Application of an integrated community analysis approach for microbial source tracking in a coastal creek

Yiping Cao, Laurie C. Van De Werfhorst[1], Bram Sercu[1], Jill L.S. Murray[2] and Patricia A. Holden[1]

Abstract

High fecal indicator bacterial (FIB) concentrations signal urban coastal water quality impairments that can threaten public health. However FIB (total and fecal coliform plus *Enterococcus* sp.) are not specific to human waste and thus microbial source tracking (MST) is employed to assess public health risks and remediation alternatives. Currently, water quality diagnosis requires several simultaneous MST assays. Relatively unexplored is a community analysis approach for MST where the overall microbial community composition is compared, via multivariate analysis, to link sources and sinks of microbial pollution. In this research, an urban coastal creek and drain sampling transect, previously diagnosed as human waste-contaminated, were evaluated for bacterial community composition relative to fecal sources; a laboratory spiking study was also performed to assess method sensitivity and specificity. Multivariate statistical analysis of community profiles clearly distinguished different fecal sources, indicated a high sensitivity for sewage spikes, and confirmed creek contamination sources. This work demonstrated that molecular microbial community analysis combined with appropriate multivariate statistical analyses is an effective addition to the MST tool box.

Introduction

Water quality impairment due to microbial contamination is a serious public health and economic concern (Rabinovici *et al.* 2004, Santo Domingo *et al.* 2007). Fecal indicator bacteria (FIB) such as total coliform, *E. coli*, and *Enterococcus*, routinely used in monitoring and regulation, are not specific to human waste. Assessing public health risks and effective remediation of impaired waters require identifying contributing sources though microbial source tracking (MST) studies (USEPA 2005).

A significant advance for MST has been the development of source-specific single indicator methods. Still, no single indicator method is absolutely effective or clearly superior owing partly to geographic and individual host instability (Griffith *et al.* 2003, Stewart *et al.* 2003). Simultaneously monitoring multiple indicators should better predict public health risks and enable more appropriate MST (Harwood *et al.* 2005), but the costs are high and indicator selection is often uncertain (Stewart *et al.* 2003). Alternatively, culture-independent microbial community analyses, where the whole detectable microbial community is characterized at once, has great potential as an ultimate multiple indicator approach for MST (Wu *et al.* 2010). Community analysis expands from the single-indicator MST concept that different host species select for different

[1]*University of California, Donald Bren School of Environmental Science and Management and Earth Research Institute, Santa Barbara, CA*

[2]*City of Santa Barbara, Santa Barbara, CA*

gut microbial communities due to their unique gut environment and diet choices (Ley *et al.* 2008). Additionally, microbial community analysis integrates biotic and abiotic factors because microbial communities in receiving waters are altered not only by source communities as inoculants, but also by changing water chemistry including nutrient addition.

However, molecular community analysis techniques have been mainly used to develop source-specific single indicators through focusing on individual components within the community (Bernhard and Field 2000, Simpson *et al.* 2004, Dick *et al.* 2005, Soule *et al.* 2006, Baertsch *et al.* 2007, Lu *et al.* 2008, Weidhaas *et al.* 2010), or to describe community diversity in fecal sources (McLellan *et al.* 2010, Shanks *et al.* 2011). A few studies applied community analysis techniques for MST in ambient water (Esseili *et al.* 2008, Unno *et al.* 2010, Wu *et al.* 2010), with limited attention to multivariate analysis of the overall community profile (Esseili *et al.* 2008, Jeong *et al.* 2008). Yet, integrating microbial community analysis with the appropriate multivariate analysis into one overall approach is widely performed in microbial ecology, and proven to be powerful (Eckburg *et al.* 2005, Cao *et al.* 2006) for recognizing the influence of perturbations (biotic and abiotic; Jernberg *et al.* 2005, Córdova-Kreylos *et al.* 2006, Cao *et al.* 2008).

This study therefore investigated the application of an integrated community analysis approach for a field scale MST study. A high-throughput fingerprinting technique, terminal restriction fragment length polymorphism (TRFLP; Liu *et al.* 1997), was selected for this study because TRFLP offers a good balance between information gained, cost, and labor intensity (Schütte *et al.* 2008), and has been frequently used for a wide range of samples (including feces, soil, and water) and for characterizing community responses to perturbations (Li *et al.* 2007, Thies 2007). Past MST studies employing TRFLP mostly focused on selecting potential source-indicative peaks from the overall profiles to then develop single indicator assays (Bernhard and Field 2000) or to discern sources in unknown samples based on the presence of such peaks in corresponding TRFLP profiles (Field *et al.* 2003, Baertsch *et al.* 2007). However, utilization of overall TRFLP community profiles in an integrated fashion for MST has been rare. This study focuses on a coastal creek system where human waste was introduced into a creek by a contaminated storm drain (Sercu *et al.* 2009). The main objective is to demonstrate the feasibility and effective use of combining molecular and statistical microbial ecological, community-based tools for application to MST of a field setting.

