Geographic sharing of ribotype patterns in *Enterococcus faecalis* for bacterial source tracking

Peter G. Hartel[1], Samuel P. Myoda[2], Kerry J. Ritter, Robin L. Kuntz[1], Karen Rodgers[1], James A. Entry[3], Sheryl A. Ver Wey[3], Eduardo C. Schröder[4], Juan Calle[4], Mercedes Lacourt[4], Janice E. Thies[5], John P. Reilly[5] and Jeffry J. Fuhrmann[6]

Abstract

The limited host range of *Enterococcus faecalis* may reduce its clonal diversity and thereby increase its geographic sharing of ribotype patterns. Such sharing would be advantageous for bacterial source tracking (BST). We determined the geographic sharing of ribotype patterns in 752 *E. faecalis* isolates obtained primarily from wastewater treatment plants in Delaware (15 locations; 490 isolates), Georgia (2 locations; 48 isolates), Idaho (1 location; 118 isolates), New York (2 locations; 48 isolates), and Puerto Rico (2 locations; 48 isolates). Isolates were ribotyped with a RiboPrinter. When pooled across all locations and analyzed at a similarity index of 100% and a tolerance level of 1.00%, the 752 *E. faecalis* isolates yielded 652 different ribotypes, of which 429 (66%) were unshared. Even when the matching criterion was relaxed by decreasing the tolerance level from 1% to 10% or lowering the similarity cutoff from 100% to 90%, half or almost half of the ribotypes were unshared. A Mantel test of zero correlation showed no statistically significant correlation between ribotype patterns and geographic distance among the 32 samples (one location at one time) at either the 1.00% ($p = 0.91$) or 10.00% ($p = 0.83$) tolerance levels. Therefore, the percentage of ribotype patterns shared between two locations did not increase as the distance between locations decreased. In the case of BST, a permanent host origin database sufficiently large to encompass these ribotype patterns would be time-consuming and expensive to construct.

Introduction

It appears, at least in the case of the fecal indicator bacterium *Escherichia coli*, that subspecies of this bacterium change with water flow (Hartel *et al.* 2004), time (Jenkins *et al.* 2003), animal diet (Hartel *et al.* 2003), habitat (Gordon *et al.* 2002), and geography (Buchan *et al.* 2001, Hartel *et al.* 2002). These results discourage researchers in bacterial source tracking (BST) who want to create a permanent host origin database to match *E. coli* isolates from environmental sources to *E. coli* isolates from known warm-blooded animal species (e.g., Parveen *et al.* 1999, Dombek *et al.* 2000) because they suggest that the host origin database must contain many thousands of isolates in order for a reasonable amount of matching to occur (Johnson *et al.* 2004). Such a permanent host origin database is time-consuming and expensive to construct.

However, constructing such a large and expensive host origin database may not be necessary if the bacterium has a limited host range. If a bacterial species is associated with only a few warm-blooded host animal species, then the bacterium may have less clonal diversity than a bacterium such as *E. coli*, which is associated with many warm-blooded host animal species. *Enterococcus faecalis* may be such a bacterium. When isolated phenotypically, the bacterium is commonly found in the feces of humans and birds, but rarely or not at all in the feces of cattle, pigs, horses, sheep, and rabbits (Pourcher *et al.*

[1] *University of Georgia, Department of Crop and Soil Sciences, Athens, GA*
[2] *Delaware Department of Natural Resources and Environmental Control, Division of Water Resources, Dover, DE*
[3] *USDA–ARS, Northwest Irrigation and Soils Research Laboratory, Kimberly, ID*
[4] *University of Puerto Rico, Department of Agronomy, Mayagüez, PR*
[5] *Cornell University, Department of Crop and Soil Sciences, Ithaca, NY*
[6] *University of Delaware, Department of Plant and Soil Sciences, Newark, DE*

1991). Wheeler *et al.* (2002) determined that the bacterium was also not found in deer feces. Even in the case of birds, not all birds harbor *E. faecalis*. For example, the bacterium is rare in pigeons (Baele *et al.* 2002). Agricultural practices may reduce the clonal diversity of *E. faecalis* further. For example, *E. faecalis* is commonly found in chicken feces, but rarely in chicken litter (Kuntz *et al.* 2004). In this case, the unfavorable moisture content of the litter (<30%) and its narrow C:N ratio (that results in self-composting and subsequent high temperature) probably limit the survival of the bacterium (Hartel *et al.* 2000).

Construction of a permanent host origin database for *E. faecalis*, as for *E. coli*, requires an understanding of subspecies changes with respect to water flow, time, animal diet, habitat, and geography. Although no research of this type has been conducted on *E. faecalis* specifically, research has been conducted with the fecal enterococci. Wiggins *et al.* (2003), using antibiotic resistance analysis, determined the geographic and temporal changes of the fecal enterococci by combining two host origin databases, one from Florida and the other from Virginia, and tested the ability of the combined database to discriminate among humans, domesticated animals, and wildlife. Because the fecal enterococci appeared to change with geography, but not with time, they concluded that databases of fecal enterococci still need to be comprised of isolates from specific regions, but the time when the isolates were collected was less important. However, the fecal enterococci still comprise many species (Manero and Blanch 1999), and it is difficult to apply these conclusions to a single enterococcal species such as *E. faecalis*.

