Southern California Bight 2013 Regional Monitoring Program

BIGHT '13

Shoreline Microbiology

Southern California Bight 2013 Regional Monitoring Program Volume IX

SCCWRP Technical Report 1005

Southern California Bight 2013 Regional Monitoring Program: Volume IX. Shoreline Microbiology

Yiping Cao[1], Meredith R. Raith[1], Paul D. Smith[1], John F. Griffith[1], Stephen B. Weisberg[1], Alexander Schriewer[2], Andrew Sheldon[3], Chris Crompton[4], Geremew G. Amenu[5], Jason Gregory[6], Joe Guzman[7], Kelly D. Goodwin[8], Laila Othman[9], Mayela Manasjan[10], Samuel Choi[11], Shana Rapoport[12], Syreeta Steele[13], Tommy Nguyen[14], and Xueyuan Yu[15]

[1]Southern California Coastal Water Research Project
[2]Weston Solutions
[3]City of Malibu
[4]Orange County Department of Public Works
[5]Los Angeles County Department of Public Works
[6]Los Angeles County Sanitation District
[7]Orange County Public Health Laboratory
[8]National Oceanic and Atmosphere Administration
[9]City of San Diego Public Utilities Department
[10]City of Encinitas
[11]Orange County Sanitation District
[12]Los Angeles Regional Water Quality Control Board
[13]Ventura County Public Health Laboratory
[14]City of Los Angeles Environmental Monitoring Division
[15]San Diego Regional Water Quality Control Board

September 2017
SCCWRP Technical Report 1005

SHORELINE MICROBIOLOGY COMMITTEE MEMBERS

Alexander Schriewer	Weston Solutions
Geremew Amenu	Los Angeles County Flood Control District
Andrew Sheldon	City of Malibu
Blythe Layton	Southern California Coastal Water Research Project
Bram Sercu	Ventura County Watershed Protection District
Chris Crompton	Orange County Department of Public Works
Helen Yu	San Diego Regional Water Quality Control Board
Jason Gregory	Los Angeles County Sanitation District
Jian Peng	Orange County Department of Public Works
JoAnn Weber	San Diego County Public Works
Joe Guzman	Orange County Public Health Laboratory
John Griffith	Southern California Coastal Water Research Project
Kathy Walker	Los Angeles County Sanitation District
Keith Kezer	County of San Diego
Kelly Goodwin	National Oceanic and Atmosphere Administration
Kelly Hahs	Ventura County Watershed Protection District
Ken Schiff	Southern California Coastal Water Research Project
Laila Othman	City of San Diego Public Utilities Department
Lara Meeker	Ventura County Watershed Protection District
Luke Christian	City of Oceanside
Mayela Manasjan	City of Encinitas
Melody McNay	Weston Solutions
Meredith Raith	Southern California Coastal Water Research Project
Michael Soto	Ventura County Public Health Laboratory
Michele Padilla	Los Angeles County Sanitation District
Mike von Winckelmann	Orange County Sanitation District
Sam Choi	Orange County Sanitation District
Shana Rapoport	Los Angeles Regional Water Quality Control Board
Sonji Romero	City of San Diego Public Utilities Department
Stephen Weisberg	Southern California Coastal Water Research Project
Syreeta Steele	Ventura County Public Health Laboratory
Tania Chiem	Orange County Public Health Laboratory
Tommy Nguyen	City of Los Angeles Environmental Monitoring Division
Yiping Cao	Southern California Coastal Water Research Project

FOREWORD

The Southern California Bight 2013 Regional Monitoring Program (Bight'13) is an integrated, collaborative effort to provide large-scale assessments of the Southern California Bight (SCB). The Bight'13 survey is an extension of previous regional assessments conducted every five years dating back to 1994. The collaboration represents the combined efforts of nearly 100 organizations. Bight'13 is organized into five elements: 1) Contaminant Impact Assessment (formerly Coastal Ecology), 2) Shoreline Microbiology, 3) Nutrients, 4) Marine Protected Areas, and 5) Trash and Debris. This report presents the results of the shoreline microbiology portion of the Bight'13 monitoring program.

ACKNOWLEDGEMENTS

This report is a result of the dedication and hard work of many individuals, including members of the Bight'13 Shoreline Microbiology Committee, as well as all support staff, students, and interns from all the participating agencies. While we have not named everyone who contributed to this report, the authors wish to thank all these individuals for their assistance with study design, sampling, sample analysis, data review, and report review. This report would not be possible without the contribution from everyone.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Quantitative polymerase chain reaction (qPCR) methods have demonstrated great potential for changing how recreational water quality is monitored, enhancing the speed to result, specificity, and sensitivity compared to traditional culture-based methods. There are two main applications for qPCR in recreational water quality monitoring. The first is a more rapid measurement of fecal indicator bacteria (FIB) to facilitate timely public notification when beach water is contaminated or a contamination event has passed. For example, U.S. EPA's 2012 revised recreational water quality criteria allows routine use of qPCR to measure *Enterococcus* spp. for rapid beach monitoring. The second application is to identify sources of fecal pollution using host-associated source markers. In this case, the *California Microbial Source Identification Manual* recommends a suite of validated qPCR-based methods to measure fecal source identification markers for identifying sources of fecal pollution in ambient waters.

The ability to transition these applications to practical management use relies on three aspects. The first is feasibility of technology transfer. Despite continued advancements in the development of qPCR-based methodologies, these tools are still primarily employed by expert-users in research labs. Furthermore, evaluative studies of the efficacy of transferring qPCR technology to public agencies responsible for routine monitoring have been limited to relatively small demonstration projects of short duration. Regional scale assessment of these methods across a variety of laboratories with varying levels of familiarity with qPCR is necessary to assess whether this technology may be realistically transitioned to end-users. Regulatory issues are the second important aspect affecting transition success. Before granting approval of qPCR quantification of enterococci as a replacement for culture-based monitoring of FIB, the U.S. EPA encourages a site-specific analysis of qPCR's performance. An implementation threshold (e.g. for beach posting) also needs to be established based on demonstrated consistent and predictable relationship with an EPA-approved benchmark method such as culture-based method. The third aspect involved in transition success is technical. While the qPCR-measured host-associated markers have been used extensively as a diagnostic tool to determine sources of fecal contamination within watersheds, there is no framework for interpreting marker results from routine monitoring to prioritize remediation actions amongst multiple watersheds.

The Microbiology Committee of the Southern California Bight '13 Regional Monitoring Program focused its efforts on addressing questions surrounding the efficacy of employing qPCR methods in southern California with three study goals:

1. Assess the ability of laboratories conducting routine water quality monitoring in the southern California region to employ qPCR methods.
2. Assess comparability of results between qPCR and traditional culture for measuring *Enterococcus* spp. at southern California beaches.
3. Determine the management value of regional monitoring for a human fecal source marker HF183 to prioritize sites for remediation.

To address the first goal, we evaluated the ability of eight water quality monitoring labs to produce comparable results using qPCR. The evaluation consisted of three phases. The first was a training workshop in which local agencies learned how to perform the qPCR methods. The second was an intercalibration study conducted two months post-training designed to assess

whether the proficiency demonstrated by participants at the completion of training persisted after they returned to their home laboratories. The third phase was a long-term assessment of laboratory performance over three subsequent years. Technology transfer of qPCR to local monitoring laboratories was found to be feasible for both the enterococci and the HF183 qPCR methods, with all participating labs demonstrating proficiency in performing the methods during the intercalibration phase. However, some laboratories did not sustain their proficiency over the 3-year implementation, as indicated by a lack of repeatability and inability to consistently produce acceptable standard curves. The higher failure rate in the long-term evaluation was laboratory specific and appeared attributable to turnover of laboratory personnel. Our results illustrate the need for formal lab training and accreditation programs for qPCR methods that includes standardized training, evaluation, and certification of laboratories performing qPCR methods for water quality monitoring.