Methods

Study Sites and Study Design

The study sites were located in the Mission Creek watershed, in the City of Santa Barbara, CA (Figure 1 and Supplemental Information available at ftp://ftp.sccwrp.org/pub/download/DOCUMENTS/AnnualReports/2012AnnualReport/ar12_21SI.pdf). The study design included laboratory and field studies. In the laboratory study, fecal sources potentially contributing to the watershed were i) sampled to test if community fingerprinting generated by TRFLP could differentiate these sources and ii) spiked into relatively pristine creek water at various concentrations to evaluate the sensitivity of TRFLP. In the field study, ambient waters were sampled from the study sites to discern potential sources of microbial contamination by TRFLP.

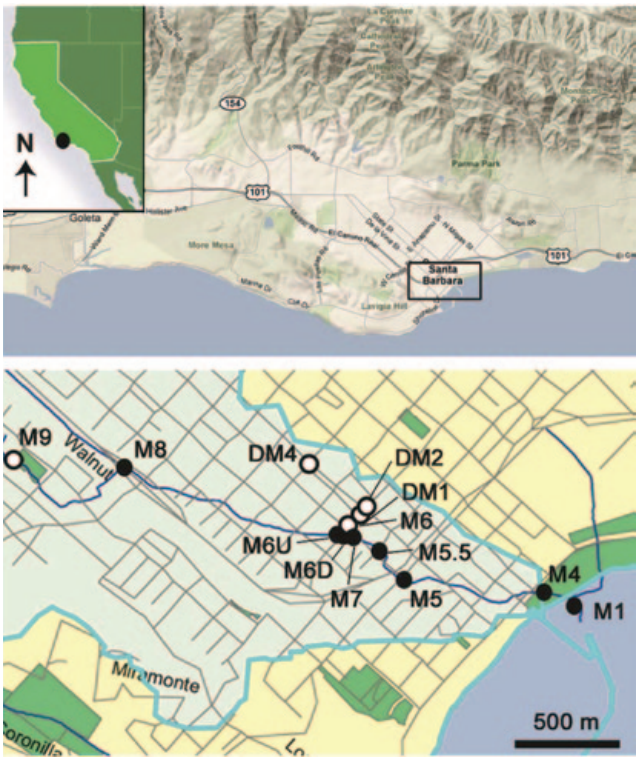

Figure 1. Study sites. Site IDs are as described in the Methods section. Open circles indicate storm drain sites. Map is modified from Figure1 in Sercu et al. 2009.

Fecal Sources and Spiking Experiment

The fecal sources, including sewage, human, gull, and raccoon, were chosen in consultation with the City of Santa Barbara Creeks Division based on sources deemed to be most relevant for the Mission Creek watershed. Sewage samples were collected from the influent of the El Estero Wastewater Treatment Plant in Santa Barbara, at four separate times across years 2004 to 2005. Human, gull, and raccoon feces were collected as described previously, each sample being a composite from ≥3 individuals (Sercu *et al.* 2009). Fecal samples were spiked into reference creek water collected from an upstream, relatively pristine, site that was located on a tributary to Mission Creek. Sewage was spiked into reference water at four dilutions with the lowest corresponding to an *Enterococcus* concentration that would minimally meet the California recreational ocean water quality standard (http://www.cdph.ca.gov/HealthInfo/environhealth/water/Pages/Beaches.aspx). Various combinations of fecal sources were selected for spiking based on anticipated co-occurrence in the field and at realistic, expected levels. Recipes used in preparing the spiked samples are listed in Supplemental Information (SI) Table SI-1 (Supplemental Information is available at ftp://ftp.sccwrp.org/pub/download/DOCUMENTS/AnnualReports/2012AnnualReport/ar10_SI.pdf).

Ambient Sampling

Working from downstream to upstream, seven sites (M1, M4-M9; Figure 1) were sampled once a day for three consecutive days (June 28-30, 2005) at approximately the same time and tidal conditions. Salinity was also measured to characterize the sites (Sercu *et al.* 2009). Two follow-up studies were performed for more spatially and temporally intensive sampling, respectively. On August 2, 2005, samples were taken from six sites that are, listed from downstream to upstream, the creek site M5, another creek site (M5.5, located between M5 and M6), the storm drain site M6, and three upstream drain locations including a continuous deflection separation unit (DM1) and two manholes (DM2, DM4). On August 4, 2005, samples were acquired over three time durations: a) 7:40-9:00, b) 11:15-12:05, and c) 13:45-14:40, from four sites that are, listed from downstream to upstream, the creek site M5.5, a creek site immediately downstream of the storm drain (M6D), the storm drain site M6, and a creek site immediately upstream of the storm drain (M6U). Samples were collected as described previously (Sercu *et al.* 2009).

Microbial Analysis

DNA extraction (Sercu *et al.* 2009) and TRFLP analysis (targeting genes encoding 16S rRNA, (Cao *et al.* 2006) and Supporting Information) were performed as described before. Four single-indicators (total coliform (TC), *E. coli* (EC), *Enterococcus* (ENT), and human-specific *Bacteroides* marker (HBM)) were measured via qPCR as described previously (Sercu *et al.* 2009); their results from the ambient sampling were reported elsewhere (Sercu *et al.* 2009).