Therefore, we conducted a study to determine changes in *E. faecalis* subspecies with geography. Samples were obtained primarily from wastewater treatment plants in Delaware, Georgia, Idaho, New York, and Puerto Rico. Based on loading rates, virtually all the *E. faecalis* isolates should originate from humans. The BST method was ribotyping, automated with a RiboPrinter.

Methods

Sampling Locations

With the exception of Delaware, only wastewater treatment plants were sampled at various locations in Georgia, Idaho, New York, and Puerto Rico (Figure 1). In Georgia, New York, and Puerto Rico, at least two different wastewater treatment plants were sampled. In Idaho, only one wastewater treatment plant was sampled, but it was sampled at 0, 2, 5, 7, and 86 days. In Delaware, one latrine (Bellevue), one cesspool (Dover), and 13 wastewater treatment plants were sampled. Of the 13 Delaware wastewater treatment plants, four (Harrington, Kent, Port Penn, and Wilmington) were sampled at two different times. The distance between any two locations was determined at http://www.indo.com/distance.

Fecal Enterococcal Sampling

Except for Delaware samples obtained from Bellevue, Kent, and Wilmington, all samples were grab samples. Kent and Wilmington samples were 24-hour composite samples obtained with an autosampler (ISCO, Lincoln, NE). In the case of the Bellevue latrine, the sample was taken with a sterile tongue depressor, and the depressor was mixed with 50 ml of sterile water. In the case of the Dover cesspool, the samples were obtained only

Figure 1. Sampling sites in Delaware, Georgia, Idaho, New York, and Puerto Rico. All sites are wastewater treatment plants, except Delaware locations Bellevue (latrine) and Dover (cesspool).

minutes apart at three locations approximately 3 m from each other. All samples were stored in sterile containers on ice in a cooler and were processed within 6 hours. Samples were processed with the Enterolert system (IDEXX Laboratories, Westbrook, ME). Briefly, samples were serially diluted with sterile distilled water to 10^{-1}, 10^{-2}, 10^{-3}, and 10^{-4} in sterile, manufacturer-supplied polystyrene bottles. A package of powdered Enterolert reagent was added to each polystyrene bottle. After the reagent was dissolved, the contents of each bottle were added to a Quanti-tray, a sterile disposable panel containing 97 wells. Each Quanti-tray was mechanically sealed, which distributed the sample uniformly into the wells. Each Quanti-tray was incubated for 24 hours at 41±0.5°C. Fluorescing (positive) wells were counted under a 365-nm UV light. The number of positive wells was converted to a Most Probable Number (MPN) value based on the dilution factor and the manufacturer-supplied MPN tables.

Obtaining *E. faecalis* Subspecies

The tops of positive Quanti-tray wells were labeled with an acetate marker and, if necessary, the Quanti-trays were shipped overnight to the University of Georgia for analysis. All positive wells, beginning with the highest dilution and working towards the lowest dilution, were sampled until between 66 and 180 wells from each location were tested. This sampling protocol should have recovered the most common *E. faecalis* subspecies from each sample. To obtain isolates of enterococci from each well, positive wells were identified on the back of the Quanti-tray with an acetate marker. The back of each Quanti-tray was surface-disinfected with 70% ethanol, and each positive well was punctured with a sterile pipet tip. A 10-µL portion was removed from each positive well, and each portion was streaked onto a 5-cm plate containing Enterococcosel agar (Becton Dickinson, Sparks, MD). Plates were incubated in Ziploc bags (DowBrands, Indianapolis, IN) at 37°C for 48 hours.

One black (indicating esculin hydrolysis) isolated colony was randomly picked from each Enterococcosel agar plate with a sterile plastic stab. Isolates were speciated as described by Wheeler *et al.* (2002). Briefly, each isolate was suspended in 125 µl of sterile saline–phosphate buffer contained in a well of a 96-well microtiter plate. Three random wells were reserved for an American Type Culture Collection (ATCC, Manassas, VA) control isolate, *E. faecalis* ATCC #19433, and three random wells were left uninoculated. Each isolate was inoculated with a replicator (Sigma, St. Louis, MO) into separate microtiter plates containing Brain Heart Infusion (BHI) agar (Difco Laboratories, Sparks, MD) with 6.5% NaCl, arginine hydrolysis medium (Difco) with and without arginine, and modified pyruvate and arabinose carbon utilization media (both Sigma). Plates were incubated at 37°C and reactions were recorded after 24 and 48 hours. Each isolate was observed under phase contrast microscopy to ensure that the cells were spherical or ovoid. A catalase test with 8.82 *M* H_2O_2 was performed in the remaining saline phosphate-cell suspension to ensure that each isolate was catalase negative. Isolates that grew in or on the two media designed specifically for fecal enterococci, the Enterolert system (bacteria possess the enzyme b-glucosidase) and Enterococcosel agar (bacteria resist sodium azide and hydrolyze esculin), were catalase negative, were spherical or ovoid in cell morphology, grew in BHI agar with 6.5% NaCl, hydrolyzed arginine, and utilized pyruvate but not arabinose, were considered *E. faecalis*. These tests are consistent with traits associated with *E. faecalis* in Standard Methods for the Examination of Water and Wastewater (Clesceri *et al.* 1998). Isolates with test results identical to *E. faecalis*, but with a different carbohydrate utilization pattern or arginine hydrolysis result, were considered to be other enterococci. Isolates that grew in or on a medium selective for fecal enterococci, but did not have round or ovoid cells, were not catalase negative, or did not grow in BHI agar with 6.5% NaCl, were considered to be unknown bacteria, regardless of their carbohydrate utilization pattern or arginine hydrolysis result. In addition, 88 isolates identified as *E. faecalis* were randomly selected and their speciation verified with a more extensive biochemical key (Manero and Blanch 1999). All isolates were confirmed as *E. faecalis*. All *E. faecalis* isolates were kept for long-term storage in a cryoprotectant mixture of saline–phosphate buffer (NaCl, 8.5 g L^{-1}; K_2HPO_4, 0.65 g L^{-1}; KH_2PO_4, 0.35 g L^{-1}; pH 7.0; 700 µl), glycerol (100 µl) and dimethyl sulfoxide (100 µl), and stored at –70°C.

