Environ. Sci. Technol. **2010,** *44,* 7777–7782

Use of Barcoded Pyrosequencing and Shared OTUs To Determine Sources of Fecal Bacteria in Watersheds

TATSUYA UNNO,[†] JEONGHWAN JANG,[†] DUKKI HAN,[†] JOON HA KIM,[†] MICHAEL J. SADOWSKY,[§] OK-SUN KIM,[∥] JONGSIK CHUN,[∥] AND HOR-GIL HUR*[,†,‡]

Department of Environmental Science and Engineering and International Environmental Research Center, Gwangju Institute of Science and Technology, Gwangju 500-712, Republic of Korea, Department of Soil, Water, and Climate and BioTechnology Institute, University of Minnesota, St. Paul, Minnesota 55108, and School of Biological Sciences, Seoul National University, NS70, 56-1 Daehak-dong, Kwanak-gu, Seoul 151-742, Republic of Korea

Received May 4, 2010. Revised manuscript received August 19, 2010. Accepted August 24, 2010.

While many current microbial source tracking (MST) methods rely on the use of specific molecular marker genes to identify sources of fecal contamination, these methods often fail to determine all point and nonpoint contributors of fecal inputs into waterways. In this study, we developed a new library-dependent MST method that uses pyrosequencing-derived shared operational taxonomy units (OTUs) to define sources of fecal contamination in waterways. A total 56,841 pyrosequencing reads of 16S rDNA obtained from the feces of humans and animals were evaluated and used to compare fecal microbial diversity in three freshwater samples obtained from the Yeongsan river basin in Jeonnam Province, South Korea. Sites included an urbanized agricultural area (Y1) (*Escherichia coli* counts ≥1600 CFU/100 mL), an open area (Y2) with no major industrial activities (940 CFU/100 mL), and a typical agricultural area (Y3) (≥1600 CFU/100 mL). Data analyses indicated that the majority of bacteria in the feces of humans and domesticated animals were comprised of members of the phyla *Bacteroidetes* or *Firmicutes*, whereas the majority of bacteria in wild goose feces and freshwater samples were classified to the phylum *Proteobacteria*. Analysis of OTUs shared between the fecal and environmental samples suggested that the potential sources of the fecal contamination at the sites were of human and swine origin. Quantification of fecal contamination was also examined by comparing the density of pyrosequencing reads in each fecal sample within shared OTUs. Taken together, our results indicated that analysis of shared OTUs derived from barcoded pyrosequencing reads provide the necessary resolution and discrimination to be useful as a next generation platform for microbial source tracking studies.

Introduction

According to the World Health Organization (WHO), diarrhea kills over 2 million people globally each year, and fecal contamination of water is one of the major contributing factors to disease incidence. Determining the sources of fecal contaminants in waterways using microbial source tracking methodologies has attracted much attention in recent years. However, current technologies often fail to consistently identify fecal contaminants in watersheds, prompting development of new and more sensitive methods (*1, 2*).

The application of 16S rDNA as a genetic marker for microbial source tracking has received recent attention, and markers purportedly specific for fecal bacteria (mainly *Bacteroidales*) originating from humans, bovine, and other animals have been developed and evaluated (*3*−*5*). Recently, however, the use of these marker genes has been questioned since there were large discrepancies in the performance of primers across different bovine populations, the lowest specificity was about 47%, and marker genes for all potential contamination sources are not available. In addition, several of the markers have been reported to have low sensitivity when used in quantitative PCR (qPCR) assays with environmental DNA (*6*), and there is generally a lack of correlation between the numbers of *Bacteriodales* present and counts of fecal indicator bacteria.

Ley et al. (*7*) used a 16S rDNA sequence library approach to examine how diet and phylogeny influences bacterial diversity. They reported that microbial diversity increases from carnivory to omnivory to herbivory, and these results indirectly support the contention that there is significant discrepancy in diversity and types of fecal bacteria in humans (omnivores) and livestock animals (herbivores). To date, several other studies have reported the use of culture-based 16S rDNA libraries to identify host specific microbial species in other animals, including *C. marimammalium* in gull feces (*8*), and *Lactobacillus sobrius/amylovorus* in pig manure (*9*).

While some of the identified marker genes hold great promise for assessing sources of fecal bacteria or feces in the environment, others have not proven useful due to issues of lack of sensitivity and cross-reactivity. This may in part be due to our lack of understanding of the broad microbial diversity present in intestinal environments, which is due, in part, to inherent deficiencies in using traditional culture based 16S rDNA library approaches. Studies of microbial diversity using reassociation kinetics for bacterial community DNA from pristine soils revealed that bacterial species diversity is more than 2 orders of magnitude greater than estimates based on using traditional culture dependent method (*10, 11*).

Pyrosequencing, a next generation sequencing technology that determines DNA sequences by using a "sequencing by synthesis" approach, has revolutionized the study of microbial diversity (*12*). While the first pyrosequencers, such as the Roche 454 FLX system, only produced average read-lengths of 100 bp, more recent advances have increased read lengths to about 400 bp, with considerable cost reduction (*13*). Pyrosequencing has been further improved by allowing for the simultaneous analysis of 1544 samples by using error-correcting barcoded primers (*14*). Short 16S rDNA pyrosequencing reads have proven useful for microbial community analysis (*15*) and has been widely applied to study microbial

* Corresponding author phone: +82-62-970-2437; fax: +82-62-970-2434; e-mail: hghur@gist.ac.kr. Corresponding author address: Department of Environmental Science and Engineering, Gwangju Institute of Science and Technology, Gwangju 500-712, South Korea.
† Department of Environmental Science and Engineering, Gwangju Institute of Science and Technology.
§ University of Minnesota.
∥ Seoul National University.
‡ International Environmental Research Center, Gwangju Institute of Science and Technology.

