and 72°C for 30 seconds; 20 cycles of: 94°C for 30 seconds, 53°C for 30 seconds, and 72°C for 90 seconds; a final elongation of 7 min at 72°C; and a final hold at 4°C. PCR products (5 μl of each) were electrophoresed on a 1% agarose gel to check for quality and quantity of the amplified product. Three

different restriction digests were performed on each sample by adding 7 μl of amplified product to 5 U of the following restriction enzymes in separate reactions: AciI (New England BioLabs, Beverly MA), HaeIII (Promega, Madison, WI), and MspI (Promega). Digested samples were precipitated with 0.1 volume of 3 M sodium acetate and 2 volumes of 100% cold ethanol and resuspended in 10 μl sterile water.

T-RFLP profiles. To 2.5 μl of the digested sample, 12 μl of deionized formamide (Applied Biosystems, Foster City, CA) and 0.5 μl ROX 500 standard (Applied Biosystems) were added. Samples were denatured at 95°C for 5 min. DNA fragments were separated using an ABI 310 sequencer (Applied Biosystems). Terminal restriction fragment sizes between 50 and 500 base pairs and with a fluorescence ≥50 were detected using 310Genescan analytical software, version 2.1.1 (Applied Biosystems), resulting in a T-RFLP profile for each sample and restriction enzyme. The ROX 500 standard was used as a reference in sample to ensure consistency in each run.

Statistical analysis. T-RFLP peak profiles were converted into binary form (peak height fluorescence ≥ 50 considered positive). The T-RFLP profiles for each enzyme (AciI, MspI, and HaeIII) were combined to form one large profile. These combination profiles were compared by calculating the Jaccard similarity coefficients for each pair of profiles and creating a dendrogram by UPGMA using BioNumerics software (version 2.5 Applied Maths; Kortrijk, Belgium). Using the BioNumerics software package, a discriminant analysis with variance was also performed to identify T-RFLP peaks associated with each host species.

Feces-contaminated water. To determine the stability of *Bacteroides-Prevotella* T-RFLP profiles once feces come into contact with the environment, a representative surface water was inoculated with fresh cow feces. An aliquot of 2 liters of surface water was collected from a pond and 1 liter of this sample was filtered onto a 0.22-μM Sterivex filter (Millipore, Bedford, MA). DNA was extracted from this filter using the PUREGENE DNA isolation kit (Gentra Systems Inc., Minneapolis, MN). Initial water and fecal sample *Bacteroides-Prevotella* T-RFLP profiles were obtained using the methods described above. The remaining liter of water sample was inoculated with 1 g of the fecal sample and thoroughly mixed in a UV-transmitting polypropylene bottle. The sample was placed outdoors in direct sunlight in a water bath maintained at 18°C and remained aerobic throughout the incubation. Sample aliquots of 50 ml were collected and filtered onto 0.22-μM Sterivex filters at the beginning of the experiment, and at days 1, 3, 5, and 12. *Bacteroides-Prevotella* T-RFLP profiles were obtained for these samples as described above.

Test of human and cow markers. To determine if previously identified markers for humans and cows could identify human and cow fecal samples in this study, 20 cow and 20 human fecal samples were tested as described by Bernhard and Field (3), with forward primer HF183 for human markers and reverse primer CF128 for the cow marker (Table 2).

RESULTS

T-RFLP patterns. The 16S rRNA gene was amplified with primers specific for the *Bacteroides-Prevotella* group and cut with single restriction enzymes AciI, MspI, and HaeIII. The data set consisted of three T-RFLP fingerprints from each of 165 fecal samples (48 cow, 20 dog, 21 deer, 14 geese, 23 horse, 20 human, and 19 pig) for a total of 495 profiles (Table 1). The initial sample set included an additional 22 geese, 22 chicken, and 14 seagull samples, however, only 14 of the 28 tested geese and 5 out of 22 chickens were amplifiable with the *Bacteroides/Prevotella* primer set. Although attempts were made to reoptimize the PCR protocol in order to amplify avian samples, they were unsuccessful. Thus, it is likely that the limited amplification was a result of the absence of the target in these

