

Use of composite data sets for source-tracking enterococci in the water column and shoreline interstitial waters on Pensacola Beach, Florida

Fred J. Genthner ^{*}, Joseph B. James, Diane F. Yates, Stephanie D. Friedman

US Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Gulf Ecology Division, Gulf Breeze, FL 32561, USA

Abstract

Sources of *Enterococcus faecalis* isolates from Pensacola Beach, FL. were identified using a library-based approach by applying the statistical method of average similarity to single and composite data sets generated from separate analyses. Data sets included antibiotic resistance analysis (ARA), rep-fingerprints, and fatty acid methyl ester (FAME) profiles. Use of a composite data set composed of ARA and rep-fingerprints, added to the confidence of the identifications. The addition of FAME data to composite data sets did not add to the confidence of identifications. Source identification was performed to better understand risk associated with higher densities of enterococci found in swash zone interstitial water (SZIW) as compared to adjacent bathing water on Pensacola Beach, FL. The “swash zone” is that area of the beach continually washed over by waves. As the potential sources of enterococci were limited in this environment, only two library units, sea gull and human, were constructed. Identification of the beach isolates using a composite data set indicated a sea gull origin. The clonality of the beach isolates suggested that the beach environment selects certain subspecies of *E. faecalis*.

Published by Elsevier Ltd.

Keywords: *Enterococcus faecalis*; Bacterial source-tracking; Swash zone; Fecal indicator bacteria; Florida

1. Introduction

The swash zone is that area of the beach where waves continuously wash up on the sand. Controversy exists regarding whether this area should be an integral part of epidemiological and microbiological studies correlating recreational water quality with health effects. Swash zone interstitial water (SZIW) generally harbors higher densities of microorganisms and fecal indicator bacteria than adjacent waters; various microbial pathogens may be encountered though direct contact (for a review see [World Health Organization, 2003](#)). Considerable expo-

sure is a likely occurrence as a significant percentage of time is spent by bathers, particularly children, on the beach itself rather than in the water. Despite these problem-suggestive data, epidemiological evidence for such risks has not been found ([Chabasse et al., 1986](#)). Regardless, the swash zone has not been sufficiently investigated to rule out potential risks. Research is needed to determine if special consideration of interstitial water and its impact on sensitive populations is warranted ([US EPA, 1999](#)).

To alert health officials to potential public health and resource risks, enteric bacteria are used as fecal pollution indicators. Bacteria source-tracking (BST) refers to a variety of techniques used to identify the warm-blooded host from which enteric bacteria were derived. Widely used are the “library-dependent” methods which

^{*} Corresponding author. Tel.: +1 850 934 9342; fax: +1 850 934 9201.
E-mail address: genthner.fred@epa.gov (F.J. Genthner).

involve the collection of phenotypic or genotypic data from enteric bacteria isolated from warm-blooded animals known to inhabit the study area. Data from each animal source is assembled into separate library units. Correlations between data from environmental isolates and data stored in library units are made using statistical methods such as discriminant analysis and maximum or average similarity (see Ritter et al., 2003 for a recent review of these methods). These correlations, used to indicate a probable source, form the basis for all library-dependent BST methods.

Few studies have addressed source of fecal contamination in beach sand or SZIW. Animals were assumed to be the source of fecal streptococci isolated from surface sand on French beaches (World Health Organization, 2003). Using rep-PCR DNA fingerprinting Whitman et al. (2001) reported sea gulls and other warm-blooded animals as the source of enterococci from sand taken from beaches along the southern shores of Lake Michigan. In a recent study Wheeler Alm et al. (2003) reported high densities of fecal indicator bacteria in swash zone sand; however, the source of these bacteria was not addressed.

In this paper we report higher densities of enterococci in swash zone interstitial water (SZIW) than from the adjacent bathing water on six Gulf of Mexico and sound-side beaches on Pensacola Beach, FL. Potential sources of these bacteria were then addressed. Library-dependent approaches traditionally use single data sets containing results from a single genotypic or phenotypic method. In our study composite data sets, synthetic experiments that combined the information from separate experiments, were tested to determine whether they could improve the confidence of the identifications. Because the diversity of warm-blooded animals on these beaches was limited, a library possessing only two units, sea gull and human, was considered adequate. Dogs are not allowed on the beaches and small rodents are not known to frequent the largely unprotected sandy shores.

Selection of the genotypic method, rep-PCR, a DNA fingerprinting technique capable of resolving to the level of subspecies (Versalovic et al., 1994), necessitated use of a single *Enterococcus* species. *Enterococcus faecalis*, the predominant *Enterococcus* in birds and humans (Wheeler et al., 2002), was selected for our study. One of the phenotypic methods, antibiotic resistance analysis (ARA), has been widely used to successfully source-track both *E. coli* and enterococci in numerous watersheds (see Scott et al., 2002 for a recent review). The other phenotypic method, fatty acid methyl esters (FAME) of membrane lipids, is commonly used as a bacterial identification method (Microbial ID, Inc. Newark, Delaware). These methods all share the common underlying assumption that differences within the host's gastrointestinal tract will select for subspecies or strains of intestinal bacteria. Thus, a major question

addressed in this study is whether phenotypic information will complement and enhance genotypic information resulting in a more confident source identification.