To address the second goal, we evaluated the level of agreement between enterococci measured in parallel by both culture and qPCR methods at 36 southern California beaches following a formal procedure outlined in an EPA technical supporting material (TSM) document. Approximately 50 samples were collected at each beach in dry weather during both summer and winter seasons. Only one site in summer and three sites in winter achieved the level of agreement required by EPA to allow replacement of the culture-based method with qPCR for posting the beach. Matrix interference, in which constituents in environmental waters interfere with the qPCR chemistry leading to underestimation, was not an obstacle for qPCR equivalency, nor was laboratory performance an impediment. The lack of demonstrable agreement was largely due to enterococci concentrations at most sites being low; even with two years of weekly sample collection, we were not able to achieve the requisite number of samples in the quantifiable range for most beaches. In addition, the useable measurements were mostly at the low end of the quantification range for both methods. This produced a limited range of values, contributing to high method variability and thus, low correlation. The few sites that did meet the EPA requirements were ones that had concentrations that were frequently above the beach warning thresholds. This suggests that some modification of the EPA protocols is necessary if qPCR is to be adopted at sites that are typically clean and subject only to transient pollution events, which are the sites where the speed advantage of qPCR is most valuable for improved public health protection.

To address the third goal, we conducted a regional assessment of human marker prevalence among drainages that discharge to the southern California ocean. Approximately 50 samples from each of 22 southern California coastal drainages were collected under summer dry weather conditions. An additional 50 samples were targeted from each of 23 drainages during wet weather, although a drought during this period prevented us from achieving the targeted number of samples at all wet weather sites. Samples were analyzed for the HF183 human fecal marker, which was found to be ubiquitous across the region; it was present at all but two sites in dry weather and at all sites during wet weather. There was considerable difference in the extent of human fecal contamination among sites. While site rankings remained consistent regardless of whether ranking was based on frequency of HF183 detection or on average concentration of HF183, site ranking differed greatly between dry and wet weather. Site ranking also differed greatly between if based on HF183 and if based on enterococci, which, unlike HF183, does not

distinguish between human and non-human fecal contamination. Although, additional work is needed to interpret results regarding human health impacts, these results illustrate the management value of HF183 as a monitoring tool to enable more effective prioritization of sites for remediation.

CHAPTER 1. TECHNOLOGY TRANSFER

Introduction

Quantitative polymerase chain reaction (qPCR) methods offer substantial advantages over culture-based methods for beach water quality monitoring. Most important, they yield results in as little as a few hours, compared to up to 3 days for some culture-based methods (Ferretti et al. 2011). This rapid time to result has the potential to enhance public health protection by enabling beach managers to notify beachgoers of unhealthful conditions prior to exposure to contaminated waters (Ferretti et al. 2013a, Griffith and Weisberg 2011). qPCR methods are also highly sensitive and specific and capable of measuring a wide range of targets that cannot be measured by culture-based methods (Boehm et al. 2013, Harwood et al. 2013). As such, qPCR technology has become a widely-used research tool in the microbial water quality community.

Most applications of qPCR for beach water testing have been conducted by researchers involved in method development (Ahmed et al. 2016, Converse et al. 2009, Noble et al. 2010). Some initial work has been done that demonstrates successful technology transfer to management agencies responsible for routine monitoring (Ferretti et al. 2013b, Griffith and Weisberg 2011, Lavender and Kinzelman 2009). Yet, these studies primarily involved a limited number of first-adopter laboratories implementing the method over a short time frame, usually a single summer.

Successful adoption of qPCR for routine use will require that all agencies operating within a region can successfully perform these methods over multiyear time frames, across a range of operators within each laboratory, and that results produced by one laboratory are comparable to those produced by other laboratories operating within the region. This study evaluated the feasibility of transferring qPCR technology to eight water quality monitoring laboratories within the southern California region over three years, using two representative qPCR methods: The *Enterococcus* qPCR method for measuring general fecal contamination, and the HF183 qPCR method for measuring human fecal contamination.

Methods

Approach

A three-phase training and evaluation approach was used. The first phase was a training workshop in which 14 local agencies learned how to perform the two qPCR methods. The second phase was an evaluation study conducted two months after the training to assess whether the proficiency participants demonstrated during the training phase persisted after they returned to their home laboratories. The third phase was an assessment of long-term laboratory performance, in which results from negative controls and standard curves produced by each of eight participating laboratories were assessed over three subsequent years as they conducted routine qPCR analyses.

qPCR methods

Two representative qPCR methods with different levels of operational complexity were used: 1) the general fecal indicator method TaqEnviron targeting *Enterococcus* spp. (Cao et al. 2012) and 2) the human fecal-associated marker method HF183/BacR287 (Green et al. 2014). The

TaqEnviron method followed EPA1611 (U.S. EPA 2012), except that it used a master mix more robust to inhibition (Cao et al. 2012) and 40 cycles of qPCR, as later documented in Methods EPA1609 (U.S. EPA 2013). The *Enterococcus* method is less complex than the HF183 method because it uses crude DNA extract that eliminated DNA purification steps, and it consisted of simplex qPCR to detect *Enterococcus* spp. Simplex refers to measuring a single target in one qPCR reaction, whereas duplex (described below) refers to measuring two targets in a single qPCR reaction. Reference material for the enterococcus assay included genomic DNA from *E. faecalis* (ATCC 29212) and whole cell calibrators of *E. faecalis*, as described in EPA protocols (U.S. EPA 2012). A sample processing control was included for every sample, which consisted of a simplex Sketa assay (sketa 22) to amplify salmon sperm DNA added during DNA extraction (U.S. EPA 2012). Although the Sketa assay required a separate reaction, qPCR thermal cycler conditions were identical to the *Enterococcus* assay so that both reactions could be run on the same plate.

The HF183/BacR287 method utilized purified DNA. DNA was extracted from samples using the GeneRite DNA EZ kit (GeneRite, North Brunswick, NJ) as described previously (Green et al. 2014). The assay consisted of a duplex assay that targeted the HF183 marker and an internal amplification control (Green et al. 2014). As with the *Enterococcus* assay, the simplex Sketa assay was used as sample processing control (U.S. EPA 2012). Reference material for the HF183 assay was plasmid DNA as described elsewhere (Green et al. 2014). The HF183 assay was included in this technology transfer demonstration because DNA purification and the duplex format was more operationally challenging and the data analysis was more complex than the *Enterococcus* assay.

qPCR training workshop

A 3-day training workshop was held at the Southern California Coastal Water Research Project Authority (SCCWRP) from August 27 to 30, 2013. More than two dozen staff from 14 agencies were taught to perform the two qPCR methods described above. The first day consisted of classroom training focused on the basic theories behind qPCR analysis. The next two days focused on hands-on training with the laboratory techniques necessary to perform qPCR. Each participating laboratory received a workshop binder consisting of all training presentations, standard operating procedures for both qPCR methods, and benchtop guides for critical steps. Hands-on training started with instruction on proper pipetting technique, which was provided by Artel, a professional pipetting training and certification company. This was followed by training on serial dilutions, performing water filtration, crude DNA extraction followed by *Enterococcus* qPCR, and finally, DNA purification followed by HF183 qPCR. Once participants had successfully produced qPCR results, they attended a classroom session on how to analyze qPCR data. Following the workshop, trainees returned to their home laboratories and continued practicing for six weeks prior to the laboratory intercalibration exercise.

Post-training laboratory intercalibration study

Eight laboratories analyzed a common set of test material provided by a centralized source (SCCWRP) on five different days of each lab's choice within two months. Three types of test material were included: blanks, standard reference material, and blind unknown samples. Blanks were used to test the ability to perform qPCR without cross contamination (i.e., cleanliness), and

included DNA-free molecular grade water in lieu of sample DNA in the qPCR reaction as no template controls (NTCs) and blank filters (membrane filters processed by filtering 100ml of 0.2µm-filtered phosphate buffered saline in lieu of sample) as negative extraction controls (NECs). Standard reference material was used to test the ability to perform serial dilution and obtain high quality standard curves. Reference material included genomic DNA from *E. faecalis* (ATCC 29212), whole *E. faecalis* cells (provided as frozen calibrator filters with 10^5 cells/filter), and synthetic linearized plasmid DNA containing the target for the HF183 marker. Detailed protocols were provided, including the exact qPCR plate setup and pre-calculated reagent volumes for each day in an Excel spreadsheet.