Ribotyping *E. faecalis* Isolates

Except for the specific restriction enzyme, bacterial isolates were processed on a RiboPrinter (DuPont Qualicon, Wilmington, DE) as described by

Kuntz *et al.* (2003). Briefly, each *E. faecalis* isolate was streaked onto a Petri plate containing BHI agar and was incubated at 37°C for 24 hours. Each RiboPrinter sample was obtained by touching the end of a sterile colony pick to a solid lawn of bacteria and mixing the sample with 40 µl of buffer. This step was repeated once before 30 µl of the buffer was transferred to a sample carrier. After the samples were inactivated by heat, 5 µl of two lysing agents were added to each sample before placing the sample carrier into the RiboPrinter system. The restriction enzyme P*st*I yielded banding patterns that discriminated the enterococci best (Hartel unpublished data), and therefore the DNA of each sample was digested with this enzyme. The DNA restriction fragments were size-separated by electrophoresis on a pre-cast agarose gel and transferred to a nylon membrane. The membrane was exposed to a series of proprietary chemical and enzymatic treatments to produce chemiluminescent DNA fragments. The banding pattern images created were captured by a CCD camera and stored as TXT files.

Statistical Analysis

Riboprint patterns were initially processed and analyzed with BioNumerics (Version 3.0, Applied Maths, Kortrijk, Belgium). The TXT files were imported and the tolerance (the maximum percentage shift allowed between two bands in different patterns for the bands to still be considered a match) was set at 1.00%. Intra- and inter-gel variations were assessed with the *E. faecalis* ATCC #19433 strain. The ribotype patterns were converted to binary data (where 0 represented the absence and 1 represented the presence of a band class) and were exported into SAS (Version 9.0, SAS Institute, Cary, NC) and S-PLUS (Version 6.2, Insightful Corp., Seattle, WA). The distribution of ribotype patterns was computed within each sample and for the host origin database as a whole, pooling all samples across locations and times. Distinct ribotype patterns were defined according to the Dice coefficient (Dice 1945) and a similarity index (percentage agreement between two patterns) of 100%. This index was chosen because intra- and inter-gel variations with the RiboPrinter and the *E. faecalis* ATCC #19433 strain were 100% (an exact match). Because a similarity index of 100% and a tolerance level of 1.00% determined by the BioNumerics software may be too stringent for defining distinct band classes (Myoda *et al.* 2003), the bands were also analyzed at a more relaxed similarity index of 90% or a more relaxed tolerance of 10.00%. If an isolate contained a ribotype pattern that was not found among any isolates in the composite host origin database, then the ribotype pattern was considered unshared; conversely, if an isolate contained a ribotype pattern that was found in the composite host origin database, then the ribotype pattern was considered shared. The number of ribotype patterns and the number of isolates sharing the same ribotype pattern were converted to a percentage to normalize unequal ribotype patterns or isolate numbers across different locations.

The correlation between the percentage of shared ribotype patterns and geographic distance was determined with a Mantel test, a nonparametric test of zero correlation between two distance matrices, where each entry represents a pairwise dissimilarity measure between two locations (Manly 1994). A product moment correlation was computed between the entries in the distance matrices. The entries in the second distance matrix were permuted and the product moment correlation was recomputed. After 10,000 repetitions, the process created a distribution by which it was possible to compute the probability of obtaining a stronger correlation than that observed in the experiment (assuming there was no initial correlation between the matrices). If the probability was ≤ 0.05 (statistically significant), then the test was rejected in favor of the existence of nonzero correlation. The first dissimilarity measure was the percentage of unshared ribotype patterns for the combined sampling of two locations (e.g., the number of ribotype patterns found in one location that were not represented by any ribotype patterns in the other location divided by the total number of distinct ribotype patterns across the two locations). The second dissimilarity measure was geographic distance between the locations in kilometers. The test was run using the "mantel.fcn" function in S-PLUS software.

To estimate the richness (or number) of ribotype patterns, an operational taxonomic unit (OTU) was defined as a ribotype pattern at a tolerance level of 1.00% and a similarity index of 100%, or separately, a tolerance level of 10.00% and a similarity index of 100%. The richness among OTU samples was assessed with a rarefaction curve, where the richness of the OTU community was assessed with two nonparametric estimators, Chao1 and abundance coverage (Hughes *et al.* 2001). To maximize the total number of isolates analyzed, 74 *E. faecalis* isolates from a wastewater treatment plant near Eulonia, GA

(Kuntz *et al.* 2003) were included in the analyses, for a grand total of 826 isolates. The richness estimates were computed using the freeware program EstimateS (Version 7.0; Colwell 1997).

