Recommendations following a multi-laboratory comparison of microbial source tracking methods

Jill R. Stewart[1], Alexandria B. Boehm[2], Eric A. Dubinsky[3], Theng-Theng Fong[4], Kelly D. Goodwin[5], John F. Griffith[6], Rachel T. Noble[1,7], Orin C. Shanks[8], Kannappan Vijayavel[9,10] and Stephen B. Weisberg[6]

Abstract

Microbial source tracking (MST) methods were evaluated in the Source Identification Protocol Project (SIPP), in which 27 laboratories compared methods to identify host sources of fecal pollution from blinded water samples containing either one or two different fecal types collected from California. This paper details lessons learned from the SIPP study and makes recommendations to further advance the field of MST. Overall, results from the SIPP study demonstrated that methods are available that can correctly identify whether particular host sources including humans, cows and birds have contributed to contamination in a body of water. However, differences between laboratory protocols and data processing affected results and complicated interpretation of MST method performance in some cases. This was an issue particularly for samples that tested positive (non-zero C_t values) but below the limits of quantification or detection of a PCR assay. Although false positives were observed, such samples in the SIPP study often contained the fecal pollution source that was being targeted, i.e., the samples were true positives. Given these results, and the fact that MST often requires detection of targets present in low concentrations, we propose that such samples be reported and identified in a unique category to facilitate data analysis and method comparisons. Important data can be lost when such samples are simply reported

[1]*University of North Carolina, Gillings School of Global Public Health, Department of Environmental Sciences and Engineering, Chapel Hill, NC*

[2]*Stanford University, Department of Civil & Environmental Engineering, Environmental and Water Studies, Stanford, CA*

[3]*University of California, Lawrence Berkeley National Laboratory, Earth Sciences Division, Berkeley, CA*

[4]*Clancy Environmental Consultants, a Tetra Tech Company, St. Albans, VT*

[5]*National Oceanic and Atmospheric Administration, Atlantic Oceanographic and Meteorological Laboratory, Miami, FL*

[6]*Southern California Coastal Water Research Project, Costa Mesa, CA*

[7]*University of North Carolina, Institute of Marine Sciences, Morehead City, NC*

[8]*US Environmental Protection Agency, Office of Research and Development, National Risk Management Research Laboratory, Cincinnati, OH*

[9]*Wayne State University, Department of Biological Sciences, Detroit, MI*

[10]*Ottawa County Health Department, Environmental Health Division, Holland, MI*

as positive or negative. Actionable thresholds were not derived in the SIPP study due to limitations that included geographic scope, age of samples, and difficulties interpreting low concentrations of target in environmental samples. Nevertheless, the results of the study support the use of MST for water management, especially to prioritize impaired waters in need of remediation. Future integration of MST data into quantitative microbial risk assessments and other models could allow managers to more efficiently protect public health based on site conditions.

Introduction

Methods are under development to provide resource managers with tools to identify sources of fecal contamination in surface waters, a field of science called microbial source tracking (MST; Griffith *et al.* 2003, Field and Samadpour 2007, Stoeckel and Harwood 2007). Some of the most promising methods were recently compared in a multiple laboratory study titled the Source Identification Protocol Project (SIPP), designed to help the State of California identify appropriate source tracking technologies for water quality monitoring and assessment. To date, the SIPP performance study represents the largest, multiple laboratory effort to assess the effectiveness of an array of molecular MST methods (Boehm *et al.* 2013).

The SIPP study was designed to identify the most sensitive and specific methods capable of identifying human and other sources of fecal contamination in water. Samples (n = 64) were prepared by mixing a single source or two sources of fecal material from California animal and human sources and adding this fecal material to water samples collected in southern California. After the samples were prepared, they were filtered according to sample processing guidelines for the assays, and samples were blinded and shipped frozen to 27 research laboratories active in the MST field. A total of 41 MST methods were included in the study, some of which were run using standard operating protocols (SOPs) shared by up to seven laboratories. Each participating laboratory then submitted their results to study organizers before the composition of samples was unblinded.

Results from the SIPP study - detailed in Boehm *et al.* 2013, Cao *et al.* 2013, Ebentier *et al.* 2013, Harwood *et al.* 2013, Layton *et al.* 2013, Schriewer *et al.* 2013, Sinigalliano *et al.* 2013, and Wang *et al.* 2013 - highlight a number of trends in the MST field. First, with the exception of methods targeting bacteriophages, none of the approaches selected by SIPP participants relied on culturing of microorganisms or employed a cultivation step. This is in stark contrast to a previous blinded MST method performance study completed 10 years ago involving 22 participating laboratories in which 50% (6 of 12) of selected approaches incorporated cultivation (Griffith *et al.* 2003). Second, PCR-based methods were the technology of choice. Almost all of the methods tested by MST experts utilized an end-point PCR or quantitative PCR (qPCR) step. Third, the majority of these qPCR methods were focused on identification of human-source contamination. Human-associated methods were tested most often (n = 20), followed by ruminant/cattle (n = 7), gull (n = 4), pig (n = 3), dog (n = 2), and horse (n = 1). Finally, a small number of laboratories tested community-based approaches (microarray and terminal restriction fragment length polymorphism [T-RFLP]) that were not restricted to a particular animal group or a single microorganism, but were able to utilize unique patterns in microbial communities for the identification of all animal fecal pollution sources tested in the study. These community-based approaches were highly specific, but not yet sensitive enough for most recreational water applications (Boehm *et al.* 2013). There remains strong interest in continuing to develop such approaches for MST applications in the near future (Cao *et al.* 2013, Shanks *et al*. 2013).

The purpose of this paper is to pave a path forward for the field of MST following completion of the SIPP study. Lessons learned and study limitations are summarized and used to define next steps for research. Potential applications for the most sensitive and specific methods identified by the SIPP study are also considered, along with other implications for water quality management.

Lessons Learned from the SIPP Study

The primary goal of the SIPP study was to characterize the performance of contributed MST methods based on sensitivity and specificity metrics to identify top performing technologies for monitoring. Based on the metrics, top performing assays included HF183 (human), CF193 and Rum2Bac (ruminant), CowM2 and CowM3 (cow), BacCan (dog), Gull2SYBR and LeeSeaGull (gull), PF163 and pigmtDNA (pig) and HoF597 (horse; Boehm *et al.* 2013). However, the SIPP study also demonstrated that a variety of factors may change the apparent performance of a given MST method (Table 1).