10.1021/es101500z
Published on Web 09/20/2010

diversity in soils (*16, 17*), fermented seafood (*18*), human saliva (*19*), wastewater treatment facilities (*20*), human and animals GI tracts (*21*−*25*), and the ecology of disease-producing microorganisms (*26*).

Dowd et al. (*27*) investigated microbial community structure in dairy cattle feces using a pyrosequencing approach. They reported that the majority of microorganisms revealed by using this approach were not found in previous studies using traditional culture-based methods. Since a barcoded pyrosequencing approach targeting bacterial 16S rDNA sequences has the necessary resolving power to reveal undiscovered microbial diversity in feces, this technique will likely prove to be useful to define future microbial markers for microbial source tracking applications.

In the studies reported here, we examined the microbial community structure in feces from humans and farm animals and further investigated the usefulness of barcoded pyrosequencing as a microbial source tracking tool. Here we report that pyrosequencing-derived shared operational taxonomy units (OTUs) and the density of pyrosequencing reads in each fecal sample within shared OTUs are useful tools to define sources of fecal contamination in waterways.

Materials and Methods

Fecal and Environmental DNA Extraction. A total 150 fecal swabs were collected from 30 randomly selected individual chickens, ducks, beef cattle, dairy cattle, and swine from three farms in Jeonnam Province, South Korea. Thirty fecal samples were obtained from wild geese resting near lake Gochunam in the same region. In addition, 30 fecal samples were obtained from 30 healthy human volunteers at the Gwangju Institute of Science and Technology, Gwangju, South Korea. All fecal swabs from identical source animals or humans were pooled into a single test tube and DNA was extracted using a stool DNA extraction kit (Bioneer Inc. South Korea). DNA was also extracted from 50 mL freshwater samples collected from three sites (Y1, Y2, and Y3) of the Yeongsan River, Jeonnam Province, South Korea using the same DNA extraction kit. The number, size, and description of samples used in this study are described in the Supporting Information and summarized in Table SI−S1.

Pyrosequencing. Genomic DNA extracted from fecal and environmental samples was amplified using primers barcoded with 8 nucleotides and targeting the V1 and V3 regions of the bacterial 16S rRNA as described previously (*28, 29*). All pyrosequencing reads obtained in this study were submitted to the Sequence Read Archive (SRA) under study accession number ERP000189.

Analyses for α Diversity of Fecal Bacteria Community. Since frequent pyrosequencing error was observed when reads were shorter than 300 bp, all reads shorter than 300 bp were removed from the entire analyses (data not shown). Operational taxonomy units (OTUs) were defined by using the CD-HIT program (*30*), qualitative species richness was estimated using Chao1 and ACE, and quantitative species richness/evenness was estimated by using nonparametric Shannon and Simpson indices by using the Mothur program v.1.8.0 (*31*).

Analyses of *β* Diversity of Fecal Bacteria Community. Species classification of each pyrosequencing read was carried out using the EzTaxon-e database (http://www.eztaxon-e.org) as described by Chun et al. (*32, 33*). Each pyrosequencing read was taxonomically assigned by using a combination of initial BLASTN-based searches (*34*) and pairwise similarity comparisons as described previously (*32*).

The relative abundance of each taxon was clustered by using Bionumerics version 6.0 software (Applied Maths, Sint Martens Latem, Belgium). The weighted UniFrac service (*35*) was used to analyze OTU divergence.

Determination of Pyrosequencing Reads Shared between Fecal and Environmental Samples. Shared OTUs between fecal and environmental samples were identified, at a dissimilarity level 0.03, by using the Mothur program. Differences in the number of reads among each fecal and environmental sample within shared OTUs were expressed as a density ratio.

Results and Discussion

Species Richness and Diversity in Fecal Bacteria Community Characterized by Barcoded Pyrosequencing. A total 56,841 pyrosequencing reads, comprised of 18,825 OTUs, were obtained from the human and animal fecal samples examined. Results in Table SI−S2 summarize the number of reads, OTUs, and indices that estimate diversity richness from the 16S rDNA samples analyzed. Relatively low species richness and diversity, at a dissimilarity level of 0.03, were found in the duck and goose fecal samples, followed by the human fecal sample (Table SI−S2 and Figure SI−S1). Approximately 20 to 30% of the pyrosequencing reads were classified at a 97% similarity level, except those from fecal samples obtained from beef and dairy cattle. Similarly, <10% of pyrosequencing reads obtained using cattle feces DNA matched sequences in publicly available 16S rDNA databases, suggesting that insufficient bacterial data from ruminants are present in current 16S rDNA libraries and databases.

About 0.04−0.07% of the reads from humans, chickens, and ducks, and 0.72% of reads from geese were identified as Eukaryota likely due to the nonspecificity of the primer sets (*36*). Reads identified as belonging to Eukaryota were greater in the environmental samples (1.3% for the Y1 site, 24.3% for the Y2 site, 0.9% for the Y3 site) than those from the fecal samples, and these were removed from subsequent analyses.