Results

In 13 different Delaware wastewater treatment plants, fecal enterococcal counts ranged from 4.61 to 6.79 $\log_{10}$ CFU 100 ml^{-1} (Table 1). Enterococcal counts from Georgia, Idaho, New York, and Puerto Rico were similar, and ranged from 4.27 to 6.73 $\log_{10}$ CFU 100 ml^{-1}. The three counts of fecal enterococci from the same Delaware cesspool averaged 4.87 $\log_{10}$ CFU 100 ml^{-1}. Therefore, all cesspool and wastewater treatment plant samples contained at least 10,000 enterococci 100 ml^{-1}.

All samples contained *E. faecalis*. The highest percentages of *E. faecalis* were observed in the Delaware latrine (96%) and cesspool (an average of 90%), whereas percentages in the wastewater treatment plants from Delaware, Idaho, New York, and Puerto Rico ranged from 14 (Winterthur, DE) to 80% (Laurel, DE). *Enterococcus faecalis* was the most common Enterococcus species in half of wastewater treatment plant samples (15 of 30 samples) and constituted a majority of the total bacterial isolates (2,058 of 3,952 isolates; 52%). Overall, the total number of unknown bacteria was low (242 isolates; 6%).

When the ribotype patterns were converted to binary data and the similarity cutoff was 100% (i.e., an exact match), there were 77 distinct banding classes at a 1.00% tolerance level, and 19 distinct banding classes at a 10.00% tolerance level. At the 1.00% tolerance level, the 752 *E. faecalis* isolates yielded 652 distinct ribotype patterns (Table 2), for an isolate:ribotype pattern ratio of approximately 1.2:1. The majority of these ribotype patterns (429/652 = 66%) were unshared. At the 10.00% tolerance level, the 752 *E. faecalis* isolates yielded 460 distinct ribotype patterns, for an isolate:ribotype pattern ratio of approximately 1.6:1. Almost a majority of these ribotype patterns (222/460 = 48%) were unshared. When the tolerance level was kept at 1.00%, but the similarity index reduced from 100 to 90%, the percentage of the 652 ribotypes that were unshared dropped from 66 to 50%. Therefore, relaxing either the similarity index or the tolerance level decreased the number of distinct ribotype patterns and increased the ribotype pattern matching among the isolates, but half or almost half of the ribotype patterns were still unshared.

The percentage of ribotype patterns shared between two locations did not increase as the distance between locations decreased. The Mantel test of zero correlation between the pairwise sharing of ribotype patterns and distance among the 32 samples (one location at one time) was not statistically significant (one-sided $p = 0.91$ at a tolerance level of 1.00%, and one-sided $p = 0.83$ at a tolerance level of 10.00%). When the geography was limited to one state, Delaware, the 424 ribotype patterns from the 490 *E. faecalis* isolates also showed no significant correlation between the percentage of shared ribotype patterns and distance (<1 to 152 km; one-sided $p = 0.38$ at a tolerance level of 1.00%). Even pooling isolates from one location over time did not increase ribotype sharing. The 118 isolates from a single sewage plant (Twin Falls, ID) that was sampled on Days 0, 2, 5, 7 and 86 resulted in 103 distinct ribotypes, of which 61 (59%) were unshared. Similarly, when one cesspool (Dover, DE) was sampled three times, with each sampling 3 m apart and each sample obtained within minutes of the other, the 74 *E. faecalis* isolates yielded 63 ribotypes, of which 81% were unshared. Therefore, even when multiple sampling was restricted to one location geographically or temporally, the majority of the *E. faecalis* isolates did not share ribotype patterns. When the tolerance level was increased to 10.00% or the similarity index was reduced to 90%, a significant percentage (≥44%) of the ribotype patterns were still unshared.

To increase the sample size, 74 *E. faecalis* isolates obtained from a wastewater treatment plant near Eulonia, GA, were combined with the 752 *E. faecalis* isolates from this study, and a) the observed richness was determined with a rarefaction curve at 1.00 and 10.00% tolerances, and b) the estimated population richness was determined with Chao1 and abundance-based coverage estimators at 1.00 and 10.00% tolerances. The observed richness and estimated population richnesses were essentially linear over the range of the 32 accumulated samples at either 1.00 or 10.00% tolerances (Figure 2). Because of this linearity, the total population richness could not be reliably estimated without additional sampling.

Discussion

Enterococcus faecalis was the most common Enterococcus species isolated from a latrine, cesspool, and 15 of 30 wastewater treatment plant

Table 1. Location, source, total count of fecal enterococci, number of bacterial isolates, and number (percent) of *Enterococcus faecalis*, other enterococci, and unknown bacteria from a latrine, cesspool, and wastewater treatment plants located in Delaware, Georgia, Idaho, New York, and Puerto Rico (see Figure 1 for locations). Percentages may not add up to 100% because of rounding.