2. Materials and methods

2.1. Sampling and analysis

Water samples were collected in triplicate using sterile 1 l bottles. Water column, "bathing water", samples were collected at a depth of 0.5 m at a point off shore where the depth of the water was 1.0 m. Swash zone interstitial water samples were obtained from holes dug approximately 70 cm deep into the sand 1–1.5 m from the shoreline and located directly inshore from where the water column samples were taken. Within 10 min a sufficient volume of interstitial water had seeped into the hole to allow sampling. Duplicate 20 ml sub-samples were removed for dissolved organic carbon (DOC) determination, filtered through GFF filters (25 mm) into small glass vials, and stored frozen (–70 °C) until analysis. Both filters and vials were pre-combusted in an oven at 500 °C for 4 h. DOC samples were analyzed with a Shimadzu TOC 5050 carbon analyzer standardized with potassium hydrogen phthalate.

2.2. Bacterial isolation, identification and enumeration

Enterococci were enumerated in water samples using the US Environmental Protection Agency's Method 1600 (US EPA, 2000). Data, expressed as the number of enterococci per 100 ml, were reported from a single sample and as the mean of 5 samples equally spaced over a 30-day period (months June and August). For library assembly enterococci were collected by rectal swab from 10 human volunteers and by cloacal swab from 10 laughing gulls (Wildlife Sanctuary of Northwest Florida, Pensacola, FL). Persons taking antibiotics at the time of sampling were excluded. Upon collection, swabs (BBL® CultureSwab®, Collection and Transport System, Becton Dickinson) saturated in transport buffer were held on ice until processed at the laboratory. Within 6 h of collection swabs were streaked onto the surface of a 0.45 µm filter (Pall Corp., Ann Arbor, MI) which lay on top of mEI (Messer and Dufour, 1998) agar medium contained in a 47 mm Petri dish. In both the enumeration procedure and library assembly isolations, all colonies on mEI agar, regardless of color, with a blue halo on the membrane were further confirmed as enterococci by growth in brain heart infusion (BHI) + NaCl (6.5%) broth and lack of catalase activity. From among these putative enterococci, *E. faecalis* were identified by PCR amplification of a species-specific internal fragment of a gene related to glycopeptide-resistance enzymes (Dutka-Malen et al., 1995). *E. faecalis* isolates,

routinely cultured in trypticase soy broth (TSB), were stored frozen (-70°C) in 20% glycerol.

Cell lysis and DNA extraction were conducted using a Mini-BeadBeater-8 (BioSpec Products, Inc., Bartlesville, OK) homogenizer and the UltraClean Microbial DNA Isolation Kit (MoBio Inc., Solana Beach, Calif). Lysis was performed according to manufacturers instructions with the addition of an enzymatic digestion (37°C for 18 h in lysozyme, 1.1 mg/ml and mutanolysin, 20 $\mu\text{g}/\text{ml}$) and three freeze–thaw cycles ($-70^{\circ}\text{C}/60^{\circ}\text{C}$) to aid in lysis.

2.3. Construction of data sets

The rep-PCR procedure was performed essentially as described by Dombek et al. (2000) with modifications as follows. Forward primer (REP2 5'NCGNCTTATCNGGCCTAC 3') and a reverse primer (REP1R 5' NCGNCGNCATCNGGC 3') were selected; MgCl_2 was adjusted to 8.5 mM per 25 μl reaction mixture. PCR thermal cycling regime was performed (5 min at 94°C , 4 min at 80°C , 5 min at 94°C ; 30 cycles with 30 s at 90°C , 30 s at 40°C , 8 min at 65°C , followed by a final extension of 16 min at 65°C) using a Gene Amp PCR System 9700 (PE Applied Biosystems). *E. faecalis* ATCC 19433 and *E. faecium* ATCC 35667 were included in each rep-PCR batch as internal controls. Gels were stained in ethidium bromide solution and the image captured using the Kodak Digital Science Image Station 440 (New Haven, CT). FAME analysis was performed according to MIDI, Inc. specifications (Paisley, 1996). Antibiotic resistance analysis (ARA) was performed as described by Harwood et al. (2000) using trypticase soy agar plates amended with the following concentrations of antibiotics: amoxicillin, 0.5, 2, and 20 $\mu\text{g}/\text{ml}$; bacitracin, 10, 25 and 100 $\mu\text{g}/\text{ml}$; cephalothin, 2, 10, and 40 $\mu\text{g}/\text{ml}$; erythromycin, 0.05, 0.5 and 10 $\mu\text{g}/\text{ml}$; nitrofurazone, 8, 24, and 80 $\mu\text{g}/\text{ml}$; penicillin G, 0.5, 3, 20 $\mu\text{g}/\text{ml}$; streptomycin 1, 20 and 100 $\mu\text{g}/\text{ml}$; tetracycline, 0.4, 20, and 100 $\mu\text{g}/\text{ml}$; trimethoprim, 0.5, 5, and 10 $\mu\text{g}/\text{ml}$; trimethoprim/sulfamethoxazole 1:19, 5, 50 and 200 $\mu\text{g}/\text{ml}$; tylosin, 0.05, 5, and 100 $\mu\text{g}/\text{ml}$; vancomycin, 0.4, 4, and 8 $\mu\text{g}/\text{ml}$.

3. Statistical analysis

Differences in concentrations of dissolved organic carbon and densities of enterococci between SZIW and bathing water were compared for statistical differences ($p < 0.05$) using the two-tailed *T*-test.