All standard reference material was prepared as single use aliquots at SCCWRP and stored at -80 ^{o}C before use. Blind unknown samples were used to test the ability to produce expected qPCR quantification. These consisted of duplicate filters spiked with three different concentrations of sewage, which were created by filtering 50ml per filter of 10-, 100- and 1000-fold diluted raw sewage (Orange County Sanitation District primary influent) in sterile PBS.

The intercalibration study took place on five days, chosen individually by each laboratory to minimize interference with routine work. It began with a simple setup of just standard curves and progressed through a setup of a full 96-well plate that included standard curves, NTC, NEC, and blind unknown samples. On Day 1, each laboratory was required to perform serial dilution and qPCR analysis for one *Enterococcus* standard curve using genomic DNA and one HF183 standard curve using the linearized plasmid DNA, each with three NTCs. On Day 2, work included crude DNA extraction from *Enterococcus* calibrator filters and serial dilution of genomic DNA and calibrator crude extracts to generate *Enterococcus* standard curves. On Day 3, work included crude DNA extraction from unknown sewage-spiked filters (a set of 6 filters) and analysis for *Enterococcus*. On Day 4, work included full DNA extraction (i.e., with DNA purification as required by the HF183 method) and analysis for the HF183 marker. On Day 5, labs repeated the same qPCR analyses as on Day 4, except that no DNA extraction was performed and all laboratories used purified DNA provided by SCCWRP (extracted at SCCWRP from replicate filters from the same batch as those used by the labs on Day 4). The purpose of the Day 5 exercise was to isolate proficiency in conducting the full DNA purification protocol, via comparing Day 5 (SCCWRP conducted DNA extraction for all labs) results to those from Day 4 (each lab did its own DNA extraction).

Long-term implementation

Laboratory performance over long-term implementation of qPCR methods was assessed through analyzing NTC, NEC, and standard curve results from each of the eight laboratories as they conducted routine qPCR analyses from July 2013 to January 2017. For *Enterococcus*, routine analysis was done in batches of 16 samples following standardized qPCR plate setup, including one standard curve, one NTC, and one NEC for each plate. For the HF183 method, routine analysis was done in batches of 22 samples following standardized qPCR plate setup, including one standard curve, one NTC, and at least two NECs for each plate. All standard reference material was prepared as single use aliquots at SCCWRP, and stored at SCCWRP or other participating laboratories at -80 ^{o}C until use.

Data analysis

All qPCR data exported from the thermocyclers were uploaded to a web portal created and maintained by SCCWRP and processed to assess proficiency in four areas: cleanliness, standard curve performance, accuracy, and repeatability. Cleanliness referred to the laboratory's ability to perform qPCR without cross contamination, and was assessed using cycle of quantification (Cq) results from negative controls such as NTC and NEC. Standard curve performance referred to the laboratory's ability to perform serial dilution and obtain high quality standard curves, and was assessed by regression. Accuracy referred to the ability to produce expected quantification of the challenge samples. Repeatability referred to the ability to produce repeatable results within a run and between runs, and was assessed by Cq standard deviation and range of mean Cq values among replicates.

The performance criteria utilized to evaluate performance for each of the four areas (Table 1) were established from published data acceptance criteria (Shanks et al. 2016, U.S. EPA 2013) and generally accepted best practices in the water quality testing field (Ebentier et al. 2013, Griffith and Weisberg 2011) or related field (Bustin et al. 2009). Some criteria are more stringent than those (refered as draft criteria hereon) provided to participants during the study, which were based on criteria used for enterococci qPCR previously (Griffith and Weisberg 2011). Standard curve regression equations (Cq vs. $\log_{10}$ concentration) were calculated with an outlier removal procedure in which a data point was deemed an outlier if its studentized residual was greater than three (Ebentier et al. 2013). While individual standard curves were assessed separately for standard curve performance criteria, a pooled master standard curve approach was used for quantification of blind unknown samples to reduce quantification bias from run-to-run variability as described previously (Ebentier et al. 2013, Sivaganesan et al. 2010).

Table 1. Performance evaluation criteria for cleanliness (NTC and NEC), standard curve parameters, accuracy, and availability, for the two qPCR methods. Criteria with notable changes between the time of the study (draft) and current recommendations are listed (see text).

	***Enterococcus* qPCR**	**HF183/BacR287 qPCR**
Cleanliness[a]	No negative control with Cq <35 <1/3 of the negative controls with Cq≥35	No negative control shows any amplification
Standard curve[b]	Draft: E: 0.87 – 1.1; R^2≥0.97 Current[e]: E: 0.85 – 1.1; R^2≥0.97	Draft: E: 0.87 – 1.1; R^2≥0.97 Current[f]: E: 0.90 – 1.1; R^2≥0.98
Accuracy[c]	Within half a log of expected values	
Repeatability[d]	Within-run standard deviation ≤ 1 Cq; Run-to-run range ≤ 2 Cq	

[a]Negative controls include no template controls (NTCs) and negative extraction controls (NECs); [b]E, amplification efficiency, and R^2 are from the standard curve regression equations: $E= 10^{-1/slope} – 1$; [c]Accuracy evaluates if the expected results are produced; [d]Repeatability evaluates variability within a run and between runs; [e]This was chosen based on experience in the region, but less strigent than that recommended in EPA1609 (E:0.90-1.1, R^2≥0.99); [f]As recommended in Shanks et al 2016.

Results

Cleanliness

All laboratories performed qPCR without cross contamination, regardless of the method complexity, during both the intercalibration phase and the 3-year long-term implementation. For *Enterococcus*, all but one lab passed the cleanliness criteria during the intercalibration (a total of 70 NTCs and 72 NECs) and all labs passed the criteria during the 3-year implementation (a total of 704 NTCs and 694 NECs). The one lab that did not pass the cleanliness criteria during the intercalibration study had more than 1/3 NECs amplified but with low signal (5 out of 9 NECs amplified, Cq ranged from 36.0 to 38.4) and no NTC amplification, indicating potential cross contamination during the crude DNA extraction step.

For HF183, no NEC (n=95) or NTC (n=72) reactions amplified in any lab during the intercalibration. Similarly, there was little sign of cross-contamination during the 3-year implementation: Among the 936 NTCs, only 2 reactions showed a low level of amplification (Cq>37). Among the 1179 NECs, only 2 amplified (Cq>36).

Standard curves

Laboratories generally obtained high quality standard curves, but there were some differences among labs and between the two qPCR methods (Figures 1 and 2). All but two labs (one for Enterococcus, and two for HF183) met the draft performance criteria during the intercalibration phase for either E or R^2 (Figure 1). These two labs received additional technical support and conducted additional practice to ensure meeting performance criteria before moving on to the 3-year implementation phase.