Comparison of Fecal Bacterial Diversity by Relative Abundance of Each Taxon. All 56,841 reads obtained from the fecal samples, and 12,430 reads that showed 100% query coverage and had ≥97% similarity, were clustered based on the relative abundance of each taxon (Figures SI−S2A and SI−S2B). Cluster profiling indicated that the chicken sample contained two topologically different clusters, and the OTUs detected in the chicken and duck fecal samples were distinctively different. However, cluster profiling of the reads having 100% query coverage and a similarity ≥97% showed that the bacterial community in the chicken and duck fecal samples clustered at a similarity level of approximately 60%.

Cluster profiling was also performed based on sequence diversity using classical Jaccard values calculated by the Mothur program (Figure SI−S2C). Unlike the clusters based on relative abundance, sequence divergence measurements separated the cattle samples from the remainder of the fecal samples analyzed. Since the poultry samples (chicken, duck, and goose) clustered together, as did those from beef and dairy cattle, these studies indicated that the OTU-based analysis better characterized difference among the fecal samples sharing similar intestinal environments than did clusters based on relative abundance (Figure SI−S2C).

Topological difference between the taxonomy- and OTU-based phylogenetic trees suggests that fecal source tracking results would be different using these two approaches. While the concept of using 100% query coverage and a similarity value ≥97% has been widely applied in screening pyrosequencing reads to assign taxa to microorganisms from various environments (*37, 38*), the relatively large number of taxonomically unassigned isolates using this strict set of criteria fails to adequately describe all microorganisms present. Thus, this approach should not be applied in microbial source tracking analyses since the data sets used need to uncover broader microbial diversity in feces.

Application of Barcoded Pyrosequencing Derived Shared OTUs for Microbial Source Tracking. Analyses of the relative

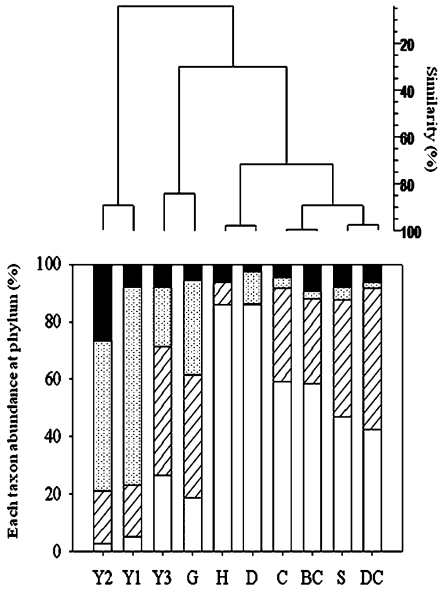

FIGURE 1. The relative abundance of phyla among OTUs from humans, animals, and environment samples: open square Firmicutes; striped square Bacteroidetes; dotted square Proteobacteria and; closed square others. Legend: H: humans; C: chickens; D: ducks; G: geese; BC: beef cattle; DC: dairy cattle; S: swine; Y1: down stream of the Yeongsan River; Y2: middle stream of the Yeongsan River; and Y3: up stream of the Yeongsan River.

abundance of each taxon at the different locations in the Yeongsan river basin were performed using all reads obtained from the fecal and environmental samples. However, due to the clear differences in cluster profiling of bacterial diversity between fecal and environmental samples (Figure 1), discrimination of the fecal samples based on this cluster profiling did not appear appropriate for our purposes. Thus, in order to better identify fecal bacteria in the environmental samples, we analyzed shared OTUs between fecal and environmental samples. Results in Table SI−S3 show that the majority of reads in the shared OTUs belonged to the phyla *Bacteroidetes, Firmicutes,* and *Proteobacteria* and were used for further analyses. Overall, the greatest numbers of the shared OTUs between fecal and environmental samples were identified with the human and swine fecal samples, followed by the fecal samples from geese. The shared OTUs between each fecal and environmental sample were also analyzed by using cluster profiling and the sample distance matrices using the UniFrac service. Results shown in Figure SI−S3 indicate that the fecal bacterial communities in swine and goose feces were most similar to the bacterial communities found at the Y1 and Y3 sites and the Y2 site, respectively. Results in Figure SI−S4 shows cluster analyses done with the shared OTUs analyzed for the phyla *Bacteroidetes, Firmicutes,* and *Proteobacteria*. Interestingly, the shared OTUs representing members of the phyla *Bacteroidetes* and *Firmicutes* in the environmental samples clustered with those from the human and/or swine fecal samples, whereas the shared OTUs from all the environmental samples clustered with the phylum *Proteobacteria* in the goose fecal sample.

Results in Table SI−S5 show the 10 most frequently observed genera in the *Proteobacteria* among the shared OTUs in the goose fecal and the environmental samples. To our knowledge, members of these genera were previously not known to be strict intestinal bacteria. Moreover, the sample distance matrix obtained for the three phyla indicated that microbial diversity between fecal and environmental samples within the phylum *Proteobacteria* were least related, compared to those in the phyla *Bacteroidetes and Firmicutes* (Table SI−S4).