Location	Source	Total count of fecal enterococci	No. of isolates	*Ent. faecalis*	Other enterococci	Unknown bacteria
		Log_{10} CFU 100 ml^{-1}			number (percent)	
Delaware						
Bellevue	Latrine	NA[1]	180	172 (96)	7 (4)	1 (1)
Bridgeville	Wastewater	5.83	90	48 (53)	37 (41)	5 (6)
Delaware City	Wastewater	5.54	90	69 (77)	17 (19)	4 (4)
Dover #1a[2]	Cesspool	4.88	60	52 (87)	8 (13)	0 (0)
Dover #1b	Cesspool	4.83	60	53 (88)	6 (10)	1 (1)
Dover #1c	Cesspool	4.90	60	55 (92)	4 (7)	1 (1)
Harrington #1a[3]	Wastewater	5.89	116	64 (55)	32 (28)	20 (17)
Harrington #1b	Wastewater	5.32	70	34 (49)	26 (37)	10 (17)
Kent #1a[3]	Wastewater	6.79	180	66 (37)	81(45)	33 (18)
Kent #1b	Wastewater	5.29	67	14 (21)	48 (72)	5 (7)
Laurel	Wastewater	5.61	90	72 (80)	18 (20)	0 (0)
Lums Pond	Wastewater	4.61	90	55 (61)	33 (37)	2 (2)
Middletown	Wastewater	5.38	180	103 (57)	67 (37)	10 (6)
Port Penn #1a[3]	Wastewater	5.43	180	138 (77)	32 (18)	10 (6)
Port Penn #1b	Wastewater	5.88	90	38 (42)	52 (58)	0 (0)
Rehoboth	Wastewater	5.99	67	25 (37)	27 (40)	5 (7)
Seaford	Wastewater	5.62	90	49 (54)	35 (39)	6 (7)
South Coastal	Wastewater	5.60	66	31 (47)	33 (50)	2 (3)
Wilmington #1a[3]	Wastewater	5.27	180	79 (44)	101 (56)	0 (0)
Wilmington #1b	Wastewater	5.11	90	20 (22)	63 (70)	7 (8)
Winterthur	Wastewater	5.54	90	13 (14)	76 (84)	1 (1)
Georgia						
Athens #1	Wastewater	4.62	180	94 (52)	84 (47)	2 (1)
Athens #2	Wastewater	5.75	180	87 (48)	80 (44)	13 (7)
Athens #3	Wastewater	5.75	180	99 (55)	68 (38)	13 (7)
Idaho						
Twin Falls[4]	Wastewater	5.44	112	55 (49)	48 (43)	9 (8)
Twin Falls	Wastewater	5.47	94	27 (29)	57 (61)	10 (11)
Twin Falls	Wastewater	5.41	125	50 (40)	57 (46)	18 (14)
Twin Falls	Wastewater	5.43	113	39 (35)	44 (39)	30 (27)
Twin Falls	Wastewater	5.31	184	75 (41)	107 (58)	2 (1)
New York						
Tivoli	Wastewater	4.29	78	33 (42)	45 (58)	0 (0)
Wappingers Falls	Wastewater	4.27	83	40 (48)	43 (52)	3 (4)
Puerto Rico						
Cabo Rojo	Wastewater	6.67	180	96 (53)	80 (44)	4 (2)
Lajas	Wastewater	6.73	180	89 (49)	76 (42)	15 (8)
San German	Wastewater	5.32	77	24 (31)	53 (69)	0 (0)
Total			3952	2058 (52)	1645 (42)	242 (6)

[1], NA, sampled with sterile tongue depressor.
[2], same cesspool sampled at three different locations within minutes of each other.
[3], one plant sampled on two different days.
[4], one plant sampled at 0, 2, 5, 7, and 86 d.

Table 2. Number of *Enterococcus faecalis* isolates from a cesspool, latrine, and wastewater treatment plants in Delaware, Georgia, Idaho, New York, and Puerto Rico (see Figure 1 for locations); and the number of ribotype patterns obtained from those isolates. Each pattern was compared to the entire database, and the number of patterns (and percentage) *not* shared at 1.00 or 10.00% tolerance, and 90 or 100% similarity index is given. Tolerance is the maximum percentage shift allowed between two bands on different patterns for the bands to still be considered a match.

Location	Source	No. of Ent. *Faecalis* isolates	Tolerance = 1.00% Similarity index = 100%		Tolerance = 10.00% Similarity index = 100%		Tolerance = 1.00% Similarity index = 90%	
			No. of ribotype patterns	No. of patterns not shared (%)	No. of ribotype patterns	No. of patterns not shared (%)	No. of ribotype patterns	No. of patterns not shared (%)
Delaware								
Bellevue	Latrine	25	20	12 (60)	10	7 (70)	20	6 (30)
Bridgeville	Wastewater	25	23	15 (65)	16	7 (44)	23	11 (48)
Delaware City	Wastewater	25	19	9 (47)	14	4 (29)	19	8 (42)
Dover #1a[1]	Cesspool	25	18	13 (72)	16	7 (44)	18	5 (28)
Dover #1b	Cesspool	24	23	21 (91)	21	14 (67)	23	13 (57)
Dover #1c	Cesspool	25	22	17 (77)	19	9 (47)	22	10 (45)
Harrington #1a[2]	Wastewater	23	20	14 (70)	17	9 (53)	20	12 (60)
Harrington #1b	Wastewater	25	24	20 (83)	22	14 (64)	24	20 (83)
Kent #1a[2]	Wastewater	23	12	12 (100)	12	8 (67)	12	10 (83)
Kent #1b	Wastewater	12	18	11 (61)	15	9 (60)	18	10 (56)
Laurel	Wastewater	24	14	9 (64)	8	4 (50)	14	5 (36)
Lums Pond	Wastewater	24	23	17 (74)	20	9 (45)	23	13 (57)
Middletown	Wastewater	25	20	14 (70)	8	2 (25)	20	6 (30)
Port Penn #1a[2]	Wastewater	25	18	7 (39)	8	4 (50)	18	5 (28)
Port Penn #1b	Wastewater	25	21	10 (48)	13	6 (46)	21	6 (29)
Rehoboth	Wastewater	25	25	18 (72)	18	12 (67)	25	15 (60)
Seaford	Wastewater	25	23	16 (70)	13	6 (46)	23	14 (61)
South Coastal	Wastewater	26	26	23 (88)	18	10 (56)	26	20 (77)
Wilmington #1a[2]	Wastewater	26	25	14 (56)	16	5 (31)	25	11 (44)
Wilmington #1b	Wastewater	20	18	12 (67)	14	5 (36)	18	11 (61)
Winterthur	Wastewater	13	12	5 (42)	12	4 (33)	12	5 (42)
Georgia								
Athens #1	Wastewater	24	19	6 (32)	10	1 (10)	19	4 (21)
Athens #2	Wastewater	24	23	14 (61)	16	8 (50)	23	12 (52)
Idaho								
Twin Falls[3]	Wastewater	25	23	14 (61)	16	7 (44)	23	11 (48)
Twin Falls	Wastewater	18	13	7 (54)	8	2 (25)	13	6 (46)
Twin Falls	Wastewater	25	25	15 (60)	15	7 (47)	25	13 (52)
Twin Falls	Wastewater	25	21	14 (67)	15	6 (40)	21	12 (57)
Twin Falls	Wastewater	25	21	11 (52)	13	9 (69)	21	7 (33)
New York								
Tivoli	Wastewater	24	21	16 (76)	17	11 (65)	21	12 (57)
Wappingers Falls	Wastewater	24	22	18 (82)	18	10 (56)	22	14 (64)
Puerto Rico								
Cabo Rojo	Wastewater	24	19	10 (53)	9	2 (22)	19	6 (32)
San German	Wastewater	24	21	15 (71)	13	4 (31)	21	14 (67)
Total database		752	652	429 (66)	460	222 (48)	652	327 (50)