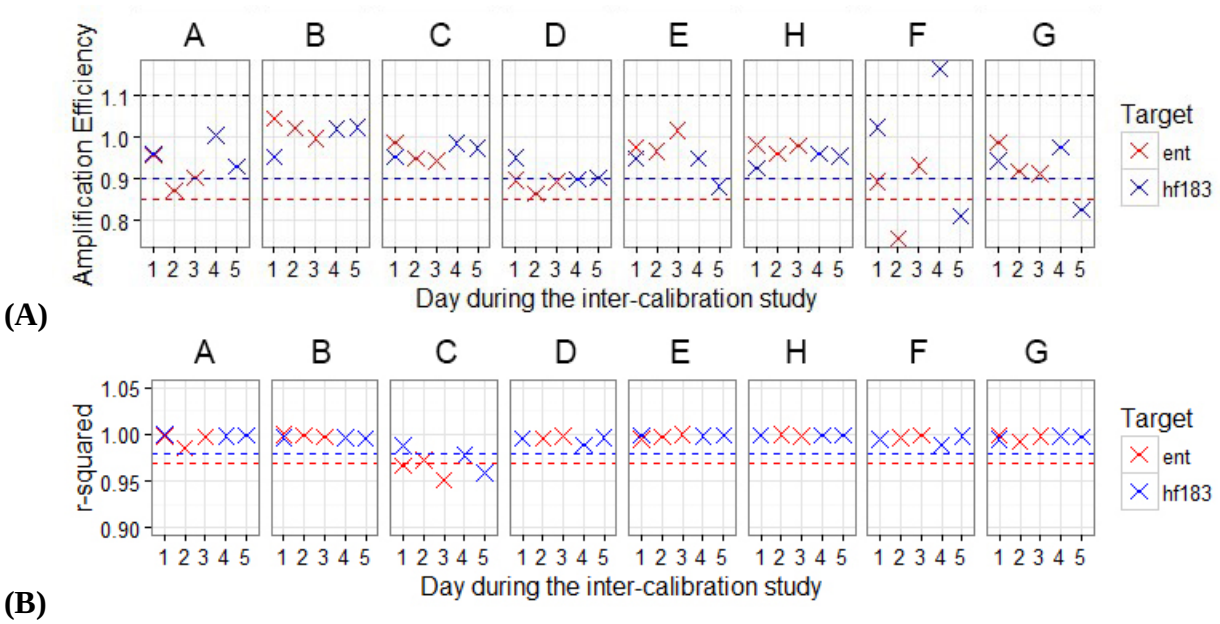
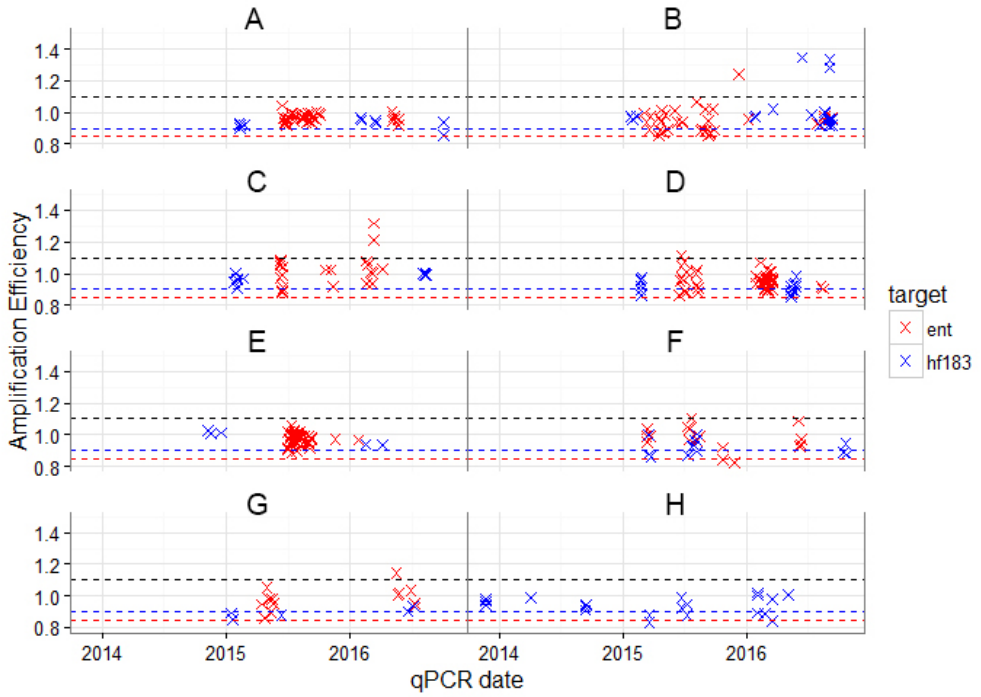
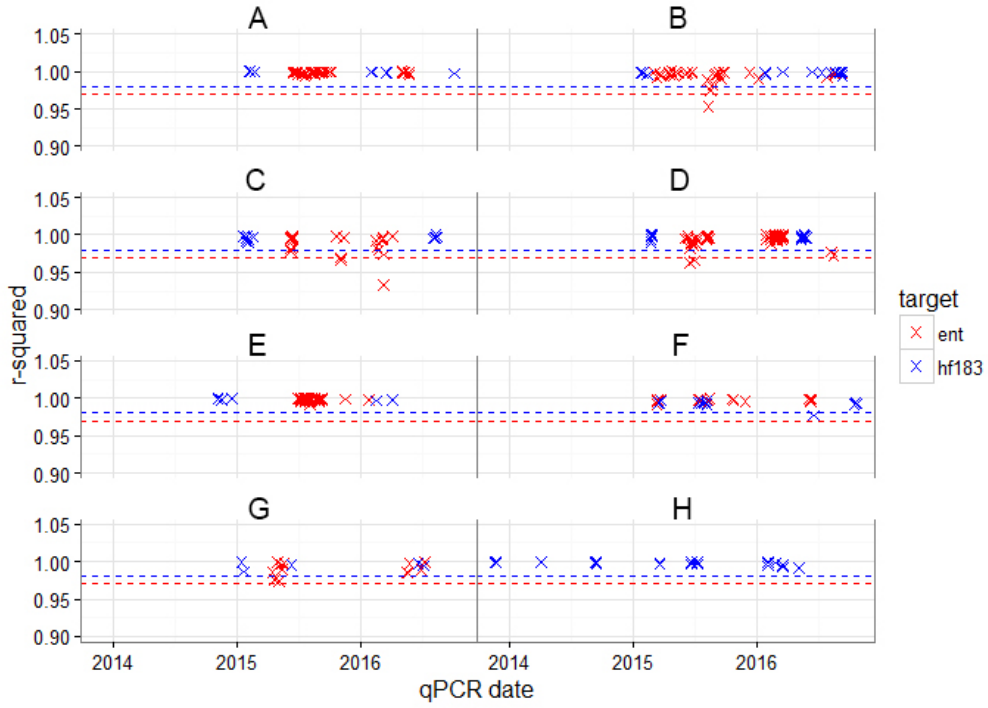


Figure 1. Amplification efficiency (A) and R^2 (B) of the standard curves during the intercalibration phase. For amplification efficiency, black (for both *Enterococcus* and HF183), blue (for HF183), and red (for *Enterococcus*) dashed lines indicate E=1.1, 0.9, 0.85, respectively, as the boundaries for acceptable E. For R^2, blue (for HF183) and red (for *Enterococcus*) dashed lines indicate 0.98 and 0.97, respectively, as lower limits of acceptable R^2.

During the long-term implementation, 7% (n=231) and 27% (n=101) failure of the amplification efficiency criteria (E, Figure 2A), and 4% and 1% failure of the R^2 criteria (Figure 2B) were observed overall for the *Enterococcus* and the HF183 qPCR assays, respectively, based on current performance criteria. Overall failure reates were reduced to 7% for HF183 amplfication efficiency if assessed based on the draft criteria. Lab differences were observed, with frequency of failure for amplification efficiency differing greatly among labs for both *Enterococcus* (0-22% and 0-40% under current and draft criteria, respectively, n=13-55 curves) and HF183 (0-60% and 0-25% under current and draft criteria, respectively, n=5-22 curves) (Table S1). Overall, labs performed best in meeting the R^2 criteria for the HF183 assay. Only one lab passed both standard curve criteria with both qPCR methods (n=47, 6 curves for enterococci and HF183, respectively), regardless if draft or current criteria was used.Movement from an efficiency criteria of 0.87 to 0.90 for the HF183 assay increased the frequency of failure for four out of the eight labs. Detailed assessment is provided in Supplementary Information by lab (Supplementary Information Table S1) and by the dates when qPCR was run (Figures S1 and S2).



(A)



(B)

Figure 2. Amplification efficiency (A) and R^2 (B) of the standard curves during the 3-year implementation. For amplification efficiency, black (for both *Enterococcus* and HF183), blue (for HF183), and red (for *Enterococcus*) dashed lines indicate E=1.1, 0.9, 0.85, respectively, as the boundaries for acceptable E. For R^2, blue (for HF183) and red (for *Enterococcus*) dashed lines indicate 0.98 and 0.97, respectively, as lower limits of acceptable R^2.