Statistical Analysis

Two suitable multivariate statistical methods were selected based on their model assumptions and results presentation to analyze the aligned TRFLP data. See Supporting Information for more details. Detrended correspondence analysis (DCA; ter Braak and Smilauer 2002) and non-metric multidimensional scaling (NMDS; Clarke and Warwick 2001) were performed in CANOCO (Microcomputer Power, Ithaca, NY, USA) and Primer (Primer-E Ltd, U.K.), respectively. DCA results are shown as 2-D plots where samples are positioned according to the (dis) similarity between their TRFLP profiles. Samples closer to each other on the plot have more similar TRFLP profiles than the more distant samples. The DCA plot axes represent latent variables that explain the multivariate TRFLP profiles; the latent variable values for a sample are determined by the sample TRF distributions. The axes are measured in Standard Deviation (SD) units, which indicate how fast the different TRFs emerge and disappear (i.e., taxonomic units turn over) along the axes. If two samples are 4 SD apart along an axis, their TRFLP profiles are considered very different with few TRFs relevant to this axis shared between them (i.e., the TRF distributions of these two samples overlap little along this axis). A statistic called total inertia, which represents the total amount of variation in the TRFLP profiles, is also reported for DCA. The percentage of the total variability explained by each DCA axis is indicated on the DCA plot. The first DCA axis is always longer (in SD units), explains a higher percentage of the total variation, and represents the most important latent variable (or gradient) in the TRFLP profiles included in the analysis. NMDS results are usually presented in 2-D plots as well,

where samples are positioned according to the ranked (dis)similarity between their TRFLP profiles.

TRFLP data from the ambient sampling events were first analyzed independently to discern trends among the ambient samples themselves. Then the reference sewage sample data were co-analyzed with ambient sample data to evaluate TRFLP profile similarities indicative of sewage contamination in ambient samples. Additionally, the NMDS results were mainly used to confirm DCA results and thus mostly are not shown.

Results

Fecal Sources

High FIB concentrations were found in all fecal samples, and neither gull nor raccoon feces were positive for HBM (Table SI-2). TRFLP profiles clearly distinguished the human from animal fecal sources included in this study (Figure 2). Although the sewage samples were acquired over a temporal span of one year, the TRFLP profile data grouped together, along with the human feces composite sample. The first DCA axis explained 33% of the total variation (inertia = 2.88) of the TRFLP profiles, and was considered an animal vs. human source-differentiating axis as the animal- and human-source samples were approximately 8 SD apart, at opposite ends of this DCA axis, indicating few TRFs were shared between them. The second DCA axis explained 18% of the total variation of the TRFLP profiles and may be considered a gull vs. raccoon differentiating axis as samples from these two animal sources were separated along this axis by at least 4 SD. NMDS indicated similar source separations between human/sewage, raccoon, and gull (not shown).

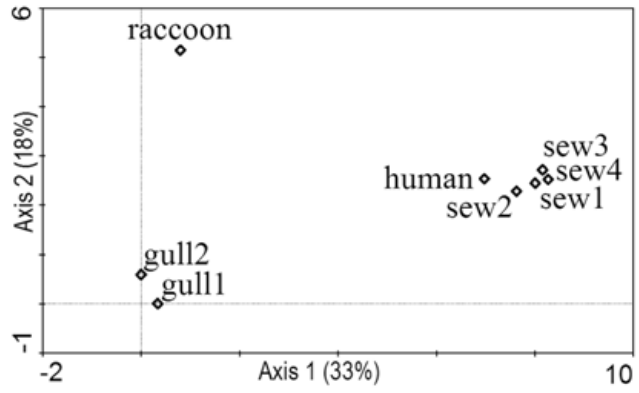

Figure 2. DCA plot for TRFLP profiles from fecal sources (inertia = 2.88). Percentage of total variability explained by each axis is indicated in parentheses.

Spiking Experiment

The single indicator results (Table SI-2) were as expected from the recipes for the spiking experiment (Table SI-1). TRFLP analysis was able to detect sewage and human feces when spiked into reference creek water, but did not detect the raccoon or gull fecal signals at the current spiking conditions in the presence of the co-spiked sewage (Figure 3). The first DCA axis explained 27% of the total variation (inertia = 3.40) of the TRFLP profiles and spanned a length of approximately 9 SD, indicating that this axis represented a significant latent variable that may be considered a differentiation between animal and human wastes versus ambient clean water. All water samples that received sewage spikes, even at the lowest (0.01%) concentration, were adjacent to the sewage sample along the first DCA axis, but distant from the animal fecal samples and the reference water sample. Although some of these sewage-spiked samples also received raccoon and/or gull fecal material, the animal feces did not shift the TRFLP profiles to the same extent as sewage. The second axis explained 13% of the total variation but was short with a length of 2 SD, indicating that the underlying latent variable was less prominent compared to the one represented by the first DCA axis. NMDS also indicated four distinct groups: raccoon, gull, reference water, and all samples containing either human or sewage spikes, with the samples among the last group indistinguishable from each other (Figure SI-1). Raw TRFLP electropherograms corroborated the multivariate findings (Figure SI-2).