TABLE 2. Primer sequences and targets used for PCR amplification

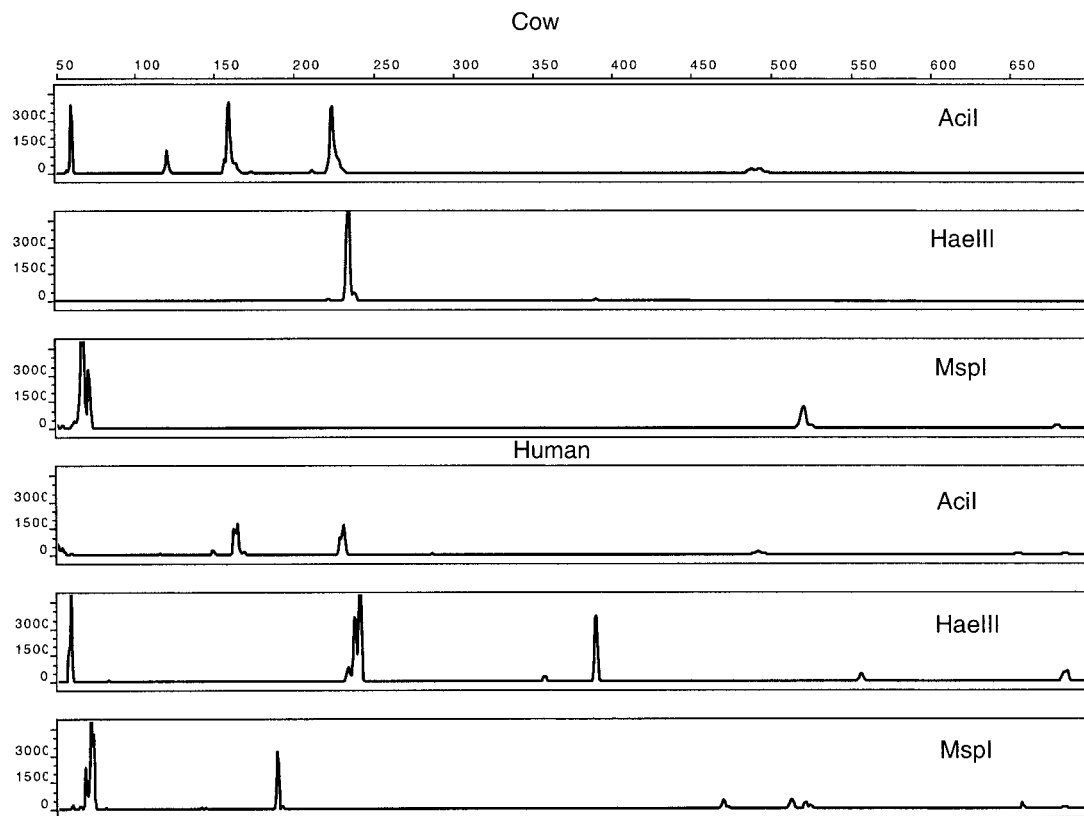| Primer	Target	Sequence, 5′–3′	Reference
Bac32F	*Bacteroides-Prevotella* group	AACGCTAGCTACAGGCTT	2
Bac708R	*Bacteroides-Prevotella* group	CAATCGGAGTTCTTCGTG	2
CF128F	Cow marker	CCAACYTTCCCGWTACTC	3
HF183F	Human marker	ATCATGAGTTCACATGTCCG	3

FIG. 1. Examples of T-RFLP peaks resulting from cow and human fecal samples cut with AciI, MspI, and HaeIII.

samples rather than poor primer performance. Therefore, with the exception of the 14 amplifiable geese samples, these samples were not included for the remainder of the study. Those geese samples that were amplifiable were only included to determine the presence of identified potential source markers.

Figure 1 displays an example of the T-RFLP that resulted from each enzyme for a representative cow and human fecal sample. While each fingerprint was completely unique, certain peaks were found in feces from all host species and other peaks were found more commonly in a single host species. To determine the homogeneity of the fecal samples and the reproducibility of the method, triplicate DNA extractions were done for a select group of samples from each source. Corresponding peak locations in T-RFLP profiles from replicate samples were nearly identical (±0.3 base pairs). However, slight differences occurred in peak height.

Cluster analysis. UPGMA cluster analysis using Jaccard similarity coefficients was performed using the combination of all three enzymes profiles. Six major clusters resulted from this analysis (similarity ≥10%; Fig. 2). Cluster A is comprised of five minor cluster groups (A1 to A5), which contains all of the cow ($n = 48$) and most of the deer (17 of 21) fecal samples. Within this major cluster group the cow and the deer samples clustered into two large groupings, A1, defined by 41 of the 48 cow samples, and A2, defined by 16 deer samples. The remaining seven cow samples are in the small clusters A3 to A5. Most similar to the cow and deer cluster is a cluster of pig samples (cluster B, 15 out of 19 pig samples). Most of the human and

dog samples make up clusters C and D, most similar to each other (cluster C, 10 out of 20 dog samples, and cluster D, all 20 human and 3 out of 20 dog samples). The majority of the horse samples clustered in group E.

For selected host species, samples were collected from multiple locations (cow, horse, deer, and dog) and seasons (deer). Cow and horse samples were collected from multiple farms and in different parts of the country. Differences in collection location were not distinguishable between samples within a source group. Likewise, deer fecal samples collected from many different locations did not cluster according to location (data not shown). For cattle fecal samples that could be identified as beef or dairy (the additional collection from West Virginia), there was a tendency for these samples to form minor beef and dairy groups within the larger cow group (Fig. 3).