Accuracy

Nearly all labs produced *Enterococcus* and HF183 quantification within the expected range, regardless of the concentration of the blind sample, the day samples were analyzed, or which replicate filters were analyzed (Figures 3, 4). The quantification also correctly reflected the 10-fold differences of the three sewage concentration levels. Additionally, HF183 results from Day 4 and Day 5 were not distinguishable in any lab (Figure 4), indicating excellent proficiency in performing the DNA purification procedure required for the HF183 qPCR method. However, one lab consistently overestimated *Enterococcus* concentrations in all three sets of blind sewage-spiked challenge samples (Figure 3), indicating a systematic bias likely caused by the actual concentration of the *Enterococcus* calibrator cells being lower than the assumed 10^5 cells per filter, possibly due to operational errors in this lab (e.g. calibrator degradation during storage, loss in calibrator DNA extraction and/or pipetting).

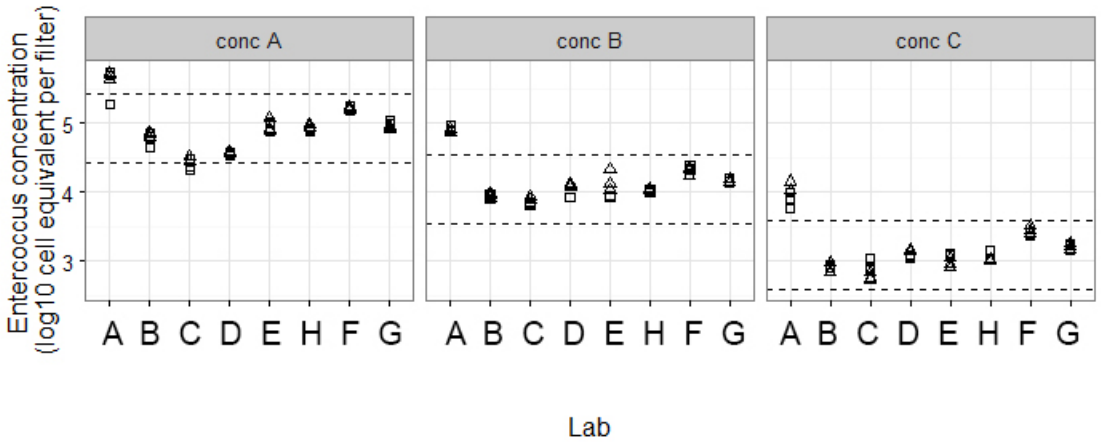


Figure 3. *Enterococcus* results for blind filter samples by each lab (x-axis) during the intercalibration study. Dashed lines indicate expected ranges of concentration (median±0.5 $\log_{10}$ unit) for each sewage spike concentration (conc A, B, and C as indicated in the grey banner of each panel). Different symbols denote replicate filters.

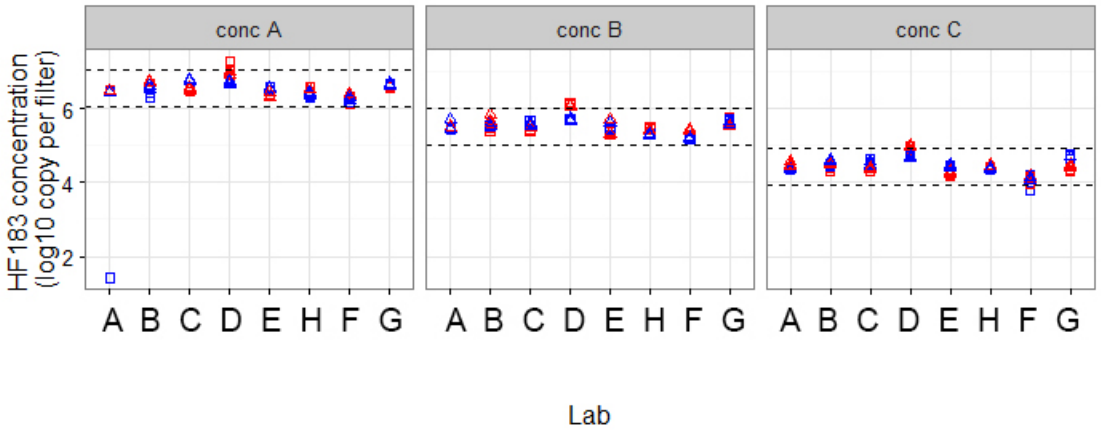


Figure 4. HF183 results for blind filter samples by each lab (x-axis) during the intercalibration study. Dashed lines indicate expected ranges of concentration (median±0.5 $\log_{10}$ unit) for each sewage spike concentration (conc A, B, and C as indicated in the grey banner of each panel). Different symbols denote replicate filters, and colors differentiate Day 4 and Day 5 of the intercalibration study.

Repeatability

Repeatability was much better in some labs than others, much better within-run than between-runs, and better during the intercalibration phase than during the 3-year implementation phase. Repeatability was also better at higher target concentrations than at lower target concentration, as indicated by the decreasing standard deviation within runs (Supplementary Information Figures S3, S4) and decreasing range of Cq values across runs (Figures 5, 6) as target concentration increased.

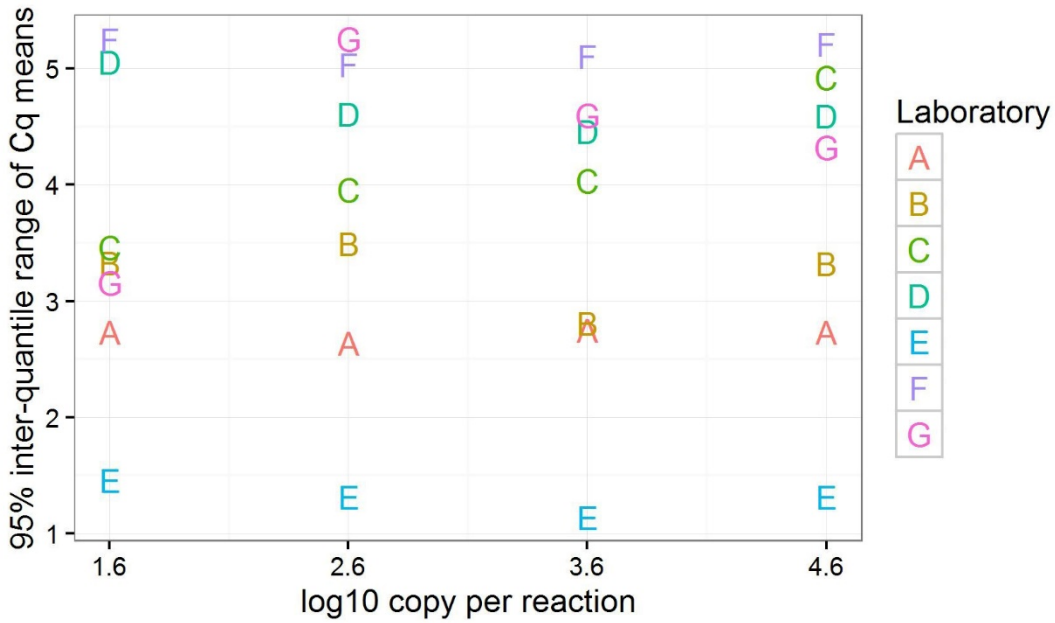


Figure 5. The 95% inter-quantile range of mean Cq values across all runs (13-55 runs) vs target concentration for *Enterococcus* standard curves in each laboratory during the 3-year qPCR implementation. The desired run-to-run Cq range is ≤2 C_q (see Table 1).