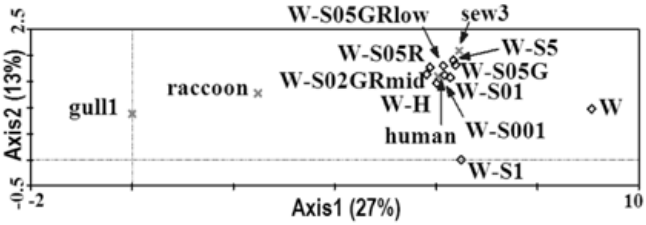

Figure 3. DCA plot for TRFLP profiles from the spiking experiment (inertia = 3.40). Sources (x) used for spiking are labeled as in Figure 2. Reference creek water (diamond) is labeled as W. Spiked samples (diamonds) are labeled as W- followed by the first letters of sources and spiking concentrations (e.g., S05 for 0.5% sewage, GRmid for medium concentrations of gull and raccoon feces). Detailed recipes for the spiking experiment are available in Supplemental Information Table SI-1.

Three-Day Ambient Sampling

There was clearly a spatial and compositional gradient (i.e., ocean to lagoon to urban creek/drain to sewage) in the microbial community composition as revealed by TRFLP profiles, and the day-to-day variation was much smaller than the site-to-site variation (Figure 4, left to right). The first DCA axis explained 18% of the total variation (inertia = 4.55) of the TRFLP profiles, and the site distribution in the DCA plot closely resembled the relative spatial locations of the sites (Figure 1) except for among the drain sites (M9, M6) and the nearby creek sites (M5,M7, M8). The ocean samples from site M1 were furthest apart (approximately 6 SD units) from the sewage samples, indicating little overlap between their TRFLP profiles. On the contrary, the drain and nearby creek sites were more similar to sewage samples based on their TRFLP profiles. In particular, along the first DCA axis, samples that were positive for human *Bacteroides* markers were towards the right side where the four sewage samples were located. The lagoon samples from site M4 were generally more similar to the ocean samples except on June 30 when M4 was more similar to creek samples near drains, which indicated a stronger influence from the upstream drain and creek water on the microbial community at M4. Salinity measurements varied spatially similarly to TRFLP patterns. While the salinity of samples from M1 and drain/creek sites were >32‰ and <1‰, respectively, salinity measurements from M4 were around 8‰ on June 28 and 29, but only 2.3‰ on June 30, reflecting a stronger freshwater influence on the latter date. The second DCA axis, explaining 6% of the total variation, was relatively short (2.5 SD) and appeared to represent the day-to-day variation (at the field sites or the wastewater treatment plant) in TRFLP profiles.

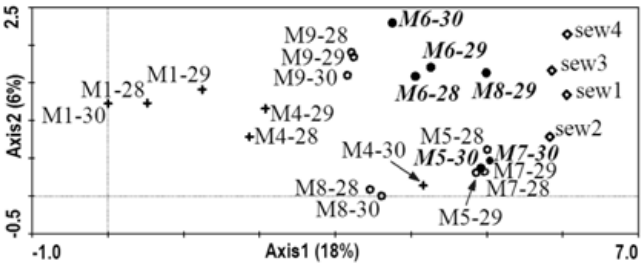

Figure 4. DCA plot for TRFLP profiles from the three-day ambient sampling (inertia = 4.55). Samples are indicated as sampling date (28, 29, and 30 for June 28, 29, and 30, respectively) following site ID. Water samples positive for human Bacteroides marker are indicated by filled circles and bold labels.

NMDS analysis indicated grouping of samples from four categories: ocean, lagoon, drain and nearby creek, and sewage (Figure SI-3).

The addition of sewage sample data did not significantly change the relative position of one ambient sample to another ambient sample in the DCA or NMDS plots. Raccoon and gull samples, when also included in the DCA, were distinctively different (they appeared as an outer group in the DCA plots) from ambient and sewage sample, making patterns among the latter group less visible. Data analysis from the other ambient sampling events (below) yielded similar observations. Therefore, only plots with ambient water and sewage sample are shown.