Important factors included lack of standardization across all aspects of a protocol (Boehm *et al.* 2013, Ebentier *et al.* 2013, Ervin *et al*. 2013) method of fecal concentration measurement and normalization (Ervin *et al*. 2013), and the definition and determination of assay limit of detection and limit of quantification (Layton *et al.* 2013).

Limits of Detection and Quantification

The SIPP study demonstrated that there can be considerable differences in results among laboratories, even among laboratories using the same protocols. These differences may be attributable to differences in PCR platforms, PCR reagents, reference DNA standards, protocols used for nucleic acid isolation, as well as the proficiency of technicians using a particular method (Ebentier *et al.* 2013). This is an area that needs to be addressed as these methods move from usage in individual research laboratories to wider-scale implementation in regulatory programs. A major factor requiring better standardization is the interpretation of positive results at low target concentrations that are near the defined limit of detection of the assay.

Several papers report different definitions and interpretations of the limit of detection (LOD) and lower limit of quantification (LLOQ), with consequences to assay performance (Boehm *et al.* 2013, Sinigalliano *et al.* 2013, Ebentier *et al.* 2013, Layton *et al.* 2013, Raith *et al.* 2013, Schriewer *et al.* 2013). The LOD is the value below which a target cannot be detected reliably while the LLOQ provides a threshold where a quantitative value can be estimated (Burns and Valdivia 2008, CODEX 2010). In general, such criteria are assessed by some form of replicate analysis or probit analysis (Spearman-Karber method (AOAC 2006, Ambruster and Pry 2008, Burd 2010). For example, the LOD50 (amount at which 50% of tests are positive) is often used in microbiological analysis (AOAC 2006), whereas the LOD95 (amount at which 95% of tests are positive) often is used in chemical analysis and PCR applications for food and clinical diagnostics (CODEX 2010, Burd 2010).

The LLOQ typically is defined as the lowest reliably detected concentration that meets some defined criterion for precision. For example, the LLOQ for qPCR has been technically defined as the LOD + 2 standard deviations (Armbruster and Pry 2008, Burd 2010, Harwood *et al*. 2011, Staley *et al*. 2012), although different definitions have been used. Among the core SIPP laboratories, the LLOQ was defined as the lowest concentration for which all of the samples in a standard curve yielded quantifiable data when analyzed using a four-point, triplicate standard curve (Boehm *et al.* 2013).

Detection or quantification of low target cell or copy concentration with qPCR/PCR is typical when

Table 1. Recommendations for next steps to advance the field of microbial source tracking based on lessons learned from and study limitations of the Source Identification Protocol Project (SIPP).

SIPP Lessons Learned	Associated Issues	Recommendation
Differences were reported among labs using the same protocols	Inter-laboratory variability	Implement standard protocols including criteria for determining positive and negative results
Different definitions are being used to define an assay limit of detection or quantification	Inconsistent data analysis	Determine assay limits and report methods used (e.g., use of an LOD50* as recommended by AOAC 2006)
Different labs have different interpretations for samples testing positive below the defined limits of an assay	Inconsistent data analysis	Classify these questionable samples as "detectable but not quantifiable (DNQ)" or "detectable but below the limit of detection (DBLOD)"
SIPP Study Limitations	**Associated Issues**	**Recommendation**
Matrix effects not tested	Poor representation of real-world conditions	Test challenge samples in more complex waters
Differential fate and transport not tested	Poor representation of real-world conditions	Test challenge samples in controlled microcosms or mesococsms
Only fecal sources relevant to coastal California were included	Applicability to other geographic regions	Test challenge samples from additional regions; include additional host sources

*LOD50 – amount at which 50% of tests are positive

analyzing ambient water samples for specific microbial contaminants and markers because of the effects of dilution as well as marker degradation. Therefore, LOD and LLOQ calculations are an important part of assay validation (Burns and Valdivia 2008, CODEX 2010), as are decisions about how to interpret positive results that are below the defined limit of detection or quantification of an assay (Layton *et al.* 2013). The MST community should reach consensus on the most appropriate method to define LOD and LLOQ. The medical and food science fields may be able to serve as guides (AOAC 2006, Bustin *et al.* 2009, CODEX 2010). In the meantime, those using MST assays should be transparent about methods of data processing and classification of samples that were detected but below the limit of quantification or detection, including the classification of samples in which some replicates fall above but others fall below a threshold. A more statistically rigorous approach needs to be undertaken to determine how best to classify weak PCR reactions for MST targets and to better elucidate the number of replicates needed to resolve equivocal results.

Based on the SIPP study, it is proposed that samples for which a molecular target is detected (non-zero C_t values) but below the limits of quantification be reported in a unique category to facilitate data analysis and method comparisons. It is critical that the result be accompanied by appropriate demonstration of negative results in field blanks, method blanks, no template controls, and negative extraction controls, all of which should be run routinely during MST sample analyses. These samples can be categorized as "detectable but not quantifiable" (DNQ) when a signal is detected between the LLOQ and the LOD (Figure 1). This term and the analogous term "detectable/below the limit of quantification (detectable/BLOQ)" have been offered previously in peer-reviewed literature (Laperche *et al.* 2011, Harrington *et al.* 2012). A sample in which the signal is below the LOD but above background signal noise (Figure 2) can be reported as detectable below the limit of detection (DBLOD). In addition, samples that did not amplify can be reported as not detected (ND) in order to clearly distinguish between samples with and without a non-zero C_t value. This approach to classify samples preserves important information while recognizing that low target concentrations can produce variable results in replicate reactions due to the Poisson distribution effect (AOAC 2006). In addition to simply re-analyzing the extract from weak

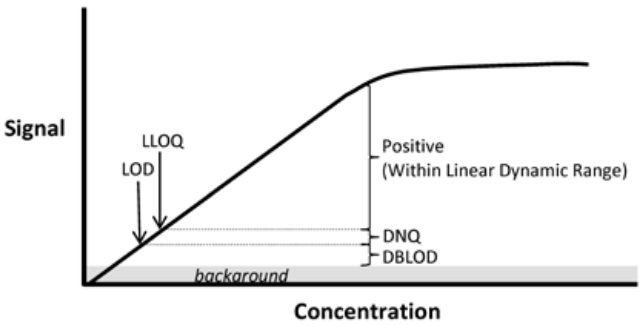

Figure1. Analytical limits to qPCR assays include an assay limit of detection (LOD) and lower limit of quantification (LLOQ). For MST applications, samples whose signal is between the LLOQ and the LOD can be reported as detectable but not quantifiable (DNQ). Samples whose signal is below the LOD but above the background fluorescence can be reported detectable below the limit of detection (DBLOD). For qPCR assays, signal can be measured using relative fluorescence units (RFU). and concentration can be measured by the PCR cycle number. Most PCR platforms can also be programed to subtract the fluorescent background from analysis.

reactions, resources permitting, application of additional MST assays to the sample may help resolve equivocal results. For example, using an assay that targets the same source but with different sensitivity and specificity characteristics can provide added confidence to sample classification. Sequencing the PCR amplicon also will verify that the target was recovered.