Discriminant analysis. To help identify peaks that separate source species, discriminant analysis was performed. Samples were manually grouped based on source species and the best discriminating components were calculated for these groups and represented on discriminant axes. This is illustrated in Fig. 4 as a three-dimensional image. The output of the discriminant analysis resulted in a set of T-RFLP peaks that corresponded to each discriminant axis that separated out the source groups. A number of these peaks were selected as potential source markers as shown in Fig. 4 and Table 3. These T-RFLP peaks were compared to the T-RFLP profiles generated for each sample. The percentage of samples containing each peak for each source group was calculated and is shown
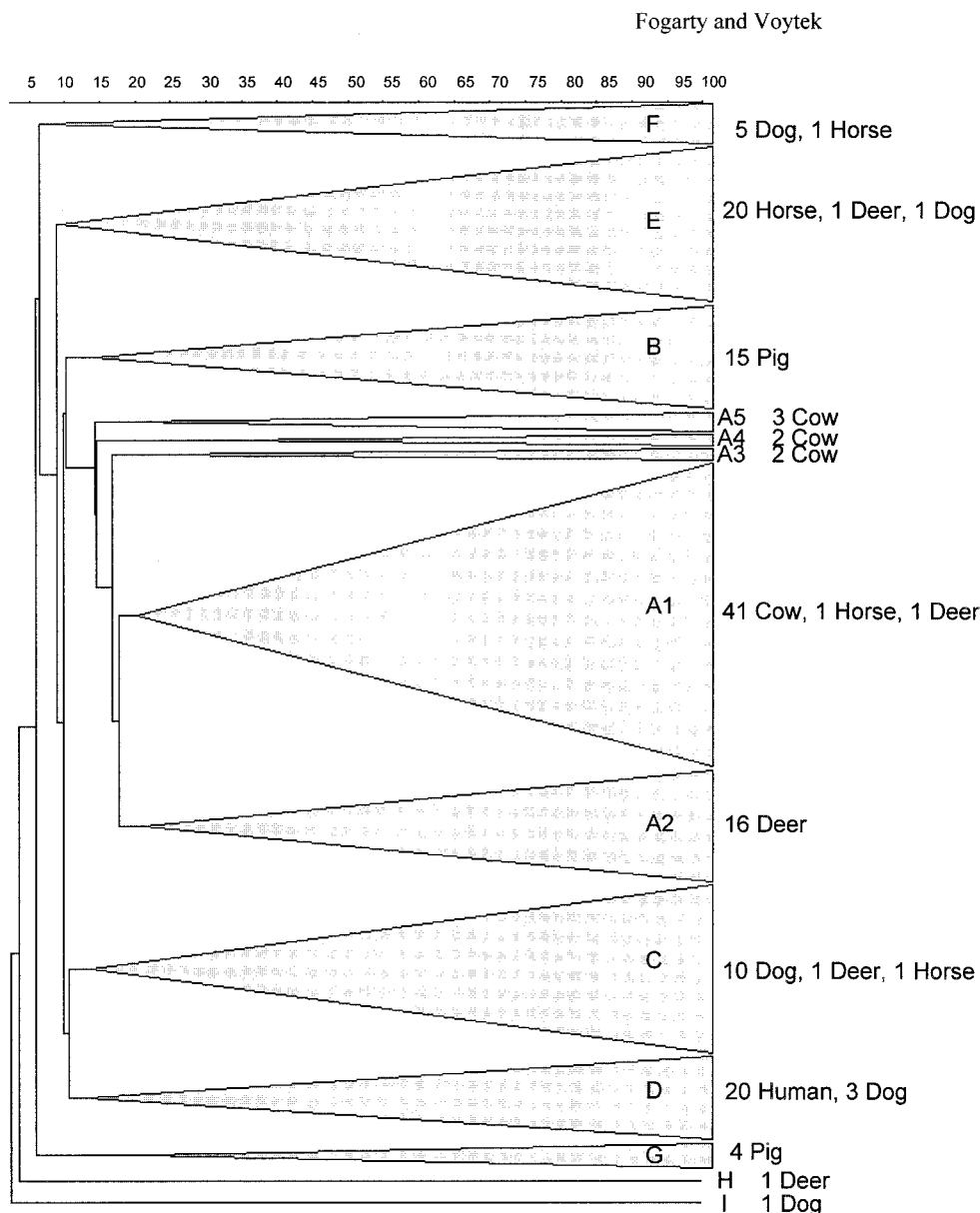
Fogarty and Voytek

FIG. 2. UPGMA cluster analysis based on Jaccard similarity coefficients of all samples with a combination of all three enzymes (AciI, MspI, and HaeIII).

in Table 3. No single peak was found in every sample belonging to a single source group and only one peak was unique to a single group, peak MspI 177 for the horse group. There were, however, many peaks that were more commonly associated with a specific source group and these were identified as potential source markers in this study.

Detection of T-RFLP profiles in water. The stability and persistence of diagnostic *Bacteroides-Prevotella* T-RFLP profiles retrieved from environmental samples were evaluated by adding fresh cattle feces to surface water collected from a local pond and monitoring the profiles over time. PCR amplification for the *Bacteroides-Prevotella* group-specific 16S rRNA gene in the pond water prior to inoculation was negative. The T-RFLP profiles of the feces-amended pond water remained constant

over the 12-day experiment and were identical to the original T-RFLP profiles for the original cow fecal sample.