Results presented in Figure 1 show that the OTUs in the goose fecal samples were different from those in the other fecal samples examined, mainly due to the presence of abundant *Proteobacterial* species and reduction in the *Firmicutes*. The goose fecal samples were obtained from migratory wild geese resting at lake Gochunam, in the same region where the Yeongsan River is located. Moreover, since the main diet for these migratory geese would likely be grasses and grain containing bacteria from soils and freshwater, they would be expected to have a GI flora different from domesticated animals fed a diet rich in grain and protein supplements. This assumption is supported by results shown in Figure 1, where the community structure of geese fecal bacteria is closely related to the bacterial community structure found in the freshwater. Therefore, the use of members of the phylum *Proteobacteria* may not be useful for microbial source tracking purposes, as the *Proteobacteria* found in the goose fecal samples may represent allochthonous microbes originating from the environment the geese live in rather than being members of the indigenous intestinal microbial community.

Analysis of the pyrosequencing data also indicated that the number of pyrosequencing reads needed for analysis of environmental samples could be dramatically reduced if primers targeting only members of the phyla *Bacteriodetes* or *Firmicutes* were used in these studies. Based on the rarefaction analyses done for reads classified to the phylum *Bacteroidetes*, less than approximately 3500 reads for one monitoring site may be sufficient to capture diversity and be microbial source tracking purposes (data not shown). On a cost and effort basis, reducing the number of pyrosequencing outputs for the environmental samples will indirectly allow for analysis of more sites. Dick et al. (*39*) developed a set of primers and a probe for 16S rDNA specific to *Bacteroidetes* that allows estimation of the number of *Bacteroidetes* in feces from humans, cows, dogs, cats, pigs, elk, deer, and gulls. These primers can be directly used for designing primers for pyrosequencing.

Results in Figure 2 show the distribution of phyla (Figure 2A), the number of the shared OTUs (Figure 2B), and the summation of the density ratio of pyrosequencing reads (Figure 2C) in the samples examined. Results in Table 1 show the percentages of shared OTUs, the number of reads, and a summation of density ratio of each fecal sample found at sites Y1, Y2, and Y3. Details of the formula used for calculations are shown in the Supporting Information. The phylum distribution of microorganisms from the site Y2 sample showed that approximately 60% of the shared OTUs were classified into the phylum *Proteobacteria* (Figure 2A). Moreover, the percentage of the shared OTUs and reads in the site Y2 sample were only 2.2 and 5.0%, respectively, which was significantly lower than that found at the other two sites (Table 1). This suggests that the Y2 site is likely not directly contaminated with fecal material originating from humans or animals. It should be noted, however, the presence of *E. coli* at this site suggested that these bacteria became naturalized to site Y2 at some time in the past.

The use of *E. coli* as a fecal indicator bacterium has been widely applied for water quality monitoring. However, recent studies have shown that *E. coli* can be an autochthonous member of the microbial community in soil and sand environments (*40*−*42*), suggesting that a new indicator bacterium is needed to monitor water quality. In the study reported here, only a few or no pyrosequencing reads in each of the fecal samples examined were classified as *E. coli*; this is in agreement with previous reports (*27*) that *E. coli* typically comprises <1% of the total intestinal bacterial population (*43*). Our results, however, also suggest that a

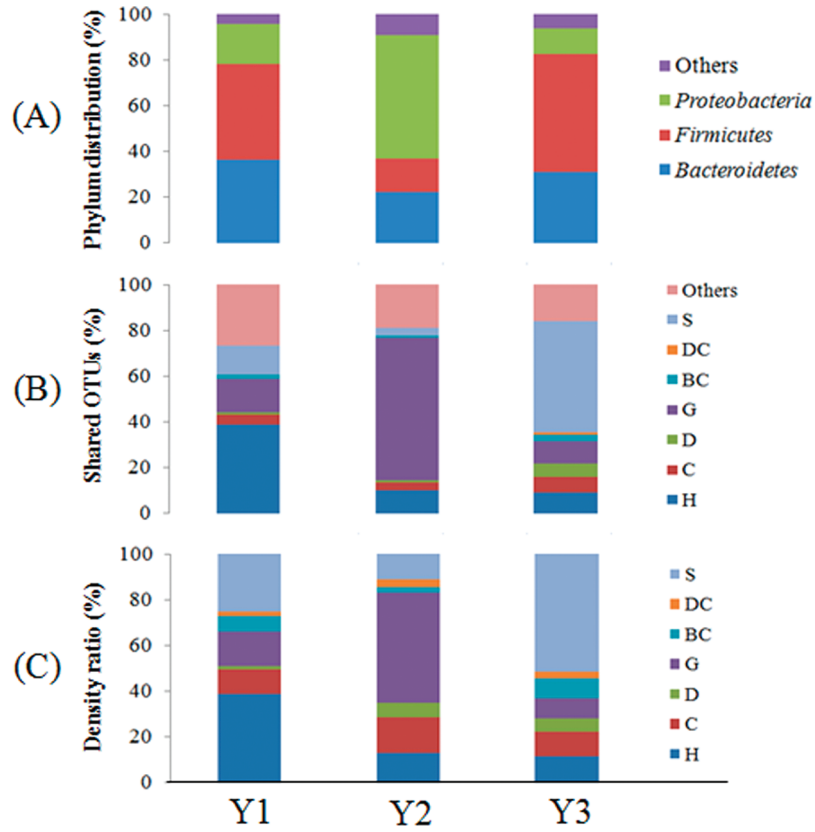

FIGURE 2. Percentage of phylum (A), shared OTUs (B), and total density ratio (C) within shared OTUs between fecal and environmental sample. Legend: H: humans; C: chickens; D: ducks; G: geese; BC: beef cattle; DC: dairy cattle; S: Slwine; Y1: down stream of the Yeongsan River; Y2: middle stream of the Yeongsan River; and Y3: up stream of the Yeongsan River.