Information Needed to Properly Compare and Interpret MST Results

The SIPP study highlighted essential information needed to properly compare results between MST studies and laboratories. It is recommended that MST scientists take care to report the performance parameters of their methods. Most MST methods are PCR-based and should use the minimum information needed for publication of quantitative PCR experiments (MIQE) guidelines (Bustin *et al.* 2009). Many MIQE guidelines could also be adapted easily for other MST technologies that are not based on PCR amplification.

Among the MIQE guidelines, a few parameters related to sample preparation and detection and quantification of MST targets are particularly important to MST studies that often target low copy numbers of a gene from complex samples (Table 2). These parameters include reporting the yield of DNA from

extraction procedures, the reference material used to generate a standard curve, and the approach taken to test for effects from PCR inhibitors common to water samples. While the MIQE guidelines provide a comprehensive list of parameters that should be reported with qPCR results, this condensed list highlights some of the parameters that are most critical for evaluation and comparison of MST results. This condensed list is also practical to include routinely in any report of MST findings without relegating descriptions to supplementary material.

Major Limitations of the SIPP Study

Impact of Environmental Sample Matrix

In the SIPP study, MST methods were tested with reference fecal pollution sources suspended in an artificial freshwater matrix. However, matrix effects could alter performance of MST technologies that include a PCR amplification step because amplification interference can vary for different primer/probe/genetic target combinations on the same DNA extract (Huggett *et al.* 2008). An environmental matrix can also significantly affect sample processing efficiency. DNA recovery and purification can be inefficient and variable from one sample to the next (Wilson 1997, Mumy and Findaly 2004, Stoeckel et al. 2009). In addition, reference samples were well-mixed in the SIPP study, and contained at most a mixture of two fecal sources. In contrast, environmental samples are heterogeneous mixtures (Manter *et al.* 2010) and frequently contain dilute and varying amounts of multiple fecal sources in addition to native microorganisms and substances that could interfere with PCR amplification. Matrix effects on nucleic acid extraction recovery and PCR inhibition need to be better characterized. Challenge samples like those tested in the SIPP study should be tested in more complex waters, including marine waters and waters high in humic acids.

Table 2. Parameters critical for evaluating and comparing results from microbial source tracking studies. Mention of trade names or commercial products are provided as examples and do not constitute endorsement or recommendation for use.

Analytical Step	Critical Parameters to be Reported with Results	Examples of Common Practice in MST
Sample Preparation	Volume of water analyzed and filtration method(s)	500 ml sample filtered with HA mixed cellulose ester membrane (pore size of 0.45 µm) (EMD Millipore Corporation, Billerica, MA)
	Method used to extract nucleic acids	DNA extracted using the MoBio PowerSoil® DNA Isolation Kit (MO BIO Laboratories, Inc., Carlsbad, CA)
	DNA yield and purity	Measured with a Nanodrop (Thermo-Scientific, Wilmington, DE) or Qubit Quantit Fluorometer (Life Technologies, Grand Island, NY)
Detection of MST Target	Instrument used for amplification	Use of a thermocycler like the ABI StepOnePlus; Biorad CFX; Cepheid SmartCycler
	Evidence for absence of inhibition	Use of internal controls; serial dilution of DNA extracts
	Definition of positive detection	Detection within the linear dynamic range of the method accompanied by expected results from controls
	Definition of limit of detection (LOD)	Lowest dilution where ≥50% of replicates are detected
Quantification of MST Target	Calibration curve with slope, y-intercept, r^2, and efficiency amplification	Y = MX +B; amplification efficiency (E) = $(10^{-1/slope})$-1
	Reference material used to generate a standard curve	Genomic DNA; plasmid DNA; RNA transcripts
	Instrument used to quantify DNA or RNA in standards	Nanodrop (Thermo-Scientific, Wilmington, DE); Qubit Quantit Fluorometer (Life Technologies, Grand Island, NY)
	Evidence for absence of contaminating DNA	No template controls; extraction method blanks; field blanks
	Definition of quantifiable replicate reactions	All replicates must be within 0.5 C_t of each other
	Evidence for absence of partial or complete inhibition	Internal amplification control spike
	Definition of lower limit of quantification (LLOQ)	Lowest concentration where ≥50% of replicates could be quantified with acceptable precision

Geographic Distribution of Reference Fecal Pollution Sources

The SIPP reference samples contained fecal material collected solely in the State of California. Several of the fecal sources, therefore, would be expected in the urban watersheds of coastal California, but other fecal sources such as those from wildlife, inland birds and chickens, cats, and pigs are possible. The ability to apply SIPP findings to other geographic regions was not tested, but it is clear that the scientific approach used to conduct the study can provide some important information to the MST community globally. Most studies that report fecal indicator bacteria concentrations are performed on a watershed or regional scale (Layton *et al*. 2006, McQuaig *et al*. 2006, Kildare *et al*. 2007, Okabe *et al*. 2007, Silkie and Nelson 2009, Wang *et al*. 2010). However, a study by Shanks and colleagues (Shanks *et al.* 2010b) reports some consistency in the distribution of general fecal indicators and many human-associated qPCR genetic markers in untreated sewage collected from various locations across the United States. In contrast, studies evaluating fecal microbial communities and the distribution of cattle-associated genetic markers within and between cattle populations have concluded that differences in animal feeding practices drastically alter community structure and the shedding of fecal indicator bacteria (Shanks *et al*. 2010a, Shanks *et al*. 2011). Based on these findings, it is likely that patterns in host-associated genetic marker shedding across geographic space will vary depending on animal source, geography, and genetic marker. Furthermore, when using viral pathogens as markers of human fecal contamination, there is a clear seasonality to infection in humans that can cause strong temporal variation in the concentration or load of viral particles in sewage during any specific season. Therefore, validation studies for all MST assays need to be performed in other geographic areas to determine the sensitivity and specificity of the assays prior to application in any given geographic region. This is an important step that can be very useful in designing appropriate studies for a particular area during a specific period of time. A strong validation study conserves valuable time and resources by permitting the generation of useful quantitative results for specific markers, and reduces the instance of pan-negative results for a study, which are of little use for prioritization or source identification.