All data processing, gel processing and statistical analyses pertaining to the *E. faecalis* isolates were performed using BioNumerics software v.3 (Applied Maths, Sint-Martens-Latem, Belgium). Similarity coeffi-

cients were calculated for each of the three data sets (REP-PCR, ARA and FAME) and for a composite data set (REP-PCR + ARA). For the rep-PCR data set, gel images were normalized and processed using the Pearson product-moment correlation, a curve-based method. This method, involving a calculation of a densitometric curve along the pattern, is insensitive to relative differences in intensity and background fluctuations between different gels. However, this method is sensitive to local differences in background fluctuations within a gel. ARA results, scored as growth or no growth, were treated as binary data. A closed character type (1 resistant/growth or 0 sensitive/no growth) was used.

Simple matching, which is the proportion of responses that match regardless of whether the attribute is present (resistant) or absent (sensitive), was used to determine similarity coefficients. FAME data was expressed as the percentage of each fatty acid methyl ester from all fatty acid methyl esters detected in a particular isolate. Data input was handled as an open character type meaning that a non-fixed number of characters (FAME's) were allowed. Pearson correlation was used to determine similarity coefficients.

For the composite data set (REP-PCR plus ARA) the “take from experiments” method was used to compute similarity coefficients. With this method the matrices from the individual experiments were averaged according to defined weights. Equal weights were placed on both experimental methods as the ARA data set contributed 36 characters (12 antibiotics at three concentrations), and the REP-PCR contributed approximately the same number of characters which, in this case, were bands on the gel.

The similarity coefficients were used to perform additional statistical analyses. Dendograms were constructed using UPGMA (unweighted pair-group method using arithmetic averages). In addition, 12 *E. faecalis* beach isolates from each data set were assigned to either the human or sea gull source group using average similarity (AS). With AS rather than predicting source membership based on proximity to a single known isolate, unknown isolates were classified into the source category yielding the largest average similarity to source-grouped isolates. In addition to source identification, the BioNumerics software offers information regarding the confidence of each assignment. This information is provided as an identification score followed by a quality factor, letters A–E. The identification score is a relative number (0–100); higher scores denoting more confident identifications. The identification scores are the similarity values obtained using the coefficient specified in the settings of the experimental type. The program calculates the similarity of the unknown entry with each of the library unit's entries and uses the average of these values. Regarding quality factors, grade “A” is considered a very confident identification; grade “B” is a good

identification; grade “C” is still considered a faithful identification whereas grade “D” is considered a doubtful identification.

4. Results

4.1. Field data

Fig. 1 shows the location of the sites where the paired samples (bathing and SZIW) were taken. The Gulf of Mexico site (N-30°19.647', W-087°09.598') was characterized by its proximity to two hotels, numerous bathers, high energy wave action, and sea gulls. The Santa Rosa Sound site (N-30°20.094', W-087°08.425') is a lower energy wave beach attracting many bathers, particularly young children. Numerous, small water craft frequent this area. Sea gulls were also commonly observed.

With a single exception (Gulf of Mexico, August) significantly higher concentrations of dissolved organic carbon (DOC) were found in the SZIW than in the adjacent bathing water (Table 1). The SZIW also harbored higher numbers of enterococci than did the adjacent bathing water (Fig. 2). These numbers were significantly higher at the Gulf of Mexico site. An average of 140 and 40 enterococci per 100 ml were found in the SZIW at the Gulf of Mexico site in June and August, respectively. Three out of five SZIW samples from June and one out of five samples from August yielded counts greater than 104 enterococci per 100 ml. In contrast, no single bathing water sample yielded a count higher than 59 enterococci per 100 ml, nor was the average number of enterococci over a one month sampling period greater than 34 enterococci per 100 ml.

Data concerning the enterococci isolated in this study are found in Table 2. Approximately 10 enterococci

were isolated from each of 10 sea gulls and 11 people. Nearly all human and sea gull isolates were identified as *E. faecalis*. In contrast, only 32% of the 98 enterococci isolated from the beach were identified as *E. faecalis*. The resolving power of Rep-PCR allowed the recognition of different sub-species through visual examination of the DNA fingerprints. Thus, by using this method we found that only a single sub-species of *E. faecalis* could be isolated from each person. *E. faecalis* colonizing the gut of sea gulls appeared less clonal as multiple sub-species were found within an individual sea gull. An inspection of the REP-PCR fingerprints derived from the *E. faecalis* isolates obtained from the beach also suggested a fairly high degree of relatedness.

4.2. Dendograms

To exclude clonal isolates from each source only those isolates which appeared unique by visual inspection of their REP-PCR fingerprints were further analyzed. Removal of clonal *E. faecalis* isolates from each source reduced the numbers to 20 from sea gulls, 11 from humans, and 12 from the beach. Four different dendograms showing similarities between these isolates were produced using genotypic and phenotypic data (Fig. 3). Fig. 3A shows the dendogram constructed from REP-PCR genomic fingerprints. The UPGMA clustering method showed three distinct phenons representing the human, sea gull and beach isolates with all isolates showing >25% similarity. The beach and human isolates formed shallow phenons; all isolates contained within each phenon show a >90% similarity. The remaining phenon, composed of sea gull isolates, was quite deep possessing a >51% similarity. The beach phenon grouped closer to the sea gull phenon (within 39%) than it grouped to the human phenon (within 64%).