[1], same cesspool sampled at three different locations within minutes of each other.

[2], one plant sampled on two different days.

[3], one plant sampled at 0, 2, 5, 7, and 86 d.

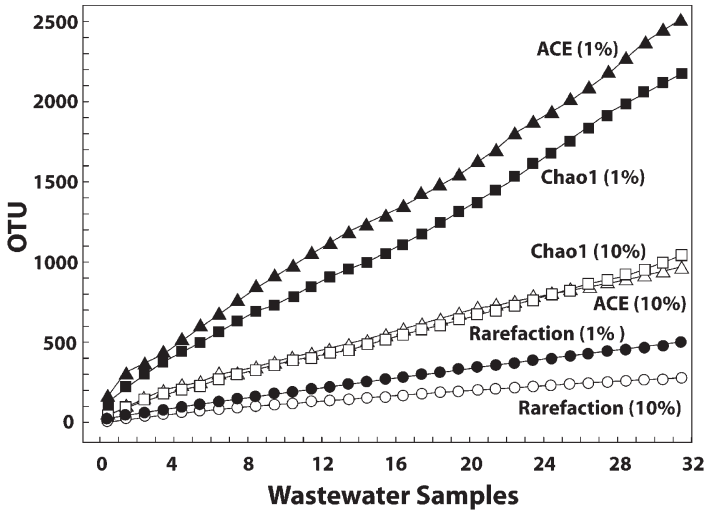

Figure 2. Sample richness (squares) using rarefaction analysis, and community richness using Chao1 (circles) and abundance-based coverage estimators (ACE; triangles) for 826 *Ent. faecalis* isolates. An OTU was defined as a ribotype pattern matching at a tolerance level of either 1.00 or 10.00%, and a similarity index of 100%. Each of the 32 samples represents one location at one time.

samples from Delaware, Georgia, Idaho, New York, and Puerto Rico. The presence of this bacterium was expected because *E. faecalis* is consistently found in the feces of birds and humans (Pourcher *et al.* 1991, Wheeler *et al.* 2002) and in urban sewage (Manero *et al.* 2002). Although the presence of bird feces in the cesspool and wastewater treatment plants is possible, it is likely that virtually all the *E. faecalis* isolates were from human feces because of the vast discrepancy in loading rates between these two sources. The predominance of *E. faecalis* in the cesspool, latrine, and wastewater treatment plant water also supports the suggestion that unless a high loading rate of bird feces is suspected, environmental water samples with a high percentage of *E. faecalis* are likely contaminated with human feces (Kuntz *et al.* 2004).

Because the host range of *E. faecalis* is essentially limited to humans and birds (Pourcher *et al.* 1991, Wheeler *et al.* 2002, Kuntz *et al.* 2004), it was hoped that fewer ribotype patterns would be needed in a permanent host origin database for good matching. This reduction did not occur. When ribotype patterns and isolates were pooled across all locations, most ribotype patterns were represented by a single isolate. Even when the matching criterion was relaxed by decreasing the tolerance level from 1.00 to 10.00% or decreasing the similarity index from 100 to 90%, the percentage of unshared ribotypes remained high.

The isolate:ribotype pattern ratio of 1.6:1 for the 752 *E. faecalis* isolates (tolerance = 1.00%; similarity index = 90%) was similar to the isolate:ribotype pattern ratio of 1.8:1 for *E. coli* isolated from yearling steers (Jenkins *et al.* 2003) and 1.7:1 isolate:ribotype pattern ratio for *E. coli* isolated from human and nonhuman sources (Parveen *et al.* 1999). However, these ratios should be used with caution because they vary with the tolerance level and the similarity index cutoff. For example, keeping the tolerance level of 1.00% but increasing the similarity index to 100% changed the *E. faecalis* isolate:ribotype pattern ratio to 1.2:1. The similarity index cutoff was 90% for Jenkins *et al.* (2003) and 75% for Parveen *et al.* (1999); if their similarity index cutoff had been more stringent, then their ratios likely would have been closer to 1:1 as well. However, relaxing the similarity index from 100 to 90% or decreasing the tolerance level from 1.00 to 10.00% (similar to what could be achieved with manual ribotyping) makes little sense because it would negate the reproducibility of the RiboPrinter, a reproducibility that was only achieved at a high cost.