Influence of Fate and Transport

One of the biggest impediments to extrapolating the SIPP study results is the use of fresh fecal material to inoculate water samples; in actuality, fecal material in receiving water often ranges in age. For example, fecal material from wildlife and livestock can remain dry in watersheds for days or weeks until washed into waterbodies by stormwater runoff. Human fecal material in a septic tank is typically contributed daily, but the mixture is an integration of fecal material and chemicals contributed over weeks to months. Even material deposited directly in water can take several days to be transported to areas that are routinely monitored for water quality. Differential fate and transport of pathogens, MST markers, and fecal indicator bacteria is a major concern for environmental health microbiologists. Differential fate and transport is expected to be particularly problematic if MST markers degrade at a different rate compared to fecal indicator bacteria such as *Enterococcus* spp. (Wang *et al.* 2013). Furthermore, the use of molecular methods to quantify MST markers as compared to culture-based methods for enumeration of fecal indicator bacteria causes confounding information, as one approach will quantify all target DNA in a sample, and the other only the metabolically active subset of organisms. Slower degradation of MST markers could lead to an overestimation of the importance of a particular fecal source. Faster degradation could potentially lead to incorrect conclusions about the absence of a particular fecal source (Walters *et al*. 2009, Schulz and Childers 2011). A few recent studies have begun to address this issue by examining degradation characteristics of individual general fecal indicator and molecular MST markers (Walters *et al*. 2009, Green *et al*. 2011, Rogers *et al*. 2011), but these studies need to be extended to include samples with known sources under controlled conditions, and also to test dynamics in a range of different receiving water body types. Differential rates of inactivation due to enzymes and UV light need to be studied. Interactions with sediments, and grazing rates among FIB, MST markers and pathogens also need to be better understood. This may be achievable through the use of microcosms or membrane diffusion chambers deployed in the field (Haznedaroglu *et al*. 2009, Schinner *et al*. 2010, Bae and Wuertz 2012).

Implications of the SIPP Study for Management

The SIPP study represents an important step toward MST method implementation by bringing experts together to address key challenges and by identifying a selection of useful tools for water quality applications (e.g., HF183, as reported in Boehm *et al.* 2013).

Application of MST to Real World Water Quality Problems

It is critical to advance methods that can identify sources of fecal contamination in water. A large number of United States waters are categorized as impaired (303(d) listed), challenging municipalities to meet water quality objectives for designated uses and to understand the factors that control FIB concentrations. A variety of remediation strategies can be undertaken to try to improve impaired waters, including structural improvements and best management practices. In general, such strategies are costly (millions to billions of US dollars), and MST can be a valuable tool in providing guidance on mitigation strategies. For example, MST coupled with physical tracing techniques (e.g., dye or smoke) has allowed for leaking infrastructure and illicit sewer connections to be identified and remediated (Dickerson *et al.* 2007). When conducted in a fully quantitative framework, MST also allows for an improved understanding of the timing and conditions of contaminant loading if flow and discharge also are quantified appropriately (Stumpf *et al*. 2010, Gentry-Shields *et al.* 2012).

One potential application of MST markers is to help to correctly identify situations where high FIB concentrations are due human or other animal fecal contamination versus being due to persistence and regrowth. *Enterococcus* spp. and *E. coli* can live, grow, and persist in a variety of benthic and aquatic environments such as sand, sediment, beach wrack, and storm drains (Yamahara *et al*. 2009, Verhougstraete *et al*. 2010, Imamura *et al*. 2011, Byappanahalli *et al*. 2012). To compound matters further, storm events that contribute stormwater runoff to receiving water bodies typically are accompanied by wind, which causes wind-driven resuspension in coastal, lake and estuarine systems. Stormwater runoff can contribute fecal contamination to the receiving waters, but when resuspended, non-fecal sources of FIB found in sediment can be contributed back into the water column. This can be problematic, as some waters may be listed as impaired due to the influence of runoff-driven fecal contamination and/or naturally occurring reservoir populations of *E. coli* and *Enterococcus* spp. Discriminating between the two requires a carefully designed research study with properly selected MST-based markers and regrowth assessment. When implementing MST-based approaches for characterizing fecal contamination in runoff, therefore, it is vital to conduct work in three major areas. First, the molecular markers to be employed must be validated in the geographic area of interest in order to determine their appropriate use. This may include collection of representative sewage effluent samples from sewage infrastructure and pump stations locally, samples from distribution boxes of septic tanks, and appropriate testing of animal-based molecular markers such as those for dogs and gulls in known scat samples. Second, it is necessary to develop a sampling and study plan that includes an appropriate schema for sample analyses. For example, this should include samples collected over a range of wet and dry weather conditions, and over the seasons of interest. Third, the sampling framework should include a total number of samples collected that permits rigorous statistical analyses of the results, preferably including power analysis to determine the minimum number of samples to be collected in any given watershed. Once these three basic steps for design of MST-based study have been followed, the researcher can then assist stakeholders in designing regrowth studies appropriate for natural source exclusion (NSE). A basic tenet of NSE is that statistically rigorous assessment of human fecal contamination or other animal sources has been conducted. If a manager or agency can credibly demonstrate that the populations of *E. coli* and *Enterococcus* spp. within a system are not due to human fecal contamination (or animal sources with comparable risk), then the effective standard used for *E. coli* or *Enterococcus* spp. may be able to be increased appropriately (California State Water Resources Control Board 2008).