Human and cow markers. Bernhard and Field developed two primers based on source-specific T-RFLP peak occurrence in order to selectively detect human or cow sources. We did not detect these T-RFLP peaks in our human and cow samples; therefore, we tested for the presence of these markers using the primers and DNA from 20 human fecal samples along with 20 cow, 6 pig, 23 horse, and 19 dog samples. The human marker was amplified from only 4 of the 20 human fecal samples tested but was not found in any of the other fecal source samples tested. The cow marker (CF128F) also described in Bernhard and Field (2) amplified all 20 cow samples tested and none of the 20 human samples tested amplified with this marker. However, all
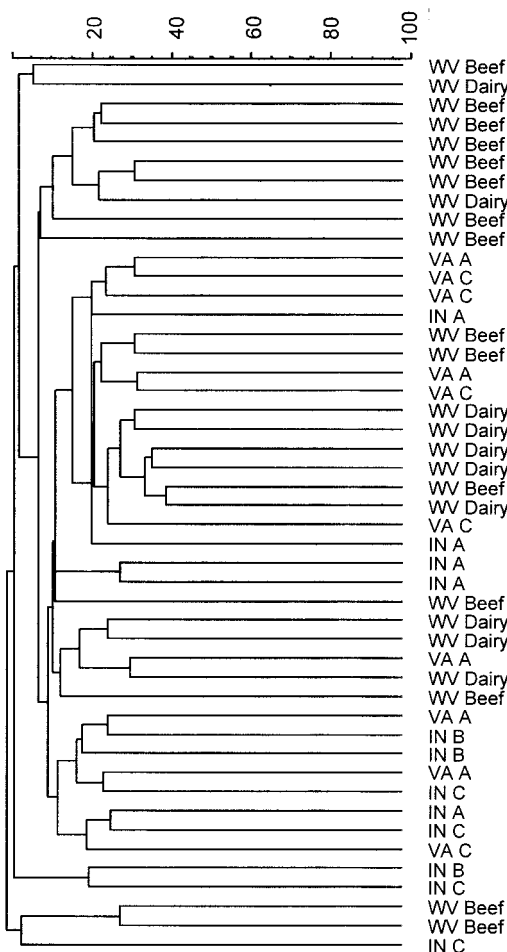
FIG. 3. UPGMA cluster analysis based on Jaccard similarity coefficients of cow samples with a combination of all three enzymes (AciI, MspI, and HaeIII). Samples were collected from farms in three locations: WV, West Virginia; VA, Virginia (two farms, A and C); and IN, Indiana (three farms, A, B, and C).

samples including the negative samples were amplified with the *Bacteroides/Prevotella* group primer set.

DISCUSSION

Many concerns have been raised regarding the appropriateness and efficacy of the bacterial water quality indicators currently in use (fecal coliforms, *E. coli*, and enterococci). In general, these fecal indicator groups represent a small percentage of the fecal flora, and therefore may be relatively insensitive indicators for contamination. They survive an indeterminable amount of time in the environment, and sometimes grow in the environment when environmental conditions are favorable (27). In addition, some fecal coliforms and enterococci species are not restricted to fecal sources, and are commonly found naturally in the environment (6, 26, 27).

The *Bacteroides-Prevotella* group has been proposed as an alternative and possibly improved water quality indicator. This group is abundant, and has thus far only been detected in human and other animal fecal sources. They may be useful to detect sources of fecal contamination, since prior studies have

shown populations to be specific to host species (1, 2, 3, 13, 28, 33). Members of the *Bacteroides-Prevotella* group have been detected in feces-contaminated surface waters in both urban and agricultural areas (1, 2, 3, 19). Unlike facultative anaerobic fecal source indicator bacteria, members of this group are obligate anaerobes. Although anaerobic bacteria could reproduce in sediments or localized anoxic environments, they would be unlikely to reproduce in aerobic aquatic environments. Therefore, detection of *Bacteroides* spp. may be a more accurate indicator of a recent and/or extensive fecal contamination event (1, 13).

There have been a few studies testing the survival of *Bacteroides* spp. For example, Kreader (19) demonstrated that the survival of *Bacteroides* species is dependent on water temperature and predation. The lower the water temperature and predation pressure, the longer *Bacteroides distasonis* survived. Survival rates for *B. distasonis* ranged from 14 days at 4°C to only 1 to 2 days at 24°C. Removal of predation pressure by filtration of river water and addition of cycloheximide to inhibit eukaryotes resulted in an increase of survival of up to 12 days at these temperatures. Our experiments to determine the stability of the *Bacteroides-Prevotella* fecal populations indicated that there was little change in the *Bacteroides-Prevotella* community profiles after inoculation. We obtained identical T-RFLP profiles from the original fecal sample and from samples collected from the inoculated pond water during the 12 days of incubation at 18°C even though the incubated samples were aerobic, subjected to natural UV exposure and natural (e.g., micrograzer) predation pressures.

In this study, we employed T-RFLP analysis of 16S rRNA gene to determine differences in populations of the *Bacteroides-Prevotella* group from nine host species (cattle, chickens, deer, dogs, geese, horses, humans, pigs, and seagulls) and to evaluate the potential use of this group and technique in fecal source identification. T-RFLP has successfully been used in many studies to describe community changes in a variety of environments (5, 10, 11, 17, 18, 24). Biases of the method are discussed in detail in recent reviews (5, 18, 22). We found T-RFLP analysis of our samples to be highly reproducible. Nearly identical T-RFLP profiles were obtained with triplicate DNA extractions from the same fecal samples, demonstrating that the protocols employed are reproducible and that the distribution of *Bacteroides-Prevotella* populations in fecal samples themselves was homogeneous. Differences in peak intensity occurred with differences in extracted DNA concentrations. Minor peaks often were not detected from samples with lower concentrations of DNA. These samples could be identified by an overall lower total intensity (i.e., the sum of the total peak areas). Despite differences in peak intensities, the T-RFLP peak locations were reproducible with a variance of ±0.3 bp.