Near-Drain Spatial Source Sampling

Further spatial sampling (M5, M5.5, M6, DM1, DM2, DM4) indicated that microbial communities within the urban drain network were similar, and that the drain was a potential microbial source to the creek water (Figure SI-4). The TRFLP profiles had relatively small total variation (inertia = 2.02) and DCA axes of intermediate length (3 SD), indicating that the drain and sewage samples were not as different as were the M1 and sewage samples (Figure 4). Specifically, sewage and drain samples were just 1.5~3 SD apart along the first DCA axis (34% of the total variation), indicating somewhat similar communities between them. Drain to creek samples fell along the second DCA axis (18% of the total variation). Although DM1 was physically closest to M6, the microbial community from DM1 was the most different from M6 compared to the upstream manhole samples, suggesting that there may be uniquely selective conditions and/or a hydrologic disconnect between DM1 and those up- and downstream in the drain network. Sites M5 and M5.5 were separated by some spatial distance on the map (Figure 1), yet they were almost indistinguishable on the DCA plot (Figure S4), indicating an in-common strong influence, perhaps from background creek water. Together with the results of human *Bacteroides* marker (e.g., higher frequency of detection and concentrations at M6, DM1, and DM2, and lower frequency of detection and concentrations at downstream sites (Sercu *et al.* 2009)), the DCA plot indicated that the drain network, containing high human waste signals, was contributing to the downstream site, but that the contribution diminished as the creek flowed further downstream. NMDS showed similar groupings by sewage and by drain

and creek samples, and also indicated a gradient among the drain and creek samples (not shown).

Within-Day Temporal Source Sampling

Further temporal sampling (M5.5, M6D, M6, M6U) revealed highly similar microbial communities from within drain site M6 and creek sites around M6, but also a clear short-term temporal variation (Figure SI-5). Although concentrations of single indicators were all high and lacked obvious temporal patterns (Figure SI-6), community profiles of three ambient samples taken close to noon were clearly different from other ambient samples and were more similar to sewage (Figure SI-5) and previously-taken drain network samples (Figure SI-7). This may point to pulsed microbial inputs around noon (the time of sampling). NMDS analysis of TRFLP profiles indicated three groups: sewage, ambient samples collected around noon, and ambient samples collected in earlier morning and later afternoon (not shown).

Discussion

Fecal sources can be clearly distinguished through multivariate analysis of microbial community profiles, and the integrated community analysis approach indicated temporal and host distribution stability in this study (Figure 2). The stability was reflected in the high similarity of TRFLP profiles between the human (composite feces from three individuals) and sewage samples (composite human wastes from thousands of individuals), between the sewage samples (collected at four different months/seasons), and between the composite gull fecal samples (obtained on two separate occasions). Such stability was also observed in another study where TRFLP community profiles from individuals of the same host species were much more similar to each other than to samples from different source species, despite the lack of unique peaks (i.e., single indicators) exclusive to any source species (Fogarty and Voytek 2005). These observations are consistent with the knowledge that microorganisms exist and function interactively, and are selected as a consortium in the gut or other ambient environments (Ley *et al.* 2008). Although individual taxa may be absent or vary among host individuals, such instability at the whole community level should be relatively low. This integrated community analysis approach is therefore a potentially valuable tool especially when single indicators do not exist for certain sources or exhibit significant host and temporal instability (Stewart *et al.* 2003, Harwood 2007). Nevertheless, this approach remains a MST tool and is not practical for routine monitoring programs.

Detection of mixed sources by community analysis, however, depends on the source composition and the resolution of particular community analysis techniques. While sewage signals were detected even at a 0.01% level against the background microbial community in the reference creek water, raccoon and gull fecal signals were lost when mixed with sewage (Figure 3). One explanation is that the spiked raccoon and/or gull feces (0.01 to 0.02g) represented very little microbial DNA compared to that in sewage (4 to 10 ml, Table SI-1). An overwhelming biomass contribution from cattle feces was also regarded as preventing the detection of other fecal sources previously (Field *et al.* 2003). In another experiment where dog, cat feces and septic solids were spiked into relatively feces-free creek water, DCA of TRFLP profiles successfully detected all 3 sources, but failed to detect septic solids at 0.1% and 0.01% (Cao *et al.* 2011). Another explanation is that the resolution of the TRFLP technique used in this study may not be sufficient to distinguish raccoon and gull in presence of sewage. TRFLP resolution using universal *bacterial* primers (this study) can be lower due to a high degree of redundancy (i.e., a single peak in a TRFLP profile can be generated by more than one species (Liu *et al.* 1997)), leading to a higher degree of overlapping patterns between sources and background (Field and Samadpour 2007). Primers that are too specific to a very small group of microorganisms (i.e., higher resolution), however, can reduce the power of community analysis that utilizes, in an integrated fashion, multiple lines of microbial population-based evidence for MST. Community analysis with primer sets targeting smaller phylogenetic (Wery *et al.* 2010) or functional (Esseili *et al.* 2008) groups, or even following culture enrichment steps (Esseili *et al.* 2008), may provide more applicable method resolution and thus warrants further investigation using the integrated community analysis approach.