Potential for a Tiered Monitoring Strategy

The SIPP study demonstrates that measurements of MST markers do not necessarily correlate with measures of FIB, but that the information gained from measurement of MST markers can be useful for identifying and mitigating contamination (Boehm *et al.* 2013, Ervin *et al*. 2013). Therefore, a tiered

monitoring strategy that incorporates all three of these measures may provide the best strategy to manage surface waters (Boehm *et al.* 2009). Although all originate from excreta, FIB, MST markers, and pathogens provide managers with different types of information and are often measured using a combination of culture-based and molecular approaches (Figure 2). Traditional FIB do not directly measure fecal sources or health risk. MST markers can directly identify sources, and pathogens can be used to estimate health risks. Given the lack of correlation between MST markers and FIB, tiered sampling should include analyses of both as a first tier assessment (Sauer *et al.* 2011, Gentry-*Shields et al*. 2012). This assessment can help prioritize impaired waters, where direct pathogen monitoring and additional resources can be directed.

It is likely that tiered monitoring may not be necessary in the future. Multiplexed detection systems (e.g., Baums *et al*. 2007) will likely be developed and improved to allow simultaneous detection of targets in a single assay. Community sequencing approaches also could enable simultaneous detection. Other important technologies likely to change the future of MST include rapid genetic sensors compatible with field deployable instruments that autonomously collect and analyze samples *in situ*, such as the Environmental Sample Processor (Preston *et al.* 2009). These systems currently have poor detection limits owing to technological challenges associated with concentrating and purifying genetic material from large volumes of water. However, in the future such systems could return information regarding FIB concentration and molecular markers indicating sources of fecal contamination in near real-time. These systems could also be used to sample frequently, increasing the chance of detecting episodic contamination events.

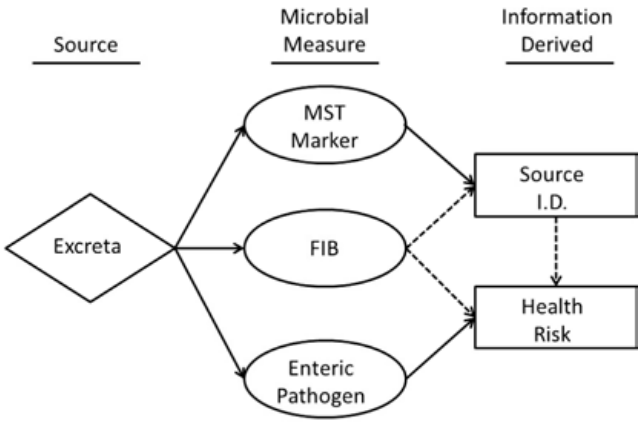

Figure 2. Microbial measures associated with fecal contamination of waters can either directly (solid line) or indirectly (dashed lines) provide information about sources and health risks. MST marker = microbial source tracking marker; FIB = fecal indicator bacteria.

Attainability of Source Allocation

The ability to test an ambient water sample and allocate the proportion of each contributing pollution source would revolutionize water quality management. During the solicitation of SIPP laboratory participants, each group was encouraged to report the proportion of fecal material present on each reference challenge filter. Ultimately, only one laboratory participant submitted source allocation data (Wang *et al*. 2013). Their work reports successful fecal allocation of SIPP reference challenge filters prepared with fresh fecal samples in artificial water, but demonstrates the additional challenges of applying this approach to environmental samples impacted by multiple sources with different fate and transport histories.

One issue related to source apportionment is that in past years library-dependent methods such as antibiotic resistance analysis (ARA) were used to predict the percent contribution of contamination from various pollution sources. This approach established a premature expectation for MST in water quality management. Until quantitative MST becomes more reliable, implementation of MST can proceed using a framework that does not rely on MST-based FIB source allocation. A framework that uses relative subsets of MST markers may be possible, particularly if host-specific proportions of a genus or group could be measured. For example, it may be possible to approach source allocation by comparing subsets of the family Bacteroidales to general *Bacteroides spp*. concentrations (Wang *et al.* 2010). However more research is needed to validate such an approach before it could be adopted for source apportionment in a management context.

Using MST to Model Water Impairment and Health Risks

Quantitative microbial risk assessment (QMRA) models are water quality tools to predict health risk from fecal contamination and could also be helpful in developing water quality criteria on a site-specific basis (USEPA 2012). Reliable MST assays would aid QMRA development (e.g., Staley *et al*. 2012). Although a number of pathogens are known to be

associated with animal wastes, human-source waste is generally considered to pose a greater risk to human health. Identification of a low amount of human waste may prompt a different management response than the identification of a low amount of animal wastes. In a QMRA framework, different classifications and criteria could be assigned to waters based on the source of contaminants. Without MST, risk assessment would be dependent on direct pathogen detection in the environment, which could be costly and difficult because pathogens are rare compared to indicators and MST markers (Straub and Chandler 2003, Savichtcheva and Okabe 2006).

Additional tools including use of GIS mapping, hydrology modeling and conditional probability statistics could be integrated with measures of FIB, MST markers and pathogens to better model water impairment and health risks. GIS tools can provide information about the density and distribution of humans and animals, or microbial measures can be compared to the location of permitted discharges and sewage infrastructure (e.g., Stewart-Pullaro *et al.* 2006). Statistical analysis can be performed to calculate conditional probabilities of correctly identifying sources (e.g., Kildare *et al*. 2007, Jenkins *et al*. 2009), and assays may be able to be combined to help discriminate between weak positive reactions that are attributable to cross-reactivity and those resulting from dilute targets. There is also a possibility for predictive modeling. For example, Gonzalez *et al.* (2012) applied empirical predictive modeling to predict measures of FIB and MST markers in recreational and shellfish harvesting waters. They found strong relationships between FIB and 5 day rainfall totals, dissolved oxygen and salinity. However, the strongest relationship between Bacteroidales markers quantified by qPCR was with antecedent dry period. These differences in predictive capability can be used in a system to better understand the mechanisms of delivery of fecal contamination to waterbodies and persistence of FIB versus molecular markers of fecal contamination in natural, complex estuarine systems. Integration of these approaches with additional MST-derived data over appropriate time and spatial scales will not only change the field of MST in the future, but will advance the science of watershed management to more effectively protect public health.