Cluster analysis indicated that the *Bacteroides-Prevotella* populations within host species are much more similar than populations between host species. Results of the cluster analysis indicated no seasonal difference in the deer source group nor any differences associated with geographic area (e.g., between cattle from Virginia versus Indiana). There was some evidence of subspecies level population difference between beef and dairy cattle (Fig. 3). On a broader level, we observed an association between cluster groups that appears to suggest
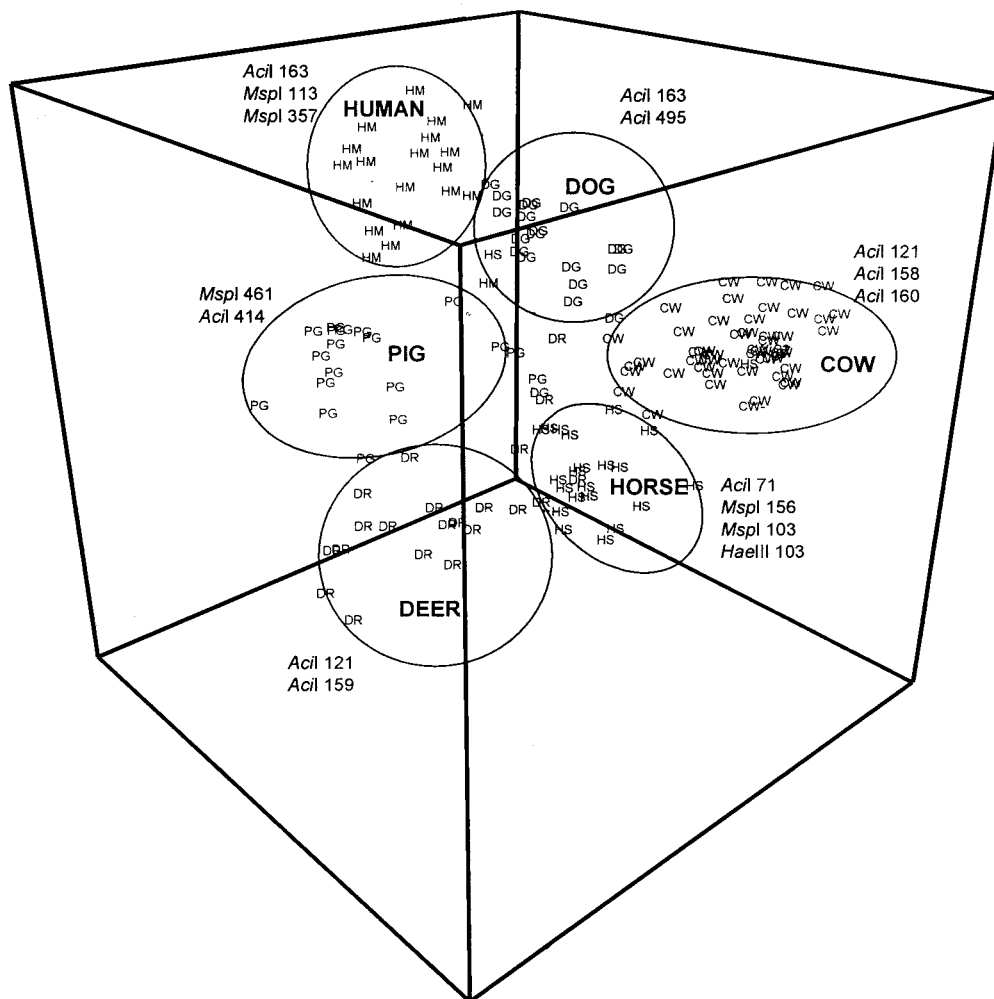
FIG. 4. Discriminant analysis displayed in three dimensions of all samples and a combination of all three T-RFLP enzymes (AciI, MspI, and HaeIII). Selected T-RFLP peaks associated with the discriminant axis that defines each source group.

a functional relation among source groups. For example, *Bacteroides-Prevotella* populations from the two ruminant animals in the study, cows and deer, were more similar to each other than to other animals (cluster A in Fig. 4). Another broad grouping included the samples from humans and dogs (clusters C and D in Fig. 2). The similarity of these populations could be a result of close contact in a shared living environment or due to similar diets and digestive systems.

Bacteroides/Prevotella was difficult to amplify from avian fecal samples and likely was not present in many samples. Fecal samples were collected from geese, seagulls, and chicken in this study. *Bacteroides* was amplifiable from only 38% (14 out of 36) of the geese fecal samples collected and 23% (5 out of 22) of the chicken samples and was unamplifiable from all 14 seagull samples (Table 1). As a result, fecal contamination by avian species would be difficult to detect using *Bacteroides/Prevotella* markers. Due to insufficient data (i.e., most avian fecal samples did not contain the *Bacteroides-Prevotella* target), seagull and chicken samples could not be included in the T-RFLP study. Although there was significant variability in peaks among the few geese samples that were amplifiable

TABLE 3. Percentage of samples that contain the identified peak from the discriminant analysis for each source group

Enzyme	T-RFLP size (bp)	% of samples[a]						
		Cow	Dog	Deer	Geese	Horse	Human	Pig
AciI	121	**96**	5	**62**	36	17	0	0
AciI	158	**90**	5	14	36	26	0	5
AciI	160	**100**	10	**48**	43	39	0	20
AciI	495	2	**50**	5	7	0	0	0
AciI	159	2	5	**62**	0	39	0	5
MspI	103	0	10	0	0	**74**	0	5
MspI	156	2	5	0	0	**80**	0	0
AciI	71	0	0	5	21	**52**	0	0
MspI	141	0	0	5	7	**57**	0	0
MspI	177	0	0	0	0	**61**	0	0
AciI	163	4	45	5	57	26	**95**	0
MspI	113	0	5	0	14	0	**60**	0
HaeII	357	0	5	0	0	0	**45**	5
MspI	461	0	0	0	0	22	0	**47**
AciI	414	6	5	10	14	4	5	**74**
Total no. of samples		48	20	21	14	23	20	19

[a] Possible source marker peaks are in bold.