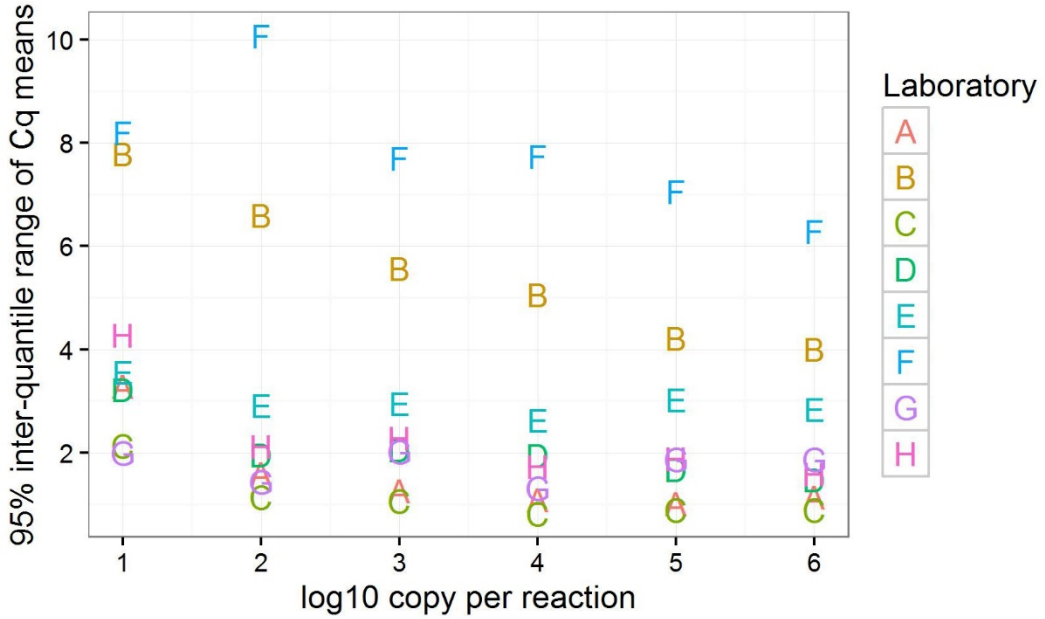


Figure 6. The 95% inter-quantile range of mean Cq values across all runs (5 -22 runs) vs target concentration for HF183 standard curves in each laboratory during the 3-year qPCR implementation. The desired run-to-run Cq range is ≤2 C_q (see Table 1).

At moderately high target concentrations (>1000 copies per reaction), all labs met the performance criteria (Table 1) for within-run repeatability in all runs during the intercalibration, regardless of qPCR method and in nearly all runs (460 out of 462 for *Enterococcus*, 300 out of 301 for HF183) during the 3-year implementation.

During the intercalibration, all labs met the run-to-run repeatability criteria for the *Enterococcus* assay and seven (out of 8) met the criteria for the HF183 assay. During the 3-year implementation, those numbers dropped to one lab and five labs meeting the run-to-run repeatability criteria for *Enterococcus* (Figure 5) and HF183 (Figure 6), respectively, even at the moderately high target concentration. Labs differed greatly in run-to-run repeatability, with high run-to-run variability observed in some labs: >5 Cq for *Enterococcus* and >10 Cq for HF183 (Figures 5, 6).

Discussion

Technology transfer of qPCR from research laboratories to local monitoring laboratories was found to be feasible for both the *Enterococcus* and the HF183 methods. The participating labs demonstrated proficiency in all four areas of qPCR performance during the intercalibration phase. This success is consistent with previous single lab validation and multiple lab intercalibration studies (Griffith and Weisberg 2011, Lavender and Kinzelman 2009, Shanks et al. 2016, Shanks et al. 2012). In previous studies, the process of extracting and purifying DNA was found to contribute to large errors in qPCR results even within research labs (Cox and Goodwin 2013, Pan et al. 2010). Yet, labs in this study demonstrated great consistency and proficiency in performing the full DNA extraction procedure (Figure 4). This success was likely due to the simple and streamlined DNA extraction procedures for the qPCR methods (Green et al. 2014, U.S. EPA 2012), compared to those involving many complicated, and often manual, steps (Pan et al. 2010). This highlights the importance of being conscious of procedural choices and mindful of difficulty for future technology transfer during method development.

Also consistent with previous studies was our observation that qPCR variability increased with decreasing target concentration (Cao et al. 2015, Wang et al. 2014) and that reference/calibration material greatly impacted qPCR quantification of unknown samples (Cao et al. 2013, Sivaganesan et al. 2011). For example, the consistent overestimation of *Enterococcus* concentration by one lab in all blind samples during intercalibration (Figure 3) was likely caused by actual concentration of the *Enterococcus* calibrator cells being lower than the assumed 10^5 cells per filter, a scenario that may result from calibrator degradation, loss during calibrator DNA extraction, or faulty pipetting.

The effort spent by each lab to achieve the performance reported was not tracked; however, the decrease in performance under current versus draft criteria for HF183 (Table S1) is interesting in this regard. It can be expected that labs would have provided additional effort (e.g., re-run plates) to meet the updated criteria. However, the need to provide such effort suggests underlying challenges in the technology transfer that may not have been fully captured or explained.
Not all labs sustained the proficiency in standard curve performance or repeatability over the 3-year long-term implementation. Overall, some labs performed better than others, but root causes are unclear. Some performance issues could have been due to degradation of standard reference material, leading to issues with efficiency and higher variability in Cq values, with higher chance of failing the criteria. However, no systematic increase in mean Cq values was observed over

time (Figures S1, S2). It was noted that labs that failed to sustain proficiency had frequent staff turnover, and it is possible that new or rotating staff members may not have received as much training or opportunity to practice prior to running these test samples. Some labs showed deterioration of performance towards the end of the project, when staff reported being under pressure to meet deadlines. For example, increased data reporting errors were also observed in labs with frequent staff turnover and with tight deadlines imposed by other projects.

The variability and difficulty observed in sustaining qPCR performance over the long term demonstrates the need for continuous formal lab training for qPCR methods and an accreditation program that regularly evaluates and certifies laboratories to ensure continuous proficiency in performing qPCR methods. Such a program will also need to provide (or identify vendors to provide) reference material for qPCR method proficiency testing, as well as routine analysis, and to establish a set of clearly defined data acceptance criteria. While lab accreditation programs for culture-based methods cannot be directly applied to qPCR methods, accreditation programs for molecular methods may be built upon guidelines from other industries (e.g., food and clinical health testing) and performance evaluation studies in the water testing community.

The advent of digital PCR as a replacement for qPCR may reduce the level of difficulty in maintaining proficiency. As digital PCR quantifies unknown samples directly without requiring standard curves, moderate variability in amplification efficiency (as those failed standard curve criteria in this study) and potential degradation/loss of reference material (as likely occurred in one lab for the *Enterococcus* calibrator filter) would no longer affect quantification accuracy (Cao et al. 2016, Cao et al. 2015). Similarly, the higher precision inherent with digital PCR over qPCR could greatly improve method repeatability across labs and across runs (Cao et al. 2016, Hindson et al. 2013).

Acknowledgments

Staff from many agencies contributed to this project. The eight core agencies are City of Los Angeles Environmental Monitoring Division (CLA), City of San Diego (CSD), Los Angeles County Sanitation District (LACSD), Orange County Public Health Lab (OCPHL), Orange County Sanitation District (OCSD), Ventura County Public Health Lab (VCPHL), Weston Solutions (Weston), and Southern California Coastal Water Research Project (SCCWRP).

References

Ahmed, W., B. Hughes and V. Harwood. 2016. Current Status of Marker Genes of Bacteroides and Related Taxa for Identifying Sewage Pollution in Environmental Waters. Water 8(6), 231.

Boehm, A.B., L.C. Van De Werfhorst, J.F. Griffith, P.A. Holden, J.A. 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(18), 6812-6828.

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(4), 611-622.

Cao, Y., J.F. Griffith, S. Dorevitch and S.B. Weisberg. 2012. Effectiveness of qPCR permutations, internal controls and dilution as means for minimizing the impact of inhibition while measuring *Enterococcus* in environmental waters. Journal of Applied Microbiology 113(1), 66-75.

Cao, Y., J.F. Griffith and S.B. Weisberg. 2016. Methods in Molecular Biology Series: Marine Genomics. Bourlat, S.J. (ed), Springer, New York, NY.

Cao, Y., M.R. Raith and J.F. Griffith. 2015. Droplet digital PCR for simultaneous quantification of general and human-associated fecal indicators for water quality assessment. Water Research 70, 337-349.

Cao, Y., M. Sivaganesan, J. Kinzelman, A.D. Blackwood, R.T. Noble, R.A. Haugland, J.F. Griffith and S.B. Weisberg. 2013. Effect of platform, reference material, and quantification model on enumeration of Enterococcus by quantitative PCR methods. Water Research 47(1), 233-241.