The integrated community analysis approach, as a comparative and unrestricted MST approach, has some potential important advantages over single indicator approaches. The latter approaches require absolute quantification of each single indicator and therefore is strongly influenced by geographic variation in indicator abundance (Jeter *et al.* 2009) and specificity (Gawler *et al.* 2007). The integrated

community analysis approach; however, utilizes the resemblance of the overall microbial community composition among ambient sites themselves and/or between sites and reference sources. This enables linking sources to sinks in the same local watershed, and naturally combines fate and transport and loading information of many (defined or undefined) single indicators in the system. For example, in this study, on June 30 (Figure 4), the microbial community at M4 (lagoon site) was much more similar to communities at the creek sites M5 and M7, suggesting a stronger input from upstream (including the drain network) towards the downstream sites on that day compared to the two days prior. This may explain the detection of the human *Bacteroides* marker at M5 and M7 on June 30 but not the two days prior, and can obviate a transport pathway plus confirm upstream storm drains as contamination sources (Sercu *et al.* 2009, Wu *et al.* 2010). Additionally, community analysis does not restrict source tracking to, or require knowledge of, the presumed host sources upon which single indicators are selected. This unrestricted nature is particularly useful for spatially tracking contamination sources (even those undefined or non-fecal such as sediment or kelp) and for screening sites for more intensive investigation. In this study, the elevated anthropogenic input around noon time, indicated only by community analysis (Figures SI-4, SI-6), could be investigated further by tracer studies.

To successfully employ an integrated community analysis approach, suitable multivariate statistical techniques are necessary. Although emphasized for community analysis in microbial ecology (Ramette 2007), multivariate analysis and selection of suitable multivariate techniques have received little attention in MST. As multivariate statistical methods differ by their model assumptions and results presentation, improper choices of methods will either hinder data interpretation or generate mathematical artifacts (e.g. if model assumptions are violated) instead of revealing true information from the community profiles (Palmer 2006, Schütte *et al.* 2008). NMDS and DCA were used in this study because i) either there is no (NMDS) or there is an appropriate (DCA) model assumption for our data, ii) in contrast to cluster analysis which forces samples to group (Ramette 2007), neither NMDS nor DCA presumes multivariate groupings of samples iii) using two different techniques aids in confirming major findings. Still, DCA provided more insights than did NMDS because data trends and patterns are interpretable quantitatively using DCA axes while NMDS axes are neither quantitative nor representing trends (i.e., latent variables). Further, NMDS may under-represent sample differences (e.g., Figure SI-1) due to a higher degree of information reduction (Clarke and Warwick 2001). As more diverse samples were included in DCA analyses, the total variation in TRFLP profiles generally increased and the percentage of total variation explained by the first axis decreased; however, meaningful latent variables were revealed (Figure 4) because a long DCA axis that explains a small percentage of the total variation is still informative (ter Braak and Prentice 1988).

In summary, this study demonstrated that a bacterial community profiling technique combined with simple and suitable multivariate statistical analyses can be effectively used for MST. Such an integrated community analysis approach is comparative and unrestricted, enabling both "library-dependent" implementation (using training samples from potential sources) and "spatial source tracking" (locating contamination sources in the field for further source identification). Relatively low cost and approachable data analysis demands for fingerprinting methods, compared to massive sequencing work, also facilitate field-scale implementation where analysis of many ambient samples are required. While this study focused on TRFLP as a technique and tracking human fecal pollution as the objective, more work utilizing other molecular community analysis techniques and involving a wider range of fecal or non fecal (such as sediment and kelp) microbial sources will add insights into the overall MST approach presented in this study.

Literature Cited

Baertsch, C., T. Paez-Rubio, E. Viau and J. Peccia. 2007. Source tracking aerosols released from land-applied class B biosolids during high-wind events. *Applied and Environmental Microbiology* 73:4522-4531.

Bernhard, A.E. and K.G. Field. 2000. Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from fecal anaerobes. *Applied and Environmental Microbiology* 66:1587-1594.

Cao, Y., G.N. Cherr, A.L. Córdova-Kreylos, T.W.-M. Fan, P.G. Green, R.M. Higashi, M.G. LaMontagne,

K.M. Scow, C.A. Vines, J. Yuan and P.A. Holden. 2006. Relationships between sediment microbial communities and pollutants in two California salt marshes. *Microbial Ecology* 52:619-633.

Cao, Y., P.G. Green and P.A. Holden. 2008. Microbial community composition and denitrifying enzyme activities in salt marsh sediments. *Applied and Environmental Microbiology* 74:7585-7595.

Cao, Y., C.H. Wu, G.L. Andersen and P.A. Holden. 2011. Community analysis-based methods. . pp. *in*: C. Hagedorn, A.R. Blanch and V.J. Harwood (eds.), Microbial Source Tracking: Methods, Applications, and Case Studies. Springer. New York, NY.

Clarke, K.R. and R.M. Warwick. 2001. Change in marine communities: An approach to statistical anlaysis and interpretation, 2nd edition. PRIMER-E. Plymouth, MA.

Córdova-Kreylos, A.L., Y. Cao, P.G. Green, H.-M. Hwang, K. Kuivila, M.G. LaMontagne, L.C.V.D. Werfhorst, P.A. Holden and K.M. Scow. 2006. Diversity, composition and geographical distribution of microbial communities in California salt marsh sediments. *Applied and Environmental Microbiology* 72:3357-3366.