TABLE 1. Suggestion of Fecal Contamination Parameter Using Results Obtained from the Barcoded Pyrosequencing

	sites in the Yeongsan River		
suggested new fecal contamination parameter	Y1	Y2	Y3
no. of shared OTUs (%)	173 (7.2)	90 (2.2)	731 (8.7)
no. of reads (%)	1806 (17.0)	544 (5.0)	3443 (11.8)
total density ratio	82.9	38.5	230.0

better indicator of water quality, and thus a better monitoring tool, would be the number of OTUs shared with intestinal microbiota revealed by deep sequencing approaches.

The number of OTUs shared between each fecal and environmental samples is summarized in Figure 2B. Only the shared OTUs found in one source animal group and environmental sample were counted, and if more than two fecal source animals (e.g., cows and humans) were found in one shared OTU, then these shared OTUs were counted as "others". In the study presented here, <30% of the shared OTUs were from more than two fecal samples, indicating that removing these shared OTUs will not significantly negatively impact the analyses. This result qualitatively suggests that the major source of fecal contamination in the environment, and the percentage of the shared OTUs for the environmental samples, may be a useful tool to indicate the degree of fecal contamination. For example, 7.2% and 8.7% of the total OTUs were identified as shared OTUs in sites Y1 and Y3, respectively (Table 1), and sites Y1 and Y3 were contaminated with human (39%) and swine (51%) feces, respectively (Figure 2C). In contrast, only 2.2% of the total OTUs were identified as shared OTUs in the site Y2 sample (Table 1). However, and as described above, most of the shared OTUs found in the goose fecal sample likely originated from the environment (Table SI−S5), thus the contamination in site Y2 may not be due to fecal contamination but rather to authoctonous bacteria.

The density ratio between each fecal and environmental sample was calculated based on number of pyrosequencing reads identified in the shared OTUs (Figure 2C). For example, since 10, 6, and 2 reads from the human and chicken fecal samples, and the site Y2 sample, respectively, were found in one shared OTU, the density ratio given to humans for this shared OTU was 0.19 ((2/10,857)/(10/10,358)) and 0.27 ((2/10,857/(6/8,761)) for chickens. If the density ratio of a shared OTU in the environmental sample was greater than that found for each fecal sample, the ratio was removed from the analysis. Based on this analysis, total density ratios (Σdensity ratio given for each shared OTU) of 82.9, 38.5, and 230 were obtained from the site Y1, Y2, and Y3 samples, respectively (Table 1). This ratio may provide a more valid estimation of the degree of fecal contamination, because bacteria originating from the environment having shared OTUs that were likely excluded from the calculation. Taken together, results of this analysis indicated that the potential sources of fecal contamination were humans and swine at sites Y1 and Y3, respectively. Moreover, the relative degree of fecal contamination at the sites was Y3 > Y1 ≫ Y2. Taxonomic classification of the shared OTUs was performed, and the 10 most frequently observed genera were analyzed further. Results in Table SI−S6 show the comparison of each fecal sample based on the density of pyrosequencing reads (number of reads within shared OTUs/number of reads in the fecal sample for genus X). For example, if 265 reads were classified to the genus *Prevotella* in the human fecal sample and 15 reads from the human fecal sample were detected within the OTUs that were solely shared between the Y1 site sample and human fecal sample, a density of 0.057 (15/265) was given to the human fecal sample. Results shown in Table

SI−S6 indicate that that the site Y1 sample contained a greater density of genera than were observed in the human fecal sample, except for *Pseudomonas*, *Ruminococcaceae_uc*, and *Veillonellaceae_uc*. This suggests that the source of the fecal contamination at site Y1 was likely of human origin. Similarly, the source of fecal contamination at Site Y2 was likely due to geese and site Y3 contained genera indicative of swine fecal contamination, except for *Pseudomonas, Bacteroides, Faecalibacterium,* and *Roseburia* strains.

Among all the environmental samples examined, the genus *Prevotella* was found most frequently in the shared OTUs. The use of primers specific for the genus *Prevotella* has been applied to identify and quantify sources of fecal contamination in freshwater environments (*44*−*46*). However, the shared OTUs found in the site Y2 sample suggest that the use of single genus for the identification of source of fecal contamination may not be appropriate. Approximately 3080 reads obtained from the human fecal sample were classified as *Faecalibacterium prausnitzii*, whereas only two reads of this OTU were found in the other fecal samples tested in this study. While this suggests that *Faecalibacterium prausnitzii* may represent a good indicator of human fecal contamination, more research is needed to determine if its presence in the environment correlates with human fecal contamination across wide geographic areas and the degree to which primers designed for this bacterium cross-react with feces from other animal species.