Converse, R.R., A.D. Blackwood, M. Kirs, J.F. Griffith, and R.T. Noble. 2009. Rapid QPCR-based assay for fecal *Bacteroides* spp. as a tool for assessing fecal contamination in recreational waters. Water Research 43(19), 4828-4837.

Cox, A.M. and K.D. Goodwin. 2013. Sample preparation methods for quantitative detection of DNA by molecular assays and marine biosensors. Marine Pollution Bulletin 73(1), 47-56.

Ebentier, D.L., K.T.Hanley, Y. Cao, B. Badgley, A. Boehm, J. Ervin, K.D. Goodwin, M. Gourmelon, J. Griffith, P. Holden, C.A. Kelty, S. Lozach, C. McGee, L. Peed, M. Raith, M.J. Sadowsky, E. Scott, J. Santodomingo, C. Sinigalliano, O.C. Shanks, L.C.V.D. Werfhorst, D. Wang, S. Wuertz and J. Jay. 2013. Evaluation of the repeatability and reproducibility of a suite of PCR-based microbial source tracking methods. Water Research 47(18), 6839-6848.

Ferretti, J.A., H.V. Tran, E. Cosgrove, J. Protonentis, V. Loftin, C.S. Conklin and R.N. Grant. 2011. Comparison of Enterococcus density estimates in marine beach and bay samples by real-time polymerase chain reaction, membrane filtration and defined substrate testing. Marine Pollution Bulletin 62(5), 1066-1072.

Ferretti, J.A., H.V. Tran, S.J. Peterson and V. Loftin. 2013a. Rapid method demonstration project at four New Jersey marine beaches using real time quantitative Polymerase Chain Reaction (qPCR). Marine Pollution Bulletin 71(1-2), 51-63.

Ferretti, J.A., H.V. Tran, S.J. Peterson and V. Loftin. 2013b. Rapid method demonstration project at four New Jersey marine beaches using real time quantitative Polymerase Chain Reaction (qPCR). Marine Pollution Bulletin 71(1), 51-63.

Green, H.C., R.A. Haugland, M. Varma, H.T. Millen, M.A. Borchardt, K.G. Field, W.A. Walters, R. Knight, M. Sivaganesan, C.A. Kelty and O.C. Shanks. 2014. Improved HF183 quantitative real-time PCR assay for characterization of human fecal pollution in ambient surface water samples. Applied & Environmental Microbiology 80(10), 3086-3094.

Griffith, J.F. and S.B. Weisberg. 2011. Challenges in implementing new technology for beach water quality monitoring: Lessons from a California demonstration project. Marine Technology Society Journal 45, 65-73.

Harwood, V.J., A.B. Boehm, L.M. Sassoubre, K. Vijayavel, J.R. Stewart, T.T. Fong, M.P. Caprais, R.R. Converse, D. Diston, J. Ebdon, J.A. Fuhrman, M. Gourmelon, J. Gentry-Shields, J.F. Griffith, D.R. Kashian, R.T. Noble, H. Taylor and M. Wicki. 2013. Performance of viruses and bacteriophages for fecal source determination in a multi-laboratory, comparative study. Water Research 47(18), 6929-6943.

Hindson, C.M., J.R. Chevillet, H.A. Briggs, E.N. Gallichotte, I.K. Ruf, B.J. Hindson, R.L. Vessella and M. Tewari. 2013. Absolute quantification by droplet digital PCR versus analog real-time PCR. Nat Methods 10(10), 1003-1005.

Lavender, J.S. and J.L. Kinzelman. 2009. A cross comparison of QPCR to agar-based or defined substrate test methods for the determination of Escherichia coli and enterococci in municipal water quality monitoring programs. Water Research 43(19), 4967-4979.

Noble, R.T., A.D. Blackwood, J.F. Griffith, C.D. McGee and S.B. Weisberg. 2010. Comparison of rapid QPCR-based and conventional culture-based methods for enumeration of *Enterococcus* spp. and *Escherichia coli* in recreational waters. Applied and Environmental Microbiology 76(22), 7437-7443.

Pan, Y., L. Bodrossy, P. Frenzel, A.G. Hestnes, S. Krause, C. Lüke, M. Meima-Franke, H. Siljanen, M.M. Svenning and P.L.E. Bodelier. 2010. Impacts of Inter- and Intralaboratory Variations on the Reproducibility of Microbial Community Analyses. Applied and Environmental Microbiology 76(22), 7451-7458.

Shanks, O.C., C.A. Kelty, R. Oshiro, R.A. Haugland, T. Madi, L. Brooks, K.G. Field and M. Sivaganesan. 2016. Data Acceptance Criteria for Standardized Human-Associated Fecal Source Identification Quantitative Real-Time PCR Methods. Applied & Environmental Microbiology 82(9), 2773-2782.

Shanks, O.C., M. Sivaganesan, L. Peed, C.A.Kelty, A.D. Blackwood, M.R. Greene, R.T. Noble, R.N. Bushon, E.A. Stelzer, J. Kinzelman, T. Anan'eva, C.Sinigalliano, D. Wanless, J. Griffith, Y, Cao, S. Weisberg, V.J.Harwood, C. Staley, K.H. Oshima, M. Varma and R.A. Haugland. 2012. Inter-laboratory general fecal indicator quantitative real-time PCR methods comparison study. Environmental Science & Technology 46(2), 945–953.

Sivaganesan, M., R.A. Haugland, E.C. Chern and O.C. Shanks. 2010. Improved strategies and optimization of calibration models for real-time PCR absolute quantification. Water Research 44(16), 4726-4735.

Sivaganesan, M., S. Siefring, M. Varma and R.A. Haugland. 2011. MPN estimation of qPCR target sequence recoveries from whole cell calibrator samples. Journal of Microbiological Methods 87(3), 343-349.

U.S. EPA. 2012. Method 1611: Enterococci in Water by TaqMan® Quantitative Polymerase Chain Reaction (qPCR) Assay. EPA-821-R-12-008, Office of Water, Washington, DC.

U.S. EPA (2013) Method 1609: Enterococci in Water by TaqMan® Quantitative Polymerase Chain Reaction (qPCR) with Internal Amplification Control (IAC) Assay. EPA-820-R-13-005, Office of Water, Washington, DC

Wang, D., H.C. Green, O.C. Shanks and A.B. Boehm. 2014. New Performance Metrics for Quantitative Polymerase Chain Reaction-Based Microbial Source Tracking Methods. Environmental Science & Technology Letters 1(1), 20-25.

Supplementary Tables

Table S1. Frequency of failure of standard curve performance criteria by lab. Performance criteria were defined in Table 1.

	Enterococcus					HF183				
		% fail draft criteria		% fail current criteria			% fail draft criteria		% fail current criteria	
Lab	n	E	R2	E	R2	n	E	R2	E	R2
A	39	0%	0%	0%	0%	10	10%	0%	10%	0%
B	35	14%	6%	6%	6%	18	17%	0%	17%	0%
C	24	17%	17%	17%	17%	11	0%	0%	0%	0%
D	55	13%	4%	11%	4%	16	25%	0%	44%	0%
E	47	0%	0%	0%	0%	6	0%	0%	0%	0%
F	18	22%	6%	22%	6%	13	23%	0%	54%	8%
G	13	15%	0%	8%	0%	5	20%	0%	60%	0%
H	-	-	-	-	-	22	9%	0%	27%	0%

Table S2. Frequency of within-run standard deviation (among qPCR replicates) meeting the performance criteria of ≤ 1 Cq, during the 3-year implementation phase. All target concentrations were lumped.

	Enterococcus		HF183	
Lab	n	% Pass	n	% Pass
A	156	100	60	95
B	140	98	108	95
C	96	98	66	97
D	220	98	96	88
E	188	100	36	92
F	72	100	130	90
G	51	94	77	85
H	-	-	30	86

Supplementary Figures

Figure S1. Mean Cq values for the highest concentration (4000 copy per reaction) in *Enterococcus* standard curve from run to run in each laboratory over the 3-year qPCR implementation period.