Summary

The SIPP study highlighted progress and unresolved issues within the field of MST. A path forward includes the following:

- Performance parameters for MST protocols, including methods used to define an assay limit of detection or quantification, and procedures for scoring samples that test positive below these limits, need to be included in all reports of MST results to facilitate method comparisons and to allow consistent and reliable application of methods across laboratories.
- Future research to advance the field of MST should include studies that broaden the geographic range of reference pollution sources, increase the number of animal sources tested, and determine decay rates of host-associated genetic markers from top performing methods. Performance of top methods also should be assessed in field studies to address issues of sample matrix effects.
- The most sensitive and specific methods from the SIPP study can be applied to identify fecal pollution sources and prioritize watersheds in violation of water quality standards. In these cases, MST data also can help direct limited resources toward the most effective mitigation strategies. MST can also be used to help model risks to public health and conditions for water impairment, opening the possibility that appropriate water quality standards can be implemented on a site-specific basis.

Literature Cited

Association of Analytical Communities International (AOAC). 2006. Final Report and Executive Summaries from the AOAC International Presidential Task Force on Best Practices in Microbiological Methodology. Report #223-01-2464. AOAC. Gaithersburg, MD.

Armbruster, D.A. and T. Pry. 2008. Limit of blank, limit of detection and limit of quantitation. *Clinical Biochemist Reviews* 29 Supplement 1:S49-52.

Bae, S. and S. Wuertz. 2012. Survival of host-associated *Bacteroidales* cells and their relationship with *Enterococcus* spp., *Campylobacter jejuni*, *Salmonella enterica serovar Typhimurium*, and

adenovirus in freshwater microcosms as measured by propidium monoazide-quantitative PCR. *Applied and Environmental Microbiology* 78:922-932.

Baums, I.B., K.D. Goodwin, T. Kiesling, D. Wanless, M.R. Diaz and J.W. Fell. 2007. Luminex detection of fecal indicators in river samples, marine recreational water, and beach sand. *Marine Pollution Bulletin* 54:521-536.

Boehm, A.B., N.J. Ashbolt, J.M. Colford, Jr., L.E. Dunbar, L.E. Fleming, M.A. Gold, J.A. Hansel, P.R. Hunter, A.M. Ichida, C,D, McGee, J.A. Soller and S.B. Weisberg. 2009. A sea change ahead for recreational water quality criteria. *Journal of Water and Health* 7:9-20.

Boehm, A.B., L. Van De Werfhorst, J.F. Griffith, P. Holden, J. Jay, O.C. Shanks, D. Wang and S.B. Weisberg. 2013. Performance of forty-one microbial source tracking methods: a twenty-seven lab evaluation study. *Water Research* 47:6812-6828.

Burd, E.M. 2010. Validation of laboratory-developed molecular assays for infectious diseases. *Clinical Microbiology Reviews* 23:550-576.

Burns, M. and H. Valdivia. 2008. Modelling the limit of detection in real-time quantitative PCR. *European Food Research and Technology* 226:1513-1524.

Bustin, S.A., V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M.W. Pfaffl, G.L. Shipley, J. Vandesompele and C.T. Wittwer. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry* 55:611-622.

Byappanahalli, M.N., T. Yan, M.J. Hamilton, S. Ishii, R.S. Fujioka, R.L. Whitman and M.J. Sadowsky. 2012. The population structure of *Escherichia coli* isolated from subtropical and temperate soils. *Science of the Total Environment* 417-418:273-279.

Cao, Y., L.C. Van der Werfhorst, E.A. Dubinsky, B.D. Badgley, M.J. Sadowsky, G.L. Andersen, J.F. Griffith and P.A. Holden. 2013. Evaluation of molecular community analysis methods for discerning fecal sources and human waste. *Water Research* 47:6862-6872.

California State Water Resources Control Board. 2008. A resolution amending the water quality control plan for the San Diego Basin (9) to incorporate implementation provisions for indicator bacteria water quality objectives to account for loading from natural uncontrollable sources within the context of a total maximum daily load. R9-2008-0028. Available on-line: http://waterboards.ca.gov/sandiego/board_decisions/adopted_orders/2008/R9-2008-0028.pdf

CODEX. 2010. Guidelines on performance criteria and validation of methods for detection, identification and quantification of specific DNA sequences and specific proteins in foods. Codex Committee on Methods of Analysis and Sampling CAC/GL 74-2010. Budapest, Hungary.

Dickerson Jr., J.W., C. Hagedorn and A. Hassall. 2007. Detection and remediation of human-origin pollution at two public beaches in Virginia using multiple source tracking methods. *Water Research* 41:3758-3770.

Ebentier, D.L., K.T. Hanley, Y. Cao, B. Badgley, A.B. Boehm, J. Ervin, K.D. Goodwin, M. Gourmelon, J.F. Griffith, P. Holden, C.A. Kelty, S. Lozach, C. McGee, L. Peed, M. Raith, M.J. Sadowsky, E. Scott, J. Santodoming, C.D. Sinagalliano, O.C. Shanks, L. Van De Werfhorst, D. Wang, S. Wuertz and J. Jay. 2013. Evaluation of the repeatibility and reproducibility of a suite of PCR-based microbial source tracking methods. *Water Research* 47:6839-6848.

Ervin, J.S., T.L. Russell, B.A. Layton, K.M. Yamahara, D. Wang, L.M. Sassoubre, Y. Cao, C.A. Kelty, M. Sivaganesan, A.B. Boehm, P. Holden, S.B. Weisberg and O.C. Shanks. 2013. Characterization of fecal concentrations in human and other animal sources by physical, cutlure, and quantitative real-time PCR methods. *Water Research* 47:6873-6882.

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

Gentry-Shields, J., J.G. Rowny and J.R. Stewart. 2012. HuBac and *nifH* source tracking markers display a relationship to land use but not rainfall. *Water Research* 46:6163-6174.

Gonzalez, R.A., K.E. Conn, J.R. Crosswell and R.T. Noble. 2012. Application of empirical predictive modeling using conventional and alternative fecal indicator bacteria in eastern North Carolina waters. *Water Research* 46:5871-5882.

Green, H., O.C. Shanks, M. Sivaganesan, R. Haugland and K. Field. 2011. Differential decay of human faecal *Bacteroides* in marine and freshwater. *Environmental Microbiology* 13:3235-3249.

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

Harrington, P.R., W. Zeng and L.K. Naeger. 2012. Clinical relevance of detectable but not quantifiable hepatitis C virus RNA during boceprevir or telaprevir treatment. *Hepatology* 55:1048-1057.