The dendogram constructed from ARA data is shown in Fig. 3B. As with the REP-PCR dendogram three major phenons were formed. A phenon, containing all beach isolates, showed that several of the isolates possessed identical antibiotic resistance patterns. Unlike the REP-PCR dendogram, however, the other two major phenons contained a mixture of sea gull and human isolates. Also, in contrast to the REP-PCR dendogram, the ARA dendogram showed higher similarities between isolates with all isolates showing a >77% similarity.

Fig. 3C shows the dendogram constructed from the FAME data. This data set, in which the input was handled as an open character type, attained a maximum of 11 different FAME results. All *E. faecalis* isolates showed an extremely high degree of similarity to each other, >94%. An *E. faecium* strain, included for comparative purposes, showed the greatest dissimilarity. Clearly separated from the single, large *E. faecalis* cluster, it was almost 15% dissimilar to the closest *E. faecalis* isolate. This FAME data set was not used in any further

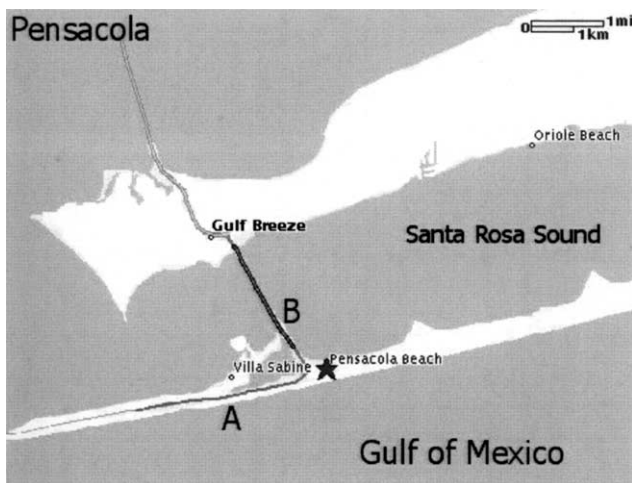


Fig. 1. Pensacola Beach (A) Gulf of Mexico site and (B) Santa Rosa Sound site.

Table 1
Dissolved organic carbon in bathing and swash zone interstitial waters (SZIW) on Pensacola Beach

Site	Dissolved organic carbon (mg/l)	
	June	August
Gulf of Mexico (GOM) Bathing water ^a /SZIW	2.0 ± 0.7 ^b /7.7* ± 0.58	2.2 ± 0.74/3.9 ± 1.6
Santa Rosa Sound Bathing water/SZIW	2.9 ± 0.47/5.0* ± 0.74	2.7 ± 0.30/4.3* ± 0.80

* Denotes a significant difference ($p < 0.05$) between means of DOC concentrations in bathing water and SZIW samples.

^a Bathing water sample taken at a depth of 0.5 m in 1 m deep water.

^b Standard deviation.

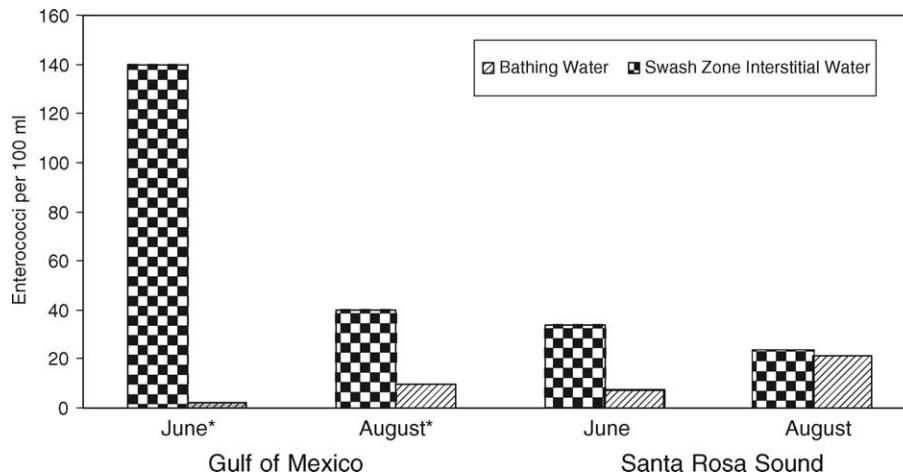


Fig. 2. Densities of enterococci in swash zone interstitial water and adjacent bathing waters on Pensacola Beach, Florida. Values shown represent the mean of five determinations taken over a 30-day period. *Denotes statistical significance between densities in the swash zone interstitial water/bathing water sampling pair.

Table 2
Enterococci isolated in this study

Source	No. of individuals sampled	No. of <i>Enterococcus</i> isolated	Percent identified as <i>Enterococcus faecalis</i>	Observations
Sea gull	10	178	98	Multiple sub-species of <i>E. faecalis</i> found within single birds
Human	11	96	100	Only one sub-species of <i>E. faecalis</i> found within a single human subject
Beach ^a	NA ^b	98	32	Clonal, difficult to distinguish individual subspecies of <i>E. faecalis</i>

^a Isolated from either bathing and swash zone interstitial water at both sites.

^b NA, not applicable.

analyses as all *E. faecalis* isolates were too tightly clustered to detect correlations between source groups.

The dendrogram derived from the composite data set of REP-PCR + ARA and constructed using similarity coefficients calculated with the “take from experiments” method is shown in Fig. 3D. All isolates showed >55% similarity. Beach and humans isolates both formed shallow clusters with >91 and >88% similarities, respectively. As in the REP-PCR dendrogram (Fig. 3A) three major phenons, each comprised entirely of either human, sea gull or beach isolates, are displayed. Also,

as with the REP-PCR dendrogram, the sea gull phenon grouped closer to the beach phenon (within 25%) than it did to the human phenon (within 34%). Calculating congruence between experiments showed all similarity values fell within one standard deviation.