(Table 3), the geese T-RFLPs were included only to determine if potential source tracking peaks would be found in geese samples.

In a recent study examining *Escherichia coli* and enterococcus populations in gull feces (by repetitive PCR and by biochemical characterization, respectively), Fogarty and coworkers found such high variability in the populations within and between individuals that they were unable to assign a representative profile for future source identification (14). Similarly, our results demonstrated significant differences in *Bacteroides-Prevotella* populations between samples from individual geese and a general lack of coherent clustering of individuals. The factors responsible for these observations are unknown but could include digestive process, migration or foraging behavior resulting in changes in food source, or eating feces from other species (35).

Several previous studies using a combination of cultivation and molecular techniques have demonstrated differences in *Bacteroides-Prevotella* populations in different animal species (2, 3, 20, 37). Two of these studies suggested specific *Bacteroides* populations might be more abundant in one animal species than others, but were unable to identify specific species or ribotypes unique to the different animal species (20, 37). Results of our study were similar; we found only one T-RFLP peak that was exclusive to host species (61% of the horse group contained peak MspI 177). Bernhard and Field (2) reported molecular markers (Table 2) obtained using T-RFLP profiles and their potential use in identifying and differentiating human and cow fecal samples (2–4). When amplified *Bacteroides-Prevotella* DNA was cut with AciI, a "cow-specific" T-RFLP peak was identified at 227 base pairs (2). When cut with HaeIII, another "cow-specific" T-RFLP peak was identified at 222 base pairs and a "human-specific" T-RFLP peak was identified at 119 base pairs (2). Using similar methods with the same PCR primers and restriction enzyme, we did not obtain these T-RFLP peaks from their respective source.

In our study, we found the human-specific T-RFLP peak at 119 bp (HaeIII) in just 1 of 20 human fecal samples; but we did find that peak in one cow fecal sample and four dog, one goose, three horse, and one pig sample. When the human-specific primer set described in the Bernhard and Field study (2) as the human marker was tested using samples from our study, only 4 out of the 20 human samples were amplifiable. Although the human marker appeared in only 20% of the human samples, it was not found in the tested cow ($n = 20$), pig ($n = 6$), horse ($n = 23$), and dog ($n = 19$) samples, suggesting good specificity. Field et al. (12) reported a similarly low success rate for individuals but a higher rate of success for sewage samples. Sources are composites of multiple individuals, and the goal would be to select a marker that occurred in the highest percentage of individuals from that one source while not occurring with other sources. In certain natural environments (i.e., not sewage), a marker present in only 20% of the population may not be high enough to provide a reliable indicator. However, finding a marker that occurs in a higher percentage of individuals from that source group would provide greater sensitivity.

The cow-specific T-RFLP peaks at 227 bp (AciI) and 222 bp (HaeIII) identified in the Berndhard and Field study (3) were found in only 25% and 19%, respectively of the 48 cow fecal samples from our study. In our study, the cow 227-bp (AciI)

T-RFLP peak was found in a larger percentage of noncow fecal samples than in the cow fecal samples, including 45% of the dog, 43% of the geese, 30% of the horse, and 58% of the pig versus 40% of the cow fecal samples. It was also found in 24% of the deer fecal samples. Despite the small number of samples containing the T-RFLP cow-specific peak, the cow-specific marker CF128F (2) proved to be a good cow marker for samples in our study. All 20 cow samples tested amplified with the cow marker primer set (CF128F and Bac32R), and none of the 20 human samples tested were amplified.

The T-RFLP profiles were used to determine if other source markers were possible, not only for cow and human fecal sources, but also for other animal sources. Our study indicates that a group of T-RFLP peaks could be used to distinguish between cow, deer, dog, horse, pig, and human fecal samples. Fecal source markers were identified by taking the three most influential peaks associated with each discriminant axis and determining the percentage of each source group that contained that peak. From that analysis, we identified potential markers for each source group. Table 3 lists the peaks identified by the discriminant analysis and the percentage of samples from each source group that contain the specific peak.

The horse group contained one peak exclusive to that group, which 61% of the horse samples and only the horse samples contained. The horse group also contained three other peaks which 50% or more of samples contained and only a small percentage of samples from other source groups also contained. Three peaks were found in 90% or more of the cow isolates, AciI 121, 158, and 160. These peaks are not, however, exclusive to the cow group. Nearly half of the deer samples also contained peaks at AciI 121 and 160. This may not be surprising since the cluster analysis showed that these two groups were very similar. However, sequencing of these peaks may differentiate them enough to design new source-specific markers. In addition, there were two peaks common to many of the human samples, and two peaks common to many of the pig samples that were found only in a small percentage of other source species. There is one dog peak most commonly associated with dogs but only 50% of the dog fecal samples contained this peak. Further studies to sequence the peaks associated with each source group might reveal new markers for tracking sources of fecal contamination.