Harwood, V.J., A.B. Boehm, L.M. Sassoubre, V. Kannappan, J.R. Stewart, T.-T. Fong, D. Diston, J. Gentry-Shields, J.F. Griffith, J.A. Furhman, M. Gourmelon, R.T. Noble and M. Wicki. 2013. Performance of viruses and bacteriophages for fecal source determination in a multi-laboratory, comparative study (SIPP). *Water Research* 47:6929-6943.

Harwood, V., K. Gordon and C. Staley. 2011. Validation of Rapid Methods for Enumeration of Markers for Human Sewage Contamination in Recreational Waters. WERF Report PATH3C09. Water Environment Research Foundation. Alexandria, VA.

Haznedaroglu, B.Z., H.N. Kim, S.A. Bradford and S.L. Walker. 2009. Relative transport behavior of *Escherichia coli* O157:H7 and *Salmonella enterica serovar pullorum* in packed bed column systems: influence of solution chemistry and cell concentration. *Environmental Science & Technology* 43:1838-1844.

Huggett, J., T. Novak, J.A. Garson, C. Green, S.D. Morris-Jones, R.F. Miller and A. Zumla. 2008. Differential susceptibility of PCR reactions to inhibitors: an important and unrecognized phenomenon. *BMC Research Notes* 1:70-79.

Imamura, G.J., R.S. Thompson, A.B. Boehm and J.A. Jay. 2011. Wrack promotes the persistence of fecal indicator bacteria in marine sands and seawater. *FEMS Microbiology Ecology* 77:40-49.

Jenkins, M.W., S. Tiwari, M. Lorente, C.M. Gichaba and S. Wuertz. 2009. Identifying human and livestock sources of fecal contamination in Kenya with host-specific *Bacteroidales* assays. *Water Research* 43:4956-4966.

Kildare, B.J., C.M. Leutenegger, B.S. McSwain, D.G. Bambic, V.B. Rajal and S. Wuertz. 2007. 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal *Bacteroidales*: A Bayesian approach. *Water Research* 41:3701-3715.

Laperche, S., F. Bouchardeau, E. Andre-Garnier, V. Thibault, A.M. Roque-Afonso, P. Trimoulet, R. Colimon, G. Duverlie, H. Leguillou-Guillemette, F. Lunel, M. Bouvier-Alias, J.M. Pawlotsky, C. Henquell, E. Schvoerer, F. Stoll-Keller, M.L. Chaix, M. Branger, C. Gaudy-Graffin, A.R. Rosenberg, B. Pozzetto, S. Vallet, Y. Baazia, J. Izopet and J.J. Lefrere. 2011. Interpretation of real-time PCR results for hepatitis C virus RNA when viral load is below quantification limits. *Journal of Clinical Microbiology* 49:1113-1115.

Layton, B., Y. Cao, D.L. Ebentier, K.T. Hanley, L. Van De Werfhorst, D. Wang, T. Madi, R.L. Whitman, M.N. Byappanahalli, E. Balleste, W. Meijier, A. Schriewer, S. Wuertz, R.R. Converse, R.T. Noble, S. Srinivasan, J.B. Rose, C.S. Lee, J. Lee, J. Shields, J.R. Stewart, G. Reischer, A. Farnleitner, C.D. Sinagalliano, R. Rodrigues, S. Lozach, M. Gourmelon, L. Peed, O.C. Shanks, J. Jay, P. Holden, A.B. Boehm and J.F. Griffith. 2013. Performance of human fecal-associated PCR-based assays: an international source identification method evaluation. *Water Research* 47:6897-6908.

Layton, A., L. McKay, D. Williams, V. Garrett, R. Gentry and G. Sayler. 2006. Development of *Bacteroides* 16S rRNA gene TaqMan-based real-time PCR assays for estimation of total, human, and bovine fecal Pollution in water. *Applied and Environmental Microbiology* 72:4214-4224.

Manter, D.K., T.L. Wier and J.M. Vivanco. 2010. Negative effects of sample pooling on PCR-based estimates of soil microbial richness and community structure. *Applied and Environmental Microbiology* 76:2086-2090.

McQuaig, S.M., T.M. Scott, V.J. Harwood, S.R. Farrah and J. Lukasik. 2006. Detection of human-derived fecal pollution in environmental waters by use of a PCR-based human polyomavirus assay. *Applied and Environmental Microbiology* 72:7567-7574.

Mumy, K.L. and R.H. Findaly. 2004. Convenient determination of DNA extraction efficiency using an

external DNA recovery standard and quantitative-competitive PCR. *Applied and Environmental Microbiology* 57:259-268.

Okabe, S., N. Okayama, O. Savichtcheva and T. Ito. 2007. Quantification of host-specific *Bacteroides-Prevotella* 16S rRNA genetic markers for assessment of fecal pollution in freshwater. *Applied Microbiology and Biotechnology* 74:890-901.

Preston, C.M., R. Marin, 3rd, S.D. Jensen, J. Feldman, J.M. Birch, E.I. Massion, E.F. Delong, M. Suzuki, K. Wheeler and C.A. Scholin. 2009. Near real-time, autonomous detection of marine bacterioplankton on a coastal mooring in Monterey Bay, California, using rRNA-targeted DNA probes. *Environmental Microbiology* 11:1168-1180.

Raith, M.R., C.A. Kelty, J.F. Griffith, A. Schriewer, S. Wuertz, S. Mieszkin, M. Gourmelon, G.H. Reischer, A.H. Farnleitner, J.S. Ervin, P.A. Holden, D.L. Ebentier, J.A. Jay, D. Wang, A.B. Boehm, T. Gim Aw, J.B. Rose, E. Balleste, W.G. Meijer, M. Sivaganesan and O.C. Shanks. 2013. Comparison of PCR and quantitative real-time PCR methods for the characterization of ruminant and cattle fecal pollution sources. *Water Research* 47:6921-6928.

Rogers, S.W., M. Donnelly, L. Peed, C.A. Kelty, S. Mondal, Z. Zhong and O.C. Shanks. 2011. Decay of bacterial pathogens, fecal indicators, and real-time quantitative PCR genetic markers in manure-ammended soils. *Applied and Environmental Microbiology* 77:4839-4848.

Sauer E.P., J.L. Vandewalle, M.J. Bootsma and S.L. McLellan. 2011. Detection of the human specific Bacteroides genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment. *Water Research* 45:4081-4091.