Rather than identifying the source of fecal contamination using existing specific source tracking genes (i.e, *Bacteroidales*-specific marker genes), in the studies presented here we report on the use of pyrosequencing as a new MST tool, that simultaneously uses multiple source markers that are defined by shared OTUs. In this way, we were able to evaluate the source of fecal contamination simultaneously using several bacterial taxa, which we feel has several advantages over the existing source tracking methods that use a limited number of specific target genes as main source tracking criteria. Moreover, since the number of shared OTUs, or the density ratio of reads within shared OTUs, can be used to indicate the source of fecal contamination, as presented in Figure 2B and 2C, the method described here allows for the determination of the source(s) of fecal contamination without assigning taxonomic classification to each pyrosequencing reads. This has obvious advantages given that a large proportion of fecal and environmental bacteria have unknown taxonomic status.

Recently, animal to animal variation in fecal microbial diversity among beef cattle was reported (*47*). Thus, in the study reported here, we attempted to limit the influence of individual variation in the microbial communities on our analyses by using a mixed fecal bacterial community from pooled fecal samples from 30 individuals obtained from three farms for sources. Despite this, however, further studies are needed to determine how this variation may influence the metagenomic analyses done in this study. Nevertheless, to our knowledge, this is the first report of the use of pyrosequencing-derived shared OTUs to determine sources of fecal bacteria in the environment. Our analyses were facilitated by the use of the Mothur Program. Since all outputs from the Mothur program are in text format and analyses can be automated with simple scripting programs, such as Perl and Python, this method may be easily adaptable to analyze even larger data sets. If verified by others, the pyrosequencing-derived shared OTUs based microbial source tracking method described here may prove to be a useful new tool to replace the many of currently used labor intensive microbial source tracking methods. This method may become even more widely used in the future as costs associated with next generation sequencing continue to decrease at a rapid rate.

Acknowledgments

This work was supported by grants from Korean Ministry of Environment as Eco-Technopia 21 project (900-20080010), Korea, and the UNU-GIST Joint Programme on Science and Technology for Sustainability operated by International Environmental Research Center, GIST, Korea.

Supporting Information Available

Experimental details, Figures SI-S1−SI-S4, and Tables SI-S1−SI-S6. This material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited

(1) Scott, T. M.; Rose, J. B.; Jenkins, T. M.; Farrah, S. R.; Lukasik, J. Microbial source tracking: current methodology and future directions. *Appl. Environ. Microbiol.* **2002**, *68* (12), 5796–803.
(2) Griffith, J. F.; Weisberg, S. B.; McGee, C. D. Evaluation of microbial source tracking methods using mixed fecal sources in aqueous test samples. *J. Water Health* **2003**, *1* (4), 141–51.
(3) Bernhard, A. E.; Field, K. G. 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.* **2000**, *66* (10), 4571–4.
(4) Dick, L. K.; Bernhard, A. E.; Brodeur, T. J.; Santo Domingo, J. W.; Simpson, J. M.; Walters, S. P.; Field, K. G. Host distributions of uncultivated fecal Bacteroidales bacteria reveal genetic markers for fecal source identification. *Appl. Environ. Microbiol.* **2005**, *71* (6), 3184–91.
(5) Layton, A.; McKay, L.; Williams, D.; Garrett, V.; Gentry, R.; Sayler, G. Development of Bacteroides 16S rRNA gene TaqMan-based real-time PCR assays for estimation of total, human, and bovine fecal pollution in water. *Appl. Environ. Microbiol.* **2006**, *72* (6), 4214–24.
(6) Shanks, O. C.; White, K.; Kelty, C. A.; Hayes, S.; Sivaganesan, M.; Jenkins, M.; Varma, M.; Haugland, R. A. Performance Assessment PCR-Based Assays Targeting Bacteroidales Genetic Markers of Bovine Fecal Pollution. *Appl. Environ. Microbiol.* **2010**, *76* (5), 1359–66.
(7) Ley, R. E.; Hamady, M.; Lozupone, C.; Turnbaugh, P. J.; Ramey, R. R.; Bircher, J. S.; Schlegel, M. L.; Tucker, T. A.; Schrenzel, M. D.; Knight, R.; Gordon, J. I. Evolution of mammals and their gut microbes. *Science* **2008**, *320* (5883), 1647–51.
(8) Lu, J.; Santo Domingo, J. W.; Lamendella, R.; Edge, T.; Hill, S. Phylogenetic diversity and molecular detection of bacteria in gull feces. *Appl. Environ. Microbiol.* **2008**, *74* (13), 3969–76.
(9) Marti, R.; Dabert, P.; Ziebal, C.; Pourcher, A. M. Evaluation of Lactobacillus sobrius/L. amylovorus as a New Microbial Marker of Pig Manure. *Appl. Environ. Microbiol.* **2010**, *76* (5), 1456–61.
(10) Armougom, F.; Raoult, D. Exploring microbial diversity using 16S rRNA high-throughput methods. *J. Comput. Sci. Syst. Biol.* **2009**, *2*, 074−092.
(11) Gans, J.; Wolinsky, M.; Dunbar, J. Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science* **2005**, *309* (5739), 1387–90.
(12) Margulies, M.; Egholm, M.; Altman, W. E.; Attiya, S.; Bader, J. S.; Bemben, L. A.; Berka, J.; Braverman, M. S.; Chen, Y. J.; Chen, Z.; Dewell, S. B.; Du, L.; Fierro, J. M.; Gomes, X. V.; Godwin, B. C.; He, W.; Helgesen, S.; Ho, C. H.; Irzyk, G. P.; Jando, S. C.; Alenquer, M. L.; Jarvie, T. P.; Jirage, K. B.; Kim, J. B.; Knight, J. R.; Lanza, J. R.; Leamon, J. H.; Lefkowitz, S. M.; Lei, M.; Li, J.; Lohman, K. L.; Lu, H.; Makhijani, V. B.; McDade, K. E.; McKenna, M. P.; Myers, E. W.; Nickerson, E.; Nobile, J. R.; Plant, R.; Puc, B. P.; Ronan, M. T.; Roth, G. T.; Sarkis, G. J.; Simons, J. F.; Simpson, J. W.; Srinivasan, M.; Tartaro, K. R.; Tomasz, A.; Vogt, K. A.; Volkmer, G. A.; Wang, S. H.; Wang, Y.; Weiner, M. P.; Yu, P.; Begley, R. F.; Rothberg, J. M. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **2005**, *437* (7057), 376–80.
(13) Rothberg, J. M.; Leamon, J. H. The development and impact of 454 sequencing. *Nat. Biotechnol.* **2008**, *26* (10), 1117–24.
(14) Hamady, M.; Walker, J. J.; Harris, J. K.; Gold, N. J.; Knight, R. Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat. Methods* **2008**, *5* (3), 235–7.
(15) Liu, Z.; Lozupone, C.; Hamady, M.; Bushman, F. D.; Knight, R. Short pyrosequencing reads suffice for accurate microbial community analysis. *Nucleic Acids Res.* **2007**, *35* (18), e120.
(16) Buee, M.; Reich, M.; Murat, C.; Morin, E.; Nilsson, R. H.; Uroz, S.; Martin, F. 454 Pyrosequencing analyses of forest soils reveal