4.3. Source determinations

The beach isolates, treated as unknown entries, were submitted to the library (composed of sea gull and human library units) for identification using average

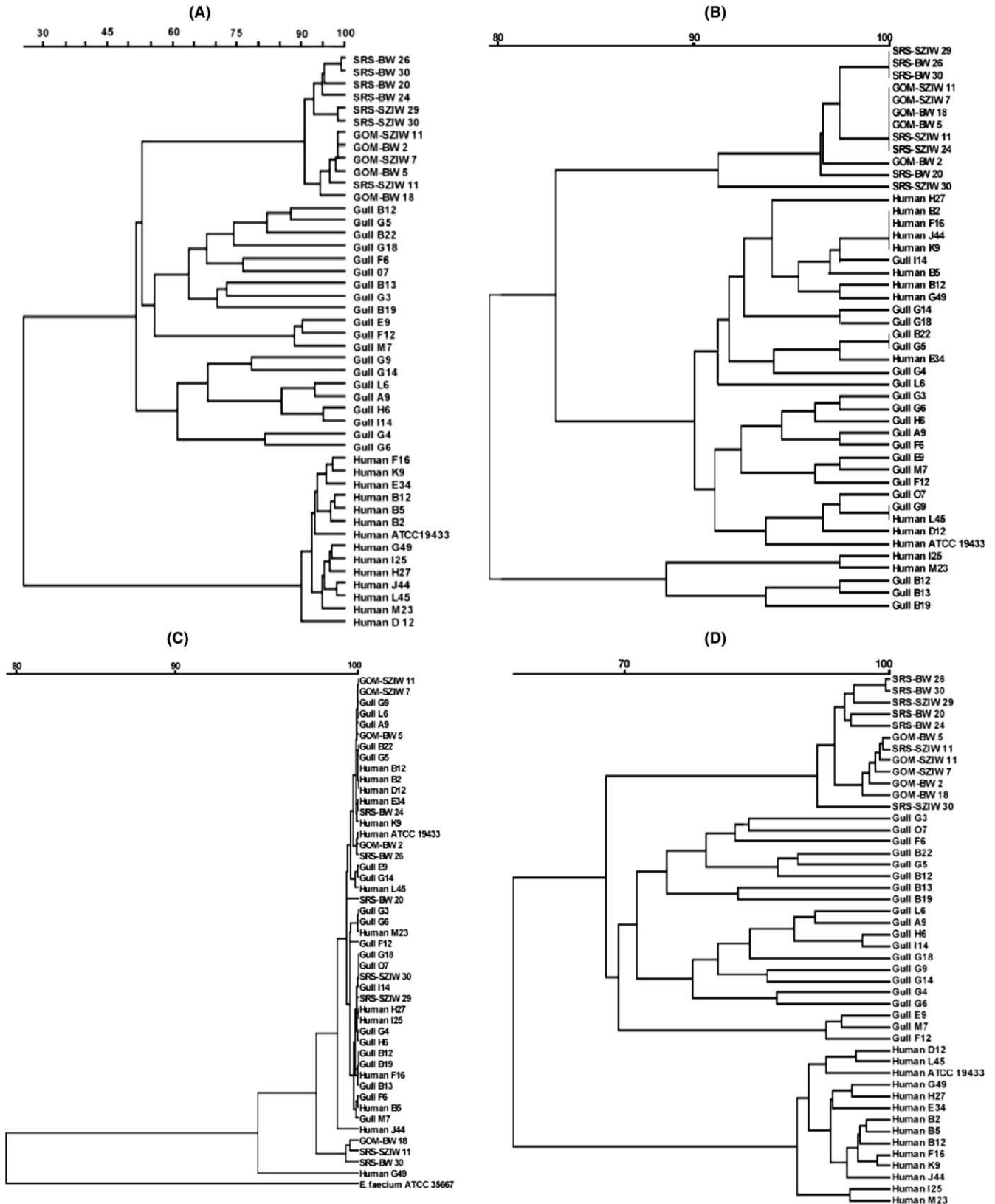


Fig. 3. Dendrograms, produced by the UPGMA clustering method, showing relatedness of *E. faecalis* strains isolated from humans, gulls, and beaches as determined by (A) REP-PCR DNA fingerprint analysis using Pearson's product-moment coefficient, (B) antibiotic resistance analysis (ARA) using simple matching, (C) fatty acid methyl ester (FAME) analysis of membrane lipids using Pearson correlation, (D) REP-PCR and ARA using "take from experiments". GOM: Gulf of Mexico site; SRS: Santa Rosa Sound site; SZIW: swash zone interstitial water; BW: bathing water.