Dick, L.K., A.E. Bernhard, T.J. Brodeur, J.W. Santo Domingo, J.M. Simpson, S.P. Walters and K.G. Field. 2005. Host distribution of uncultivated fecal *Bacteroidales* bacteria reveal genetic markers for fecal srouce identification. *Applied and Environmental Microbiology* 71:3184-3191.

Eckburg, P.B., E.M. Bik, C.N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S.R. Gill, K.E. Nelson and D.A. Relman. 2005. Diversity of the human intestinal microbial flora. *Science* 308:1635-1638.

Esseili, M.A., I.I. Kassen and V. Sigler. 2008. Optimization of DGGE community fingerprinting for characterizing *Escherichia coli* communities associated with fecal pollution. *Water Research* 42:4467-4476.

Field, K.G., E.C. Chern, L.K. Dick, J. Fuhrman, J. Griffith, P.A. Holden, M.G. LaMontagne, B. Olson and M.T. Simonich. 2003. A comparative study of culture-independent, library-independent genotypic methods of fecal source tracking. *Journal of Water and Health* 1:181-194.

Field, K.G. and M. Samadpour. 2007. Fecal source tracking, the indicator paradigm, and managing water quality. *Water Research* 41:3517-3538.

Fogarty, L.R. and M.A. Voytek. 2005. Comparison of *Bacteroides-Prevotella* 16S rRNA genetic markers for fecal samples from different animal species. *Applied and Environmental Microbiology* 71:5999-6007.

Gawler, A.H., J.E. Beecher, J. Brandão, N.M. Carroll, L. Falcão, M. Gourmelon, B. Masterson, B. Nunes, J. Porter, A. Rincé, R. Rodrigues, M. Thorp, J. Martin Walters and W.G. Meijer. 2007. Validation of host-specific *Bacteriodales* 16S rRNA genes as markers to determine the origin of faecal pollution in Atlantic Rim countries of the European Union. *Water Research* 41:3780-3784.

Griffith, J.F., S.B. Weisberg and C.D. McGee. 2003. Evaluation of microbial source tracking methods using mixed fecal sources in aqueous test samples. *Journal of Water and Health* 1:141-151.

Harwood, V.J. 2007. Assumptions and limitations of microbial source tracking methods. pp. *in*: J. Santo Domingo and M.J. Sadowsky (eds.), Microbial Source Tracking. ASM Press. Washington, DC.

Harwood, V.J., A.D. Levine, T.M. Scott, V. Chivukula, J. Lukasik, S.R. Farrah and J.B. Rose. 2005. Validity of the indicator organism paradigm for pathogen reduction in reclaimed water and public health protection. *Applied and Environmental Microbiology* 71:3163-3170.

Jeong, J.-Y., K.-I. Gil, K.-H. Lee and J.-O. Ka. 2008. Molecular identification of fecal pollution sources in water supplies by host-specific fecal DNA markers and terminal restriction fragment length polymorphisms profiles of 16S rRNA genes. *The Journal of Microbiology* 46:599-607.

Jernberg, C., Å. Sullivan, C. Edlund and J.K. Jansson. 2005. Monitoring of antibiotic-induced alterations in the human intestinal microflora and detection of probiotic strains by use of terminal restriction fragment length polymorphism. *Applied and Environmental Microbiology* 71:501-506.

Jeter, S.N., C.M. McDermott, P.A. Bower, J.L. Kinzelman, M.J. Bootsma, G.W. Goetz and S.L. McLellan. 2009. *Bacteroidales* diversity in Ring-Billed Gulls (*Laurus delawarensis*) residing at Lake

Michigan beaches. *Applied and Environmental Microbiology* 75:1525-1533.

Ley, R.E., M. Hamady, C. Lozupone, P.J. Turnbaugh, R.R. Ramey, J.S. Bircher, M.L. Schlegel, T.A. Tucker, M.D. Schrenzel, R. Knight and J.I. Gordon. 2008. Evolution of mammals and their gut microbes. *Science* 32:1647-1651.

Li, F., M.A.J. Hullar and J.W. Lampe. 2007. Optimization of terminal restriction fragment polymorphism (TRFLP) analysis of human gut microbiota. *Journal of Microbiological Methods* 68:303-311.

Liu, W.-T., T.L. Marsh, H. Cheng and L.J. Forney. 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Applied and Environmental Microbiology* 63:4516-4522.

Lu, J.R., J.W. Santo Domingo, R. Lamendella, T. Edge and S. Hill. 2008. Phylogenetic diversity and molecular detection of bacteria in gull feces. *Applied and Environmental Microbiology* 74:3969-3976.

McLellan, S.L., S.M. Huse, S.R. Mueller-Spitz, E.N. Andreishcheva and M.L. Sogin. 2010. Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent. *Environmental Microbiology* 12:378-392.