Our results support the use of molecular techniques to characterize *Bacteroides-Prevotella* populations as a means to improve the ability to track sources of fecal contamination, but also show the need for more development of these methods. The sample populations in this study were admittedly small; more investigations are needed to determine the variation in identified markers between source animals, and to determine how these markers may vary due to geography, farming practices, season, etc. In addition to the methodological considerations, there are numerous factors that may influence the fecal bacteria community and thereby affect any analysis and subsequent interpretation. These factors include physiological factors such as digestive process, sex or age, lifestyle factors such as domestication, diet, and antibiotic exposure, and seasonal activities such as migration and food limitation (32). Few, if any, studies have used molecular techniques to examine the effect of these factors on *Bacteroides-Prevotella* fecal community dynamics in a variety of animal species. In order to address

some of these factors, we analyzed fecal samples from animals with different digestive systems (i.e., ruminants and nonruminants) and from different geographical locations and times of year and from wild and domesticated animals. Diet, age, and antibiotic exposure were not specifically identified for the fecal samples and therefore could not be addressed. Studies of this nature are important before there can be wider application of any technique.

To date, published studies of library-independent microbial source tracking tools have been largely limited to human and cows. This study has contributed to the investigation of fecal indicators and source tracking by including more fecal sources and more sources from different areas. Future work to refine this approach should include additional fecal samples from the sources investigated in this study and from other potential fecal sources such as sewage, other confined livestock such as swine, sheep, and goats or common and/or abundant wildlife. Samples need to be collected from a variety of locations at various times of the year. Future work should also include DNA sequence analysis of the identified fecal source markers to clarify the specificity of each marker. Development of specific probes based on these markers, as in Bernhard and Field (2, 3), would permit water quality specialists to probe an environment and quantify sources of contamination by quantitative PCR or direct hybridization techniques (i.e., molecular probing requiring no amplification).

Our findings suggest that source identification by the interpretation of a T-RFLP profile generated from an environmental sample contaminated by a single source, should be relatively straightforward. However, in many environmental situations there is often more than one contributing source. As the environmental setting becomes more complex (e.g., with multiple fecal inputs), it may not be possible to develop a single cost-effective diagnostic tool for source tracking studies; rather, an integrated approach that incorporates microbiological, chemical, and hydrological data may be required.