Table 3

Assignment of *E. faecalis* beach isolates to either the human or sea gull source group using average similarity from REP-PCR, ARA and REP-PCR + ARA data sets

Beach isolate ^a	REP-PCR data set Source ^b /ID score ^c /QF ^d	ARA data set Source/ID score/QF	REP-PCR + ARA data set Source/ID score/QF
GOM-SZIW 11	Sea gull/55.0/C	Sea gull/81.1/D	Sea gull/68.1/C
GOM-SZIW 7	Sea gull/54.4/C	Sea gull/81.1/D	Sea gull/67.7/C
GOM-BW 18	Sea gull/50.8/C	Sea gull/81.1/D	Sea gull/66.0/C
GOM-BW 2	Sea gull/55.6/C	Sea gull/80.6/D	Sea gull/68.1/C
SRS-SZIW 30	Sea gull/51.6/C	Human/78.0/E	Sea gull/63.2/C
SRS-SZIW 29	Sea gull/52.5/C	Sea gull/79.2/D	Sea gull/65.8/C
SRS-BW 24	Sea gull/47.0/C	Sea gull/81.1/D	Sea gull/64.0/C
SRS-BW 30	Sea gull/49.8/C	Sea gull/79.2/D	Sea gull/64.5/C
SRS-BW 26	Sea gull/48.7/C	Sea gull/79.2/D	Sea gull/63.9/C
SRS-BW 20	Sea gull/50.7/C	Sea gull/80.0/D	Sea gull/65.4/C
SRS-SZIW 11	Sea gull/54.5/C	Sea gull/81.1/D	Sea gull/67.8/C
GOM-BW 5	Sea gull/55.7/C	Sea gull/81.1/D	Sea gull/68.4/C

^a GOM: Gulf of Mexico site; SRS: Santa Rosa Sound site; SZIW: swash zone interstitial water; BW: bathing water.

^b Source (human or sea gull) assignment.

^c ID score, Identification score is an indication of the confidence of the source assignment, a relative concept (0–100). The higher the score the more confident the assignment.

^d QF (quality factor) indicates how well the isolate fits within the assignment. Grades are A–E. The grade of C is considered a faithful identification whereas a grade of D is considered a doubtful identification.

similarity (AS). Identifications were performed with two single (REP-PCR, ARA) and one composite (REP-PCR + ARA) data sets (Table 3). With one exception, all beach isolates, when analyzed by all data sets, were assigned to the sea gull source group. The exception, isolate SRS-SZIW 30, was assigned to the human source group when analyzed by the ARA data set. The quality factor of this assignment, however, was rated an E which is considered an improbable identification (Bio-numerics Manual v 2.5, 2000). In addition, the quality factors of the remaining assignments analyzed by the ARA data set were also low with each receiving a doubtful identification rating. Despite the low quality factors, higher identification scores were achieved using the ARA data set than when using either the REP-PCR or REP-PCR + ARA data sets. The average identification score using ARA was 80.2 whereas the average scores for the REP-PCR and REP-PCR + ARA assignments were 52.2 and 66.1, respectively. The REP-PCR + ARA assignments were clearly better than either the REP-PCR or the ARA assignments showing higher identification scores than the REP-PCR assignments and retaining a quality factor “C” indicating a faithful identification which was lacking in the ARA assignments “D”.

5. Discussion

This study documents higher densities of enterococci in SZIW than in adjacent bathing waters on Pensacola Beach (Fig. 2). Entrapment may partially account for increased bacteria densities, however, biological factors (nutrients, protection from predation) and physical fac-

tors (particulate matter, periodic wetting and drying, protection from solar irradiation) may not only allow the enhanced survival of bacteria but may actually provide a growth-promoting environmental niche on the beach (Solo-Gabriele et al., 2000).

Recommendations for maintaining the quality of bathing beach waters issued by the EPA are as follows: the geometric mean density of enterococci (five samples taken over a 30-day period) should not exceed 35 colonies per 100 ml in marine waters and any single count exceeding 104 per 100 ml warrants immediate beach closure (US EPA, 1986). We report that three out of five SZIW samples from June and one out of five SZIW samples from August yielded counts which would trigger immediate closure. In contrast, enterococci densities in the adjacent bathing water samples never exceeded the recommended exposure limits. Although concern about health risks from exposure to beach sand has been expressed (Nestor et al., 1984; Mendes et al., 1997), the real extent of this threat to public health or living marine resources (LMR) remains unknown (World Health Organization, 2003).

Our objective was to provide information regarding source of *E. faecalis* isolated during the enumeration phase of this study to gain insight into potential risk due to exposure in this contaminated environment. Source determinations aid in assessing impact of human vs. non-human contamination on illness (Wade et al., 2003; US EPA, 1999). Epidemiological studies that formed the basis of US EPA's recommended sanitary standards for primary contact, correlated exposure to gastroenteritis in bathers. These studies were performed on beaches where sewage was the source of contamination (Cabelli et al., 1982). Although additional epidemi-

ological research is needed, it is widely assumed that exposure to recreational waters possessing high densities of fecal indicator bacteria (*E. coli* and enterococci), originating from wild animal sources such as bird droppings, would not pose the same health risk as if the source of these indicator bacteria were human.

E. faecalis was the predominant *Enterococcus* species isolated from both humans and sea gulls (Table 2). This is in agreement with the results of Wheeler et al. (2002), who reported the host-range of *E. faecalis* was limited to dogs, humans and poultry, and the results of Pourcher et al. (1991) who found *E. faecalis* to be prominent in humans and sea gulls. Interestingly, only about one-third of all enterococci isolated from the beach environment were identified as *E. faecalis* (Table 2). Possible reasons for this higher proportion of other *Enterococcus* species may be that other species of enterococci deposited on the beach, either directly from other warm-blooded animals or indirectly through runoff, persist and grow in this environment.