an unexpectedly high fungal diversity. *New Phytol.* **2009**, *184* (2), 449–56.
(17) Jones, R. T.; Robeson, M. S.; Lauber, C. L.; Hamady, M.; Knight, R.; Fierer, N. A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *ISME J.* **2009**, *3* (4), 442–53.
(18) Roh, S. W.; Kim, K. H.; Nam, Y. D.; Chang, H. W.; Park, E. J.; Bae, J. W. Investigation of archaeal and bacterial diversity in fermented seafood using barcoded pyrosequencing. *ISME J.* **2009**, *4* (1), 1–16.
(19) Nasidze, I.; Quinque, D.; Li, J.; Li, M.; Tang, K.; Stoneking, M. Comparative analysis of human saliva microbiome diversity by barcoded pyrosequencing and cloning approaches. *Anal. Biochem.* **2009**, *391* (1), 64–8.
(20) Sanapareddy, N.; Hamp, T. J.; Gonzalez, L. C.; Hilger, H. A.; Fodor, A. A.; Clinton, S. M. Molecular diversity of a North Carolina wastewater treatment plant as revealed by pyrosequencing. *Appl. Environ. Microbiol.* **2009**, *75* (6), 1688–96.
(21) Andersson, A. F.; Lindberg, M.; Jakobsson, H.; Backhed, F.; Nyren, P.; Engstrand, L. Comparative analysis of human gut microbiota by barcoded pyrosequencing. *PLoS One* **2008**, *3* (7), e2836.
(22) Mshvildadze, M.; Neu, J.; Shuster, J.; Theriaque, D.; Li, N.; Mai, V. Intestinal microbial ecology in premature infants assessed with non-culture-based techniques. *J. Pediat.r* **2010**, *156* (1), 20–5.
(23) Turnbaugh, P. J.; Hamady, M.; Yatsunenko, T.; Cantarel, B. L.; Duncan, A.; Ley, R. E.; Sogin, M. L.; Jones, W. J.; Roe, B. A.; Affourtit, J. P.; Egholm, M.; Henrissat, B.; Heath, A. C.; Knight, R.; Gordon, J. I. A core gut microbiome in obese and lean twins. *Nature* **2009**, *457* (7228), 480–4.
(24) Armougom, F.; Raoult, D. Use of pyrosequencing and DNA barcodes to monitor variations in Firmicutes and Bacteroidetes communities in the gut microbiota of obese humans. *BMC Genomics* **2008**, *9*, 576.
(25) Armougom, F.; Henry, M.; Vialettes, B.; Raccah, D.; Raoult, D. Monitoring bacterial community of human gut microbiota reveals an increase in Lactobacillus in obese patients and Methanogens in anorexic patients. *PLoS One* **2009**, *4* (9), e7125.
(26) Wolcott, R. D.; Gontcharova, V.; Sun, Y.; Dowd, S. E. Evaluation of the bacterial diversity among and within individual venous leg ulcers using bacterial tag-encoded FLX and titanium amplicon pyrosequencing and metagenomic approaches. *BMC Microbiol.* **2009**, *9*, 226.
(27) Dowd, S. E.; Callaway, T. R.; Wolcott, R. D.; Sun, Y.; McKeehan, T.; Hagevoort, R. G.; Edrington, T. S. Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC Microbiol.* **2008**, *8*, 125.
(28) Chakravorty, S.; Helb, D.; Burday, M.; Connell, N.; Alland, D. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *J. Microbiol. Methods* **2007**, *69* (2), 330–9.
(29) Chun, J.; Kim, K. Y.; Lee, J. H.; Choi, Y. The Analysis of Oral Microbial Communities of Wild-type and Toll-Like Receptor 2-deficient Mice Using a 454 GS FLX Titanium Pyrosequencer. *BMC Microbiol.* **2010**, *10* (1), 101.
(30) Li, W.; Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **2006**, *22* (13), 1658–9.
(31) Schloss, P. D.; Westcott, S. L.; Ryabin, T.; Hall, J. R.; Hartmann, M.; Hollister, E. B.; Lesniewski, R. A.; Oakley, B. B.; Parks, D. H.; Robinson, C. J.; Sahl, J. W.; Stres, B.; Thallinger, G. G.; Van Horn, D. J.; Weber, C. F. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **2009**, *75* (23), 7537–41.