There are two possible explanations for the results: 1) the *E. faecalis* isolates contained different ribotype patterns that changed with geography, or 2) the *E. faecalis* isolates contained different ribotype patterns that did not change with geography, and only the small sampling size gave the appearance

of change. These explanations are similar to those given by Hartel *et al.* (2004) for the geographic changes of *E. coli* isolates in the Chattahoochee River of Georgia.

In the first explanation, no significant correlation was observed between the percentage of shared ribotype patterns and distance between individual samples obtained from Delaware, Georgia, Idaho, New York, and Puerto Rico wastewater treatment plants ($p = 0.91$), or when the *E. faecalis* isolates were limited to Delaware ($p = 0.38$; <1 to 152 km between locations). Even when *E. faecalis* isolates were obtained from three sites only meters apart in a cesspool, the majority of ribotype patterns obtained from these isolates were unshared. These results do not agree with those of Wiggins *et al.* (2003), who suggested that matching of antibiotic resistance patterns for fecal enterococci increased with decreased distance. There are several reasons that may account for this difference. First, ribotyping is much more discriminatory (and thus generates more patterns) than antibiotic resistance analysis (Farber 1996). Second, this lack of matching may be animal-host specific. Hartel *et al.* (2002) observed that the percentage of shared *E. coli* ribotype patterns within an animal species increased with decreased distance for cattle and horses, but not for swine and chicken. It may be that there is less sharing of ribotype patterns from *E. faecalis* isolates from humans because the human diet is more variable or because a significant number of humans are traveling. Why chickens and swine are similar to humans and dissimilar to horses and cattle is unclear, but may also be because of diet. Hartel *et al.* (2003) showed that ribotype pattern diversity of *E. coli* isolates in deer was affected by diet.

The second explanation, that none of the ribotype patterns changed geographically and only the small sample size gave the appearance of these changes, is supported by the observed richness, determined with a rarefaction curve, and total population richness curves, estimated with Chao1 and accumulation-based coverage estimators, none of which leveled off. Even relaxing the stringency by increasing the tolerance level from 1.00 to 10.00% (and thus reducing the number of band classes) did not affect the results. Therefore, the clonal diversity must be high, and these estimates represented, at best, only the low bounds of richness, and suggest that much more sampling is needed to obtain a more precise estimate of total richness.

Whether the *E. faecalis* isolates contained different ribotype patterns that changed with geography, or that only the small sampling size gave the appearance of these changes, the data are discouraging, because, just like the data for *E. coli* (Johnson *et al.* 2004) and fecal enterococci (Hagedorn *et al.* 1999, Wiggins *et al.* 2003), they suggest that the host origin database must still contain many thousands of *E. faecalis* isolates in order for a reasonable amount of pattern matching to occur, even when the host range of the indicator bacterium is essentially limited to humans. Because such a permanent host origin database is time-consuming and expensive to construct, the data suggest that a continued emphasis on targeted sampling as a prelude to BST (Kuntz *et al.* 2003), coupled with various genotypic (e.g., Ram *et al.* 2004), phenotypic (e.g., Wheeler *et al.* 2002), or chemical (e.g., optical brighteners; Hagedorn *et al.* 2003) methods that are host origin database independent, may be more appropriate for identifying and source tracking of human fecal contamination in waterways.

The results obtained here suggested that either ribotype patterns of *E. faecalis* change with geography, or that the small sampling size gave the appearance of ribotype patterns changing with geography. Given the cost and time to construct a permanent host origin database where good matching might be expected, designing and testing alternative methods for BST seem preferable.

Literature Cited

Baele, M., L.A. Devriese, P. Butaye and F. Haesebrouck. 2002. Composition of enterococcal and streptococcal flora from pigeon intestines. *Journal of Applied. Microbiology* 92:348-351.

Buchan, A., M. Alber and R.E. Hodson. 2001. Strain–specific differentiation of environmental *Escherichia coli* isolates via denaturing gradient gel electrophoresis (DGGE) analysis of the 16S–23S intergenic spacer region. *FEMS Microbiology Ecology* 35:313-321.

Clesceri, L.S., A.E. Greenberg and A.D. Eaton. 1998. Standard Methods for the Examination of Water And Wastewater, 20th ed. American Public Health Association., American Water Works Association, and Water Environmental Federal Regulations. Washington, DC.

Colwell, R.K. 1997. EstimateS: Statistical estimation of species richness and shared species from

samples. Version 5. User's guide and application published at http://viceroy.eeb.uconn.edu/estimates

Dice, L.R. 1945. Measures of the amount of ecologic association between species. *Ecology* 26:297-302.

Dombek, P.E., L.-A.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. *Applied Environmental Microbiology* 66:2572-2577.

Farber, J.M. 1996. An introduction to the hows and whys of molecular typing. *Journal of Food Protection* 59:1091-1101.

Gordon, D.M., S. Bauer and J.R. Johnson. 2002. The genetic structure of *Escherichia coli* populations in primary and secondary habitats. *Microbiology* 148:1513-1522.