Finding only a single sub-species of *E. faecalis* in each human and multiple sub-species of *E. faecalis* in individual sea gulls was probably due to the differences in the digestive systems of these two warm-blooded animals. Humans possess a much longer transit time than sea gulls and would tend to harbor an autochthonous population of *Enterococcus*. Flight requires that sea gulls eliminate their wastes rapidly thereby discouraging a stable, resident population of *E. faecalis*. However, finding only *E. faecalis* and no other species of *Enterococcus* in the cloaca of sea gulls suggest that there were additional factors favoring persistence of *E. faecalis*.

The beach isolates analyzed by REP-PCR were as tightly clustered as the human isolates. The beach environment studied here probably does not receive a large, continual input of enterococci. However, once deposited in the water or sand, environmental conditions may have selected for genotypically similar *E. faecalis* through growth and/or survival. The high degree of similarity shown from the REP-PCR fingerprints plus finding that several isolates possessed identical antibiotic resistance patterns suggests that the *E. faecalis* beach isolates were the same or closely related sub-species. Thus, the beach environment, just as the GI tract of humans, seems to exert selective pressure promoting the persistence of certain sub-species of *E. faecalis*. Perhaps, after an appropriate length of time in sand or seawater, the *E. faecalis* population originating from sea gull droppings would closely resemble the *E. faecalis* population found in the beach environment. In other words, the beach isolates may have been a subset of the sea gull isolates selected for by environmental conditions.

As shown in the dendrogram derived from FAME data (Fig. 3C) all *E. faecalis* isolates were >94% similar as determined by Pearson's correlation. An *E. faecium* strain (ATCC 35667) included as control was clearly

separated from the *E. faecalis* phenon. This result demonstrated that FAME only permits a phylogenetic classification to the species level and, therefore, does not possess the resolving power required in the approach to bacterial source-tracking used in this study.

Similarity coefficients used to construct the dendrogram from the composite data set were calculated using the "take from experiments" method. In this straightforward approach the matrices from the individual experiments are averaged according to their defined weights, thus making it useful when comparing two very different techniques. In our analysis equal weights were placed on both experiments because the number of characters (approximately 35 bands and 12 antibiotics using three concentrations) were approximately equal. The dendrogram from the composite data set was similar to the dendrogram produced from the similarity coefficients derived from the REP-PCR experiment in that the beach, sea gull and human isolates grouped into separate phenons.

Only two very small library units were constructed for our study because of the occurrence of many clonal isolates. The inclusion of many clonal isolates within a library artificially raises the average rate of correct classification and should be avoided (M. Sadowsky, personal communication). Thus, we only used isolates which yielded unique banding patterns. Small libraries are not representative of the bacterial diversity in the host population and are subject to random clustering (Whitlock et al., 2002). However, the objective of our study was not to conduct an exhaustive library-based bacterial source-tracking study in the swash zone. Rather, our objectives were to explore the feasibility of using a DNA fingerprinting technique with enterococci, rather than the commonly used *E. coli*, as the indicator organism and determine whether composite data sets add to the confidence of identifications. Nevertheless, the result of this study showed that the beach isolates were comparatively more similar to sea gull isolates than to the human isolates.

In comparison to *E. coli*, additional issues needed to be addressed when using enterococci for source-tracking. For example, in the digestive tracts of warm-blooded animals *E. coli* is the predominant species belonging to the genus *Escherichia* (Conway, 1995; Neill et al., 1994). Not so with *Enterococcus*, as different warm-blooded animals harbor different species belonging to this genus (Pourcher et al., 1991; Wheeler et al., 2002). When using methods that resolve at the level of subspecies (e.g., DNA fingerprinting), a single species of *Enterococcus* must be selected. Selection of a single *Enterococcus* species can present difficulties in library construction in areas contaminated with domestic animal and human waste. For example Wheeler et al. (2002) found *E. faecalis* was predominant in humans, dogs and poultry whereas *E. gallinarum* was

predominant in cattle and deer. In our study *E. faecalis* was the selected species because the beach environment was likely impacted by sea gulls and humans in which *E. faecalis* predominates (Table 2). However, as only 32% of the enterococci isolated from the beach were identified as *E. faecalis*, any conclusions regarding sources of all *Enterococcus* sp. in this environment should be considered with this finding in mind.

Results presented in Table 3 show that the combined data set (REP-PCR + ARA) improved the confidence of the identifications. Although the identification scores were higher in ARA than the other two data sets, the quality factors indicated doubtful identifications. Both the REP-PCR and REP-PCR + ARA data sets produced faithful sea gull identifications, yet the identification scores were consistently higher in the combined data set. Thus, the increased complexity provided by the combined data set allowed for greater confidence in the identifications. Although the use of combined data sets requires a higher level of effort, this approach may be of value when using smaller data sets.

Acknowledgements

The authors gratefully acknowledge the assistance of M. Sadowsky, L. Johnson, P. Dombek, K. Smith, B. Barrett, V. Harwood, and J. Whitlock.