(32) Chun, J.; Lee, J. H.; Jung, Y.; Kim, M.; Kim, S.; Kim, B. K.; Lim, Y. W. EzTaxon: a web-based tool for the identification of prokaryotes based on 16S ribosomal RNA gene sequences. *Int. J. Syst. Evol. Microbiol.* **2007**, *57* (Pt 10), 2259–61.
(33) Chun, J.; Kim, K. Y.; Lee, J. H.; Choi, Y. The Analysis of Oral Microbial Communities of Wild-type and Toll-Like Receptor 2-deficient Mice Using a 454 GS FLX Titanium Pyrosequencer. *BMC Microbiol.* **2010**, *10* (1), 101.
(34) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215* (3), 403–10.
(35) Lozupone, C.; Hamady, M.; Knight, R. UniFrac--an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinf.* **2006**, *7*, 371.
(36) Huys, G.; Vanhoutte, T.; Joossens, M.; Mahious, A. S.; De Brandt, E.; Vermeire, S.; Swings, J. Coamplification of eukaryotic DNA with 16S rRNA gene-based PCR primers: possible consequences for population fingerprinting of complex microbial communities. *Curr. Microbiol.* **2008**, *56* (6), 553–7.
(37) Soininen, E. M.; Valentini, A.; Coissac, E.; Miquel, C.; Gielly, L.; Brochmann, C.; Brysting, A. K.; Sonstebo, J. H.; Ims, R. A.; Yoccoz, N. G.; Taberlet, P. Analysing diet of small herbivores: the efficiency of DNA barcoding coupled with high-throughput pyrosequencing for deciphering the composition of complex plant mixtures. *Front Zool.* **2009**, *6*, 16.
(38) Valentini, A.; Miquel, C.; Nawaz, A. M.; Bellemain, E.; Coissac, E.; Pompanon, F.; Gielly, L.; Cruaud, C.; Nascetti, G.; Wincker, P.; Swenson, E. J.; Taberlet, P. New perspective in diet analysis based on DNA barcoding and parallel pyrosequencing: the *trn*L approach. *Mol. Ecol. Resour.* **2009**, *9*, 51–60.
(39) Dick, L. K.; Field, K. G. Rapid estimation of numbers of fecal Bacteroidetes by use of a quantitative PCR assay for 16S rRNA genes. *Appl. Environ. Microbiol.* **2004**, *70* (9), 5695–7.
(40) Brennan, F. P.; Abram, F.; Chinalia, F. A.; Richards, K. G.; O'Flaherty, V. Characterization of Environmentally Persistent Escherichia coli Isolates Leached from an Irish Soil. *Appl. Environ. Microbiol.* **2010**, *76* (7), 2175–80.
(41) Ishii, S.; Ksoll, W. B.; Hicks, R. E.; Sadowsky, M. J. Presence and growth of naturalized Escherichia coli in temperate soils from Lake Superior watersheds. *Appl. Environ. Microbiol.* **2006**, *72* (1), 612–21.
(42) Walk, S. T.; Alm, E. W.; Calhoun, L. M.; Mladonicky, J. M.; Whittam, T. S. Genetic diversity and population structure of Escherichia coli isolated from freshwater beaches. *Environ. Microbiol.* **2007**, *9* (9), 2274–88.
(43) Drasar, B. S.; Barrow, P. *Intestinal Microbiology*; Wokingham, UK, Van Nostrand Reinhold: 1985.
(44) Okabe, S.; Okayama, N.; Savichtcheva, O.; Ito, T. Quantification of host-specific Bacteroides-Prevotella 16S rRNA genetic markers for assessment of fecal pollution in freshwater. *Appl. Microbiol. Biotechnol.* **2007**, *74* (4), 890–901.
(45) Savichtcheva, O.; Okabe, S. Qualitative and quantitative estimation of host-specific fecal pollution using Bacteroides-Prevotella 16S rRNA genetic markers by T-RFLP and real-time PCR analyses. *Water Sci. Technol.* **2009**, *59* (9), 1831–40.
(46) Fogarty, L. R.; Voytek, M. A. Comparison of Bacteroides-Prevotella 16S rRNA genetic markers for fecal samples from different animal species. *Appl. Environ. Microbiol.* **2005**, *71* (10), 5999–6007.
(47) Durso, L. M.; Harhay, G. P.; Smith, T. P.; Bono, J. L.; Desantis, T. Z.; Harhay, D. M.; Andersen, G. L.; Keen, J. E.; Laegreid, W. W.; Clawson, M. L. Animal-to-animal variation in fecal microbial diversity among beef cattle. *Appl. Environ. Microbiol.* **2010**, *76* (14), 4858–62.

ES101500Z