Hagedorn, C., R.B. Reneau, M. Saluta and A. Chapman. 2003. Impact of onsite wastewater systems on water quality in coastal regions. Virginia Coastal Resources Management Program Memorandum of Agreement 50312–01–13–PT. Virginia Department of Conservation and Recreation, Virginia Department of Health, 1/28/02 - 6/30/03.

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. *Applied Environmental Microbiology* 65:5522-5531.

Hartel, P.G., J.D. Summer and W.I. Segars. 2003. Deer diet affects ribotype diversity of *Escherichia coli* for bacterial source tracking. *Water Research* 37:3263-3268.

Hartel, P.G., E.A. Frick, A.L. Funk, J.L. Hill, J.D. Summer and M.B. Gregory. 2004. Sharing of ribotype patterns of *Escherichia coli* isolates during baseflow and stormflow conditions. US Geological Survey Scientific Investigations Report 2004–5004. Available at http://water.usgs.gov/pubs/sir/2004/5004 (posted 22 Apr 2004).

Hartel, P.G., W.I. Segars, J.D. Summer, J.V. Collins, A.T. Phillips and E. Whittle. 2000. Survival of fecal coliforms in fresh and stacked broiler litters. *Journal of Applied Poultry Research* 9:505-512.

Hartel, P.G., J.D. Summer, J.L. Hill, J.V. Collins, J.A. Entry and W.I. Segars. 2002. Geographic variability of *Escherichia coli* ribotypes from animals in Idaho and Georgia. *Journal of Environment Quality* 31:1273-1278.

Hughes, J.B., J.J. Hellmann, T.H. Ricketts and B.J.M. Bohannan. 2001. Counting the uncountable: Statistical approaches to estimating microbial diversity. *Applied Environment Microbiology* 67:4399-4406.

Jenkins, M.B., P.G. Hartel, T.J. Olexa and J.A. Stuedemann. 2003. Putative temporal variability of *Escherichia coli* ribotypes from yearling steers. *Journal of Environment Quality* 32:305-309.

Johnson, L.–A.K., M.B. Brown, E.A. Carruthers, J.A. Ferguson, P.E. Dombek and M.J. Sadowsky. 2004. Sample size, library composition, and genotypic diversity among natural populations of *Escherichia coli* from different animals influence accuracy of determining sources of fecal contamination. *Applied Environment Microbiology* 70:4478-4485.

Kuntz, R.L., P.G. Hartel, K. Rodgers and W.I. Segars. 2004. Presence of *Enterococcus faecalis* in broiler litter and wild bird feces for bacterial source tracking. *Water Research* 38:3351-3557.

Kuntz, R.L., P.G. Hartel, D.G. Godfrey, J.L. McDonald, K.W. Gates and W.I. Segars. 2003. Targeted sampling protocol as prelude to bacterial source tracking with *Enterococcus faecalis*. *Journal of Environment Quality* 32:2311-2318.

Manero, A. and A.R. Blanch. 1999. Identification of Enterococcus spp. with a biochemical key. *Applied Environment Microbiology* 65:4425-4430.

Manero, A., X. Vilanova, M. Cerdà–Cuéllar and A.R. Blanch. 2002. Characterization of sewage waters by biochemical fingerprinting of Enterococci. *Water Research* 36:2831-2835.

Manly, B.F.J. 1994. Multivariate Statistical Methods: A Primer, 2nd ed. Chapman and Hall. New York, NY.

Myoda, S.P., C.A. Carson, J.J. Fuhrman, B.-K. Hahm, P.G. Hartel, H. Yampara–Iquise, L.-A. Johnson, R.L. Kuntz, C.H. Nakatsu, M.J. Sadowsky and M. Samadpour. 2003. Comparison of genotypic-based microbial source tracking methods requiring a host origin database. *Journal of Water Health* 1:167-180.

Parveen, S., K.M. Portier, K. Robinson, L. Edmiston and M.L. Tamplin. 1999. Discriminant analysis of ribotype profiles of *Escherichia coli* for differentiating human and nonhuman sources of fecal pollution. *Applied Environment Microbiology* 65:3142-3147.

Pourcher, A.M., L.A. Devriese, J.F. Hernandez and J.M. Delattre. 1991. Enumeration by a miniaturized method of *Escherichia coli*, Streptococcus bovis, and enterococci as indicators of the origin of fecal pollution of waters. *Journal of Applied Bacteriology* 70:525-530.

Ram, J.L., R.P. Ritchie, J. Fang, F.S. Gonzales and J.P. Selegean. 2004. Sequence-based source tracking of *Escherichia coli* based on genetic diversity of b-glucuronidase. *Journal of Environment Quality* 33:1024-1032.

Wheeler, A.L., P.G. Hartel, D.G. Godfrey, J.L. Hill and W.I. Segars. 2002. Combining ribotyping and limited host range of *Enterococcus faecalis* for *microbial source tracking. Journal of Environment Quality* 31:1286-1293.

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. *Applied Environment Microbiology* 69:3399-3405.

Acknowledgments

The authors thank Dan Hall and Tyler Laporte for their technical assistance. This research was partially supported by the Delaware Department of Natural Resources and Environmental Control, Division of Water Resources, New York Multistate Hatch S-297, Puerto Rico Water Resources and Environmental Research Institute, and the State of Georgia Environmental Protection Division through the US Environmental Protection Agency.