ACKNOWLEDGMENTS

REFERENCES

1. **Allsop, K., and D. J. Stickler.** 1985. An assessment of *Bacteroides fragilis* group organisms as indicators of human faecal pollution. J. Appl. Bacteriol. **58:**95–99.
2. **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.
3. **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.
4. **Bernhard, A. E., T. Goyard, M. T. Simonich, and K. G. Field.** 2003. Application of a rapid method for identifying fecal pollution sources in a multi-use estuary. Water Res. **37:**909–913.
5. **Blackwood, C. B., T. Marsh, S. Kim, and E. A. Paul.** 2003. Terminal restriction fragment length polymorphism data analysis for quantitative comparison of microbial communities. Appl. Environ. Microbiol. **69:**926–932.
6. **Carrillo, M., E. Estrada, and T. C. Hazen.** 1985. Survival and enumeration of the fecal indicators *Bifidobacterium adolescentis* and *Escherichia coli* in a tropical rain forest watershed. Appl. Environ. Microbiol. **50:**468–476.
7. **Carson, C. A., B. L. Shear, M. R. Ellersieck, and A. Asfaw.** 2001. Identification of fecal *Escherichia coli* from humans and animals by ribotyping. Appl. Environ. Microbiol. **67:**1503–1507.
8. **Centers for Disease Control and Prevention.** 2002. Surveillance for waterborne disease outbreaks—United States, 1999–2000. Morb. Mortal. Wkly. Rep. **51:**1–28
9. **Dombek, P. E., L. K. Johnson, S. T. Zimmerley, and M. J. Sadowsky.** 2000. Use of repetitive DNA sequences and the PCR to differentiate *Escherichia coli* isolates from human and animal sources. Appl. Environ. Microbiol. **66:**2572–2577.
10. **Dunbar, J., L. O. Ticknor, and C. R. Kuske.** 2000. Assessment of microbial diversity in four southwestern United States soils by 16S rRNA gene terminal restriction fragment analysis. Appl. Environ. Microbiol. **66:**2943–2950.
11. **Dunbar, J., L. O. Ticknor, and C. R. Kuske.** 2001. Phylogenetic specificity and reproducibility and new method for analysis of terminal restriction fragment profiles of 16S rRNA genes from bacterial communities. Appl. Environ. Microbiol. **67:**190–197.
12. **Field, K., E. C. Chern, L. K. Dick, J. Fuhrman, J. Griffith, P. A. Holden, M. G. LaMontagne, J. Le, B. Olson, and M. T. Simonich.** 2003. A comparative study of culture-independent, library-independent genotypic methods of fecal source tracking. J. Water Health **1:**181–194.
13. **Fiksdal, L., J. S. Maki, S. J. LaCroix, and J. T. Staley.** 1985. Survival and detection of *Bacteroides* spp., prospective indicator bacteria. tAppl. Environ. Microbiol. **49:**148–150.
14. **Fogarty, L. R., S. K. Haack, M. J. Wolcott, and R. L. Whitman.** 2003. Abundance and characteristics of the recreational water quality indicator bacteria *Escherichia coli* and enterococci in gull faeces. J. Appl. Microbiol. **94:**865–878.
15. **Hagedorn, C., S. L. Robinson, J. R. Filtz, S. M. Grubbs, T. A. Angier, and R. B. Reneau, Jr.** 1999. Determining sources of fecal pollution in a rural Virginia watershed with antibiotic resistance patterns in fecal streptococci. Appl. Environ. Microbiol. **65:**5522–5531.
16. **Harwood, V. J., J. Whitlock, and V. Withington.** 2000. Classification of antibiotic resistance patterns of indicator bacteria by discriminant analysis: Use in predicting the source of fecal contamination in subtropical waters. Appl. Environ. Microbiol. **66:**3698–3704.
17. **Kaplan, C. W., J. C. Astaire, M. E. Sanders, Bandaru, S. Reddy, and C. L. Kitts.** 2001. 16S ribosomal DNA terminal restriction fragment pattern analysis of bacterial communities in feces of rats fed *Lactobacillus acidophilus*. Appl. Environ. Microbiol. **67:**1935–1939.
18. **Kitts, C. L.** 2001. Terminal restriction fragment patterns: A tool for comparing microbial communities and assessing community dynamics. Curr. Issues Intest. Microbiol. **2:**17–25.
19. **Kreader, C. A.** 1998. Persistence of PCR-detectable *Bacteroides distasonis* from human feces in river water. Appl. Environ. Microbiol. **64:**4103–4105.
20. **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.
21. **Malakoff, D.** 2002. Water quality. Microbiologists on the trail of polluting bacteria. Science **295:**2352–2353.
22. **Marsh, T. L.** 1999. Terminal restriction fragment length polymorphism (T-RFLP): an emerging method for characterizing diversity among homologous populations of amplification products. Curr. Opin. Microbiol. **2:**323–327.
23. **McOrist, A. L., M. Jackson, and A. R. Bird.** 2002. A comparison of five methods for extraction of bacterial DNA from human faecal samples. J. Microbiol. Methods **50:**131–139.
24. **Nagashima, K., T. Hisada, M. Sato, and J. Mochizuki.** 2003. Application of new primer-enzyme combinations to terminal restriction fragment length polymorphism profiling of bacterial populations in human feces. Appl. Environ. Microbiol. **69:**1251–1262.
25. **Parveen, S., N. C. Hodge, R. E. Stall, S. R. Farrah, and M. L. Tamplin.** 2001. Phenotypic and genotypic characterization of human and nonhuman *Escherichia coli*. Water Res. **35:**379–386.
26. **Rhodes, M. W., and H. Kator.** 1988. Survival of *Escherichia coli* and *Salmonella* spp. in estuarine environments. Appl. Environ. Microbiol. **54:**2902–2907.
27. **Rose, J. B., and D. J. Grimes.** 2001. Reevaluation of microbial water quality: powerful new tools for detection and risk assessment. ASM Press, Washington, D.C.
28. **Salyers, A. A.** 1984. *Bacteroides* of the human lower intestinal tract. Annu. Rev. Microbiol. **38:**293–313.
29. **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.
30. **Stoeckel, D. M., M. V. Mathes, K. E. Hyer, C. Hagedorn, H. Kator, J. Lukasik, T. L. O'Brien, T. W. Fenger, M. Samadpour, K. M. Strickler, and B. A. Wiggins.** 2004. Comparison of seven protocols to identify fecal contamination sources using *Escherichia coli*. Environ. Sci. Technol. **38:**6109–6117.

31. **Suau, A., R. Bonnet, 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.

32. **Tannok, G. W.** 1997. Modification of the normal microbiota by diet, stress, antimicrobial agents, and probiotics, p. 434–465. In R. J. Mackie and B. A. White (ed.) Gastrointestinal microbiology, vol. 2: gastrointestinal microbes and host interactions. Academic Press, New York, NY.

33. **U.S. Environmental Protection Agency.** 2000. The quality of our nation's waters. EPA841–S–00–0011–20. U.S. Environmental Protection Agency, Washington, DC.

34. **U.S. Environmental Protection Agency.** 1986. Ambient water quality criteria for bacteria-1986. EPA440-5-84-002. U.S. Environmental Protection Agency, Washington, DC.

35. **Vispo, C., and W. H. Karasov.** 1997. The interaction of avian gut microbes and their host: an elusive symbiosis, p 116–155. *In* R. I. Mackie and B. A. White (ed.), Gastrointestinal microbiology, vol. 1: gastrointestinal ecosystems and fermentations. Academic Press, New York, NY.

36. **Wiggins, B. A., P. W. Cash, W. S. Creamer, S. E. Dart, P. P. Garcia, T. M. Gerecke, J. Han, B. L. Henry, K. B. Hoover, E. L. Johnson, K. C. Jones, J. G. McCarthy, J. A. McDonough, S. A. Mercer, M. J. Noto, H. Park, M. S. Phillips, S. M. Purner, B. M. Smith, E. N. Stevens, and A. K. Varner.** 2003. Use of antibiotic resistance analysis for representativeness testing of multiwatershed libraries. Appl. Environ. Microbiol. **69:**3399–3405.

37. **Wood, J., K. P. Scott, G. Avgustin, 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.