Palmer, M.W. Ordination methods for ecologists: The ordination webpage. http://http://ordination.okstate.edu/

Rabinovici, S.J.M., R.L. Berhknopf, A.M. Wein, D.L. Coursey and R.L. Whitman. 2004. Economic and health risk trade-offs of swim closures at a Lake Michigan Beach. *Environmental Science & Technology* 38:2737-2745.

Ramette, A. 2007. Multivariate analyses in microbial ecology. *FEMS Microbiology Ecology* 62:142-160.

Santo Domingo, J.W., D.G. Bambic, T.A. Edge and S. Wuertz. 2007. Quo vadis source tracking? Towards a strategic framework for environmental monitoring of fecal pollution. *Water Research* 41:3539-3552.

Schütte, U.M.E., Z. Abdo, S.J. Bent, C. Shyu, C.J. Williams, J.D. Pierson and L.J. Forney. 2008. Advances in the use of terminal restriction fragment length polymorphism (T-RFLP) analysis of 16S rRNA genes to characterize microbial communities. *Applied Microbiology & Biotechnology* 80:365-380.

Sercu, B., L.C. Van De Werfhorst, J. Murray and P.A. Holden. 2009. Storm drains are sources of human fecal pollution during dry weather in three urban southern California watersheds. *Environmental Science & Technology* 43:293-298.

Shanks, O.C., C.A. Kelty, S. Archibeque, M. Jenkins, R.J. Newton, S.L. McLellan, S.M. Huse and M.L. Sogin. 2011. Community structures of fecal bacteria in cattle from different animal feeding operations. *Applied and Environmental Microbiology* 77:2992-3001.

Simpson, J.M., J.W. Santo Domingo and D.J. Reasoner. 2004. Assessment of equine fecal contamination: the search for alternative bacterial source-tracking targets. *FEMS Microbiology Ecology* 47:65.

Soule, M., E. Kuhn, F. Loge, J. Gay and D.R. Call. 2006. Using DNA microarrays to identify library-independent markers for bacterial source tracking. *Applied and Environmental Microbiology* 72:1843-1851.

Stewart, J.R., R.D. Ellender, J.A. Gooch, S. Jiang, S.P. Myoda and S.B. Weisberg. 2003. Recommendations for microbial source tracking: Lessons from a methods comparison study. *Journal of Water and Health* 1:225-231.

ter Braak, C.J.F. and I.C. Prentice. 1988. A theory of gradient analysis. . *Advances in Ecological Research* 18.

ter Braak, C.J.F. and P. Smilauer. 2002. CANOCO reference manual and CanoDraw for windows user's guide: Software for canonical community ordination (version 4.5). Microcomputer Power. Ithaca, NY, USA.

Thies, J.E. 2007. Soil microbial community analysis using terminal restriction fragment length polymorphisms. *Soil Science Society of America Journal* 71:579-591.

United States Environmental Protection Agency (USEPA). 2005. Microbial Source Tracking Guide Document. USEPA Office of Research and Development. Cincinnatti, OH.

Unno, T., J. Jang, D. Han, J.H. Kim, M.J. Sadowsky, O.-S. Kim, J. Chun and H.-G. Hur. 2010. Use of barcoded pyrosequencing and shared OTUs to determine sources of fecal bacteria in watersheds. *Environmental Science & Technology* 44:7777-7782.

Weidhaas, J.L., T.W. Macbeth, R.L. Olsen, M.J. Sadowsky, D. Norat and V.J. Harwood. 2010. Identification of a *Brevibacterium* marker gene specific to poultry litter and development of a quantitative PCR assay. *Journal of Applied Microbiology* 109:334-347.

Wery, N., C. Monteil, A.M. Pourcher and J.J. Godon. 2010. Human-specific fecal bacteria in wastewater treatment plant effluents. *Water Research* 44:1873-1883.

Wu, C.H., B. Sercu, L.C. Van De Werfhorst, J. Wong, T.Z. DeSantis, E.L. Brodie, T.C. Hazen, P.A. Holden and G.L. Andersen. 2010. Characterization of coastal urban watershed bacterial communities leads to alternative community-based indicators. *PLoS One* 5:e11285. doi:11210.11371/journal.pone.0011285.

Acknowledgements

This research was supported by the City of Santa Barbara through Measure B funding, and by the Switzer Foundation through a Leadership Grant to the City. The NSF-funded Santa Barbara Long Term Ecological Research project (NSF OCE 9982105 and OCE 0620276) is also acknowledged. Additional assistance from the City of Santa Barbara was from Jill Zachary, Tim Burgess, Harry Slicker, Steve Mack, Rebecca Bjork, and staff of the El Estero Wastewater Treatment Plant. Nancy Callahan of the Santa Barbara Wildlife Care Network provided the source of raccoon feces. Santa Barbara Cottage Hospital provided the human feces. The authors would like to thank Steve Weisberg, Ken Schiff, John Griffith, and Donna Ferguson for helpful review comments.

Supplemental Information

Supplemental Information is available at ftp://ftp.sccwrp.org/pub/download/DOCUMENTS/AnnualReports/2012AnnualReport/ar12_21SI.pdf