Savichtcheva, O. and S. Okabe. 2006. Alternative indicators of fecal pollution: relations with pathogens and conventional indicators, current methodologies for direct pathogen monitoring and future application perspectives. *Water Research* 40:2463-2476.

Schinner, T., A. Letzner, S. Liedtke, F.D. Castro, I.A. Eydelnant. and N. Tufenkji. 2010. Transport of selected bacterial pathogens in agricultural soil and quartz sand. *Water Research* 44:1182-1192.

Schriewer, A., C.D. Sinigalliano, K.D. Goodwin, A. Cox, D. Wanless, J. Bartkowiak, D.L. Ebentier, K.T. Hanley, J. Ervin, L.A. Deering, O.C. Shanks, L.A. Peed, W.G. Mijer, J.F. Griffith, J. Santo Domingo, J.A. Jay, P.A. Holden and S. Suertz. 2013. Performance evaluation of canine-associated *Bacteroidales* assays in a multi-laboratory comparison study. *Water Research* 47:6909-6920.

Schulz, C.J. and G.W. Childers. 2011. Fecal *Bacteroidales* diversity and decay in response to variations in temperature and salinity. *Applied and Environmental Microbiology* 77:2563-2571.

Shanks, O.C., C.A. Kelty, S.L. 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.

Shanks, O.C., R.J. Newton, C.A. Kelty, S.M. Huse, M.L. Sogin and S.L. McLellan. 2013. Comparison of microbial community structure in untreated wastewater from different geographic locales. *Applied and Environmental Microbiology* 79:2906-2913.

Shanks, O.C., K. White, C.A. Kelty, S. Hayes, M. Sivaganesan, M. Jenkins, M. Varma and R.A. Haugland. 2010a. Performance assessment of cattle-associated PCR and quantitative real-time PCR assays targeting *Bacteroidales* Genes. *Applied and Environmental Microbiology* 76:1359-1366.

Shanks, O.C., K. White, C.A. Kelty, M. Sivaganesan, J. Blannon, M. Meckes, M. Varma and R.A. Haugland. 2010b. Performance of PCR-based assays targeting *Bacteroidales* genetic markers of human fecal pollution in sewage and fecal samples. *Environmental Science and Technology* 44:6281-6288.

Silkie, S.S. and K.L. Nelson. 2009. Concentrations of host-specific and generic fecal markers measured by quantitative PCR in raw sewage and fresh animal feces. *Water Research* 43:4860-4871.

Sinigalliano, C., J. Ervin, L. Van De Werfhorst, B. Badgley, E. Ballesté, J. Bartkowiak, A. Boehm, M. Byappanahalli, K. Goodwin, M. Gourmelon, J. Griffith, P. Holden, J. Jay, B. Layton, C. Lee, J. Lee, W. Meijer, RT, Nobel., M. Raith, H. Ryu, M. Sadowsky, A. Schriewer, D. Wang, D. Wanless, R. Whitman, S. Wuertz, J. Santo Domingo. 2013.

Multi-Laboratory Evaluations of the Performance of *Catellicoccus marimammalium* PCR Assays Developed to Target Gull Fecal Sources. *Water Research* 47:6883-6896.

Staley, C., K.V. Gordon, M.E. Schoen and V.J. Harwood. 2012. Performance of two quantitative PCR methods for microbial source tracking of human sewage and implications for microbial risk assessment in recreational waters. *Applied and Environmental Microbiology* 78:7317-7326.

Stewart-Pullaro, J., J.W. Daugomah, D.E. Chestnut, D.A. Graves, M.D. Sobsey and G.I. Scott. 2006. F+ RNA coliphage typing for microbial source tracking in surface waters. *Journal of Applied Microbiology* 101:1015-1026.

Stoeckel, D.M. and V.J. Harwood. 2007. Performance, design, and analysis in microbial source tracking studies. *Applied and Environmental Microbiology* 73:2405-2415.

Stoeckel, D.M., E.A. Stelzer and L.K. Dick. 2009. Evaluation of two spike-and -recovery controls for assessment of extraction efficiency in microbial source tracking studies. *Water Research* 43:4820-4827.

Straub, T.M. and D.P. Chandler. 2003. Towards a unified system for detecting waterborne pathogens. *Journal of Microbiological Methods* 53:185-197.

Stumpf, C.H., M.F. Piehler, S. Thompson and R.T. Noble. 2010. Loading of fecal indicator bacteria in North Carolina tidal creek headwaters: hydrographic patterns and terrestrial runoff relationships. *Water Research* 44:4704-4715.

United States Environmental Protection Agency (USEPA). 2012. 2012 Recreational Water Quality Criteria, EPA Office of Water, Washington, DC. Criteria Document 820-F-12-058. Available on-line: http://water.epa.gov/scitech/swguidance/standards/criteria/health/recreation/index.cfm

Verhougstraete, M.P., M.N. Byappanahalli, J.B. Rose and R.L. Whitman. 2010. Cladophora in the Great Lakes: Impacts on beach water quality and human health. *Water Science and Technology* 62:68-76.

Walters, S.P., K.M. Yamahara and A.B. Boehm. 2009. Persistence of nucleic acid markers of health-relevant organisms in seawater microcosms: Implications for their use in assessing risk in recreational waters. *Water Research* 43:4929-4939.

Wang, D., A. Farnleitner, K. Field, H.C. Green, O.C. Shanks and A.B. Boehm. 2013. *Enterococcus* and *Escherichia coli* fecal source apportionment with microbial source tracking genetic markers: is it feasible? *Water Research* 47:6849-6861.

Wang, D., S.S Silkie, K.L. Nelson and S. Wuertz. 2010. Estimating true human and animal host source contribution in quantitative microbial source tracking using the Monte Carlo method. *Water Research* 44:4760-4775.

Wilson, I.G. 1997. Inhibition and facilitation of nucleic acid amplification. *Applied and Environmental Microbiology* 63:3741-3751.

Yamahara, K.M., S.P. Walters and A.B. Boehm. 2009. Growth of enterococci in unaltered, unseeded beach sands subjected to tidal wetting. *Applied and Environmental Microbiology* 75:1517-1524.

Acknowledgements

Funding for this project has been provided in part through an agreement with the California State Water Resources Control Board. The contents of this document do not necessarily reflect the views and policies of the California State Water Resources Control Board, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.