References

- BioNumerics Manual v 2.5, 2000. Applied Maths. Kortrijk, Belgium.
- Cabelli, V.J., Dufour, A.P., McCabe, L.J., Levin, M.A., 1982. Swimming-associated gastroenteritis and water quality. *Am. J. Epidemiol.* 115, 606–616.
- Chabasse, D., Laine, P., Simitzis-Le-Floch, A.M., Martineau, B., el Hourch, M., Beaud, J.P., 1986. Sanitary study of surface water and of the beach of a water sports and leisure complex. *J. Hyg.* 96, 393–401.
- Conway, P.L., 1995. Microbial ecology of the human large intestine. In: Gibson, G.R., Macfarlane, G.T. (Eds.), *Human Colonic Bacteria: Role in Nutrition, Physiology, and Pathology*. CRC Press, Boca Raton, FL, pp. 1–24.
- Dombek, P.E., Johnson, L.K., Zimmerley, S.T., Sadowsky, M.J., 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.
- Dutka-Malen, S., Evers, S., Courvalin, P., 1995. Detection of glycopeptide resistance genotypes and identification to the species level of clinically relevant enterococci by PCR. *J. Clin. Microbiol.* 33, 24–27.
- Harwood, V.J., Whitlock, J., Withington, V., 2000. Classification of antibiotic resistance patterns of indicator bacteria by discriminate analysis: use in predicting the source of fecal contamination in subtropical waters. *Appl. Environ. Microbiol.* 66, 3698–3704.
- Mendes, B., Urbano, P., Alves, C., Lapa, N., Norais, J., Nascimento, J., Oliveira, J.F.S., 1997. Sanitary quality of sands from beaches of Azores islands. *Water Sci. Technol.* 35, 147–150.
- Messer, J.W., Dufour, A.P., 1998. A rapid, specific membrane filtration procedure for enumeration of enterococci in recreational water. *Appl. Environ. Microbiol.* 64, 678–680.
- Neill, M.A., Tarr, P.I., Taylor, D.N., Trofa, A.F., 1994. *Escherichia coli*. In: Hui, Y.H., Gorham, J.R., Murell, K.D., Cliver, D.O. (Eds.), *Foodborne Disease Handbook*. Marcel Dekker Inc., New York, pp. 169–213.
- Nestor, I., Costin-Lazar, L., Sovrea, D., Ionescu, N., 1984. Detection of enteroviruses in sea water and beach sand. *Zentr. Bakter. Mikr. Hyg. ABT 1 (178)*, 527–534.
- Paisley, R., 1996. *MIS Whole Cell Fatty Acid Analysis by Gas Chromatography*. MIDI, Inc., Newark, DE.
- Pourcher, A.M., Devriese, L.A., Hernandez, J.F., Delattre, J.M., 1991. Enumeration by a miniaturized method of *Escherichia coli*, *Streptococcus bovis*, and enterococci as indicators of the origin of fecal pollution of waters. *J. Appl. Bacteriol.* 70, 525–530.
- Ritter, K.J., Carruthers, E., Carson, C.A., Ellender, R.D., Harwood, V.J., Kingsley, K., Nakatsu, C., Sadowsky, M., Shear, B., West, B., Whitlock, J.E., Wiggins, B.A., Wilbur, J.D., 2003. Assessment of statistical methods used in microbial source tracking. *J. Water Health* 01, 209–223.
- Scott, T.M., Rose, J.B., Jenkins, T.M., Farrah, S.R., Lukasik, J., 2002. Microbial source-tracking: current methodology and future directions. *Appl. Environ. Microbiol.* 68, 5796–5803.
- Solo-Gabriele, H.M., Wofert, M.A., Desmarais, T.R., Palmer, C.J., 2000. Sources of *Escherichia coli* in a coastal subtropical environment. *Appl. Environ. Microbiol.* 66, 230–237.
- US Environmental Protection Agency, 1986. *Ambient Water Quality Criteria for Bacteria—1986*. Office of Water Regulations and Standards Division, Washington, DC. EPA-440/5-84-002.
- US Environmental Protection Agency, 1999. *Action Plan for Beaches and Recreational Waters*. Office of Research and Development, Office of Water, Washington, DC. EPA/600/R-98/079.
- US Environmental Protection Agency, 2000. *Method 1600: Improved Enumeration Methods for the Recreational Water Quality Indicators: Enterococci and Escherichia coli*. Office of Science and Technology, Washington, DC. EPA/821/R-97/004.
- Versalovic, J., Schneider, M., de Bruijn, F.J., Lupski, J.R., 1994. Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction methods. *Mol. Cell. Biol.* 5, 25–40.
- Wade, T.J., Pai, N., Eisenberg, J.N.S., Colford, J.M., 2003. Research Reviews—Do US Environmental Protection Agency water quality guidelines for recreational waters prevent gastrointestinal illness? A systematic review and meta-analysis. *Environ. Health Perspect.* 111, 1–39.
- Wheeler Alm, E., Burke, J., Spain, A., 2003. Fecal indicator bacteria are abundant in wet sand at freshwater beaches. *Water Research* 37, 3978–3982.
- Wheeler, A.L., Hartel, P.G., Godfrey, D.G., Hill, J.L., Segars, W.I., 2002. Potential of *Enterococcus faecalis* as a human fecal indicator for microbial source tracking. *J. Environ. Qual.* 31, 1286–1293.
- Whitlock, J.E., Jones, D.T., Harwood, V.J., 2002. Identification of the sources of fecal coliforms in an urban watershed using antibiotic resistance analysis. *Water Res.* 36, 4273–4282.
- Whitman, R., Olyphant, G., Wolcott, M., Haack, S., Horvath, T., Goodrich, M., Nevers, M., 2001. *Characterization of E. coli at 63rd Street Beach*. Report to the Chicago Park District and the City of Chicago, 135 pp.
- World Health Organization, 2003. *Microbiological aspects of beach quality*. In: *Coastal and Fresh Waters. Guidelines for Safe Recreational Water Environments*, vol. 1. Geneva, Switzerland (Chapter 5).