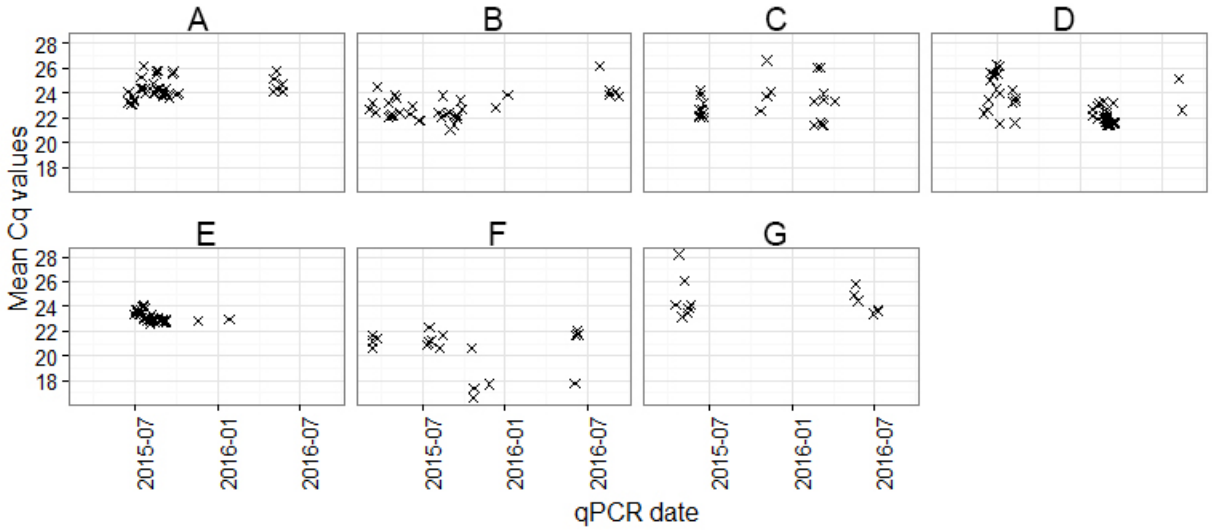


Figure S2. Mean Cq values for the highest concentration (1×10^6 copy per reaction) in HF183 standard curve from run to run in each laboratory over the 3-year qPCR implementation period.

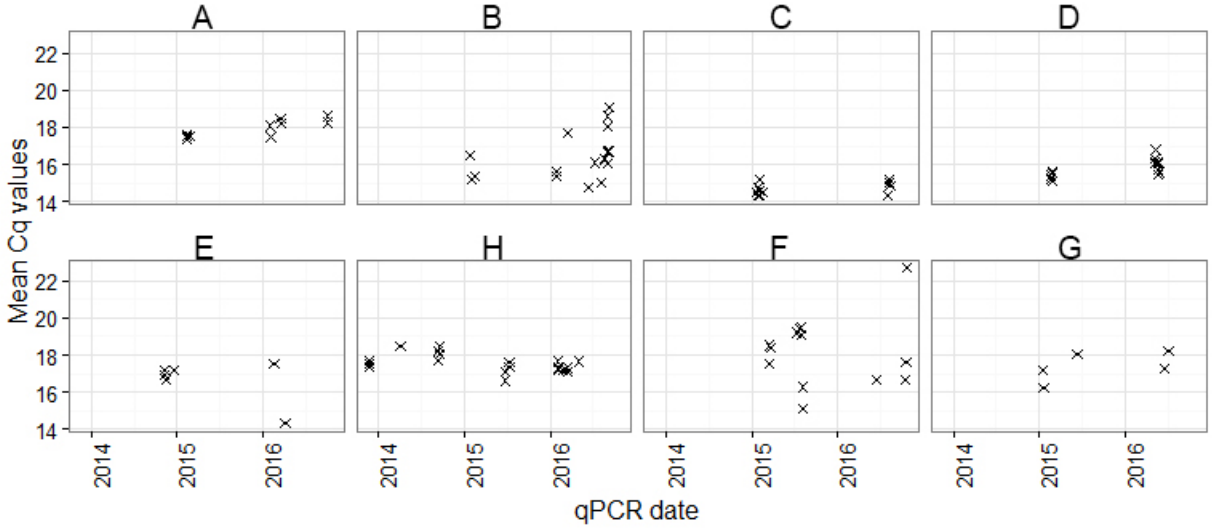


Figure S3. Within-run repeatability (measured by standard deviation between qPCR replicates within a run) at all concentration levels by lab, for the *Enterococcus* qPCR assays during the 3-year implementation phase. Dashed lines indicate the upper limit of acceptable standard deviation (1 Cq).

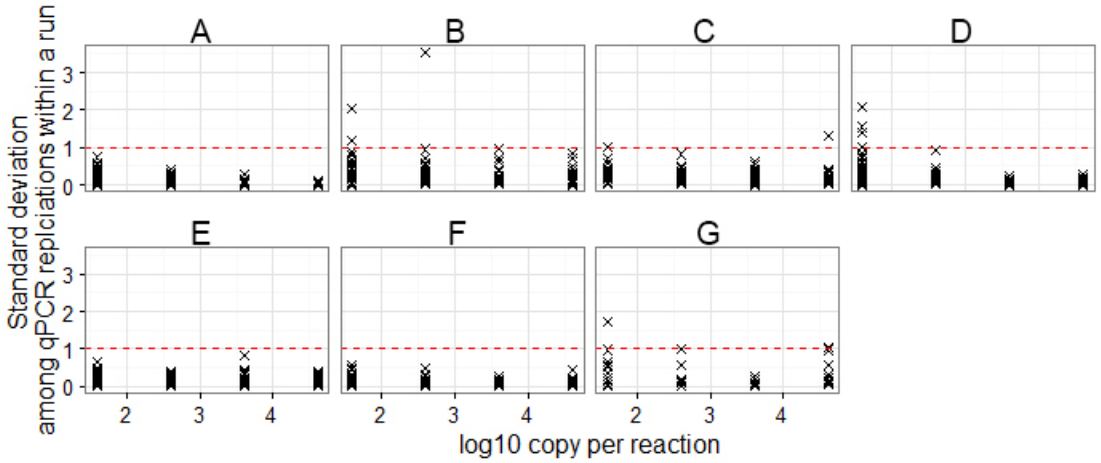
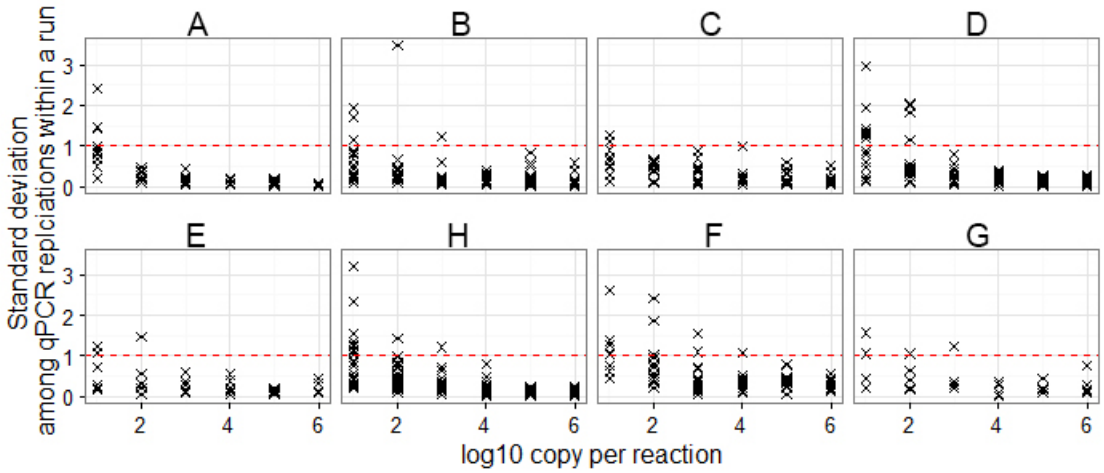


Figure S4. Within-run repeatability (measured by standard deviation between qPCR replicates within a run) at all concentration levels by lab, for the HF183 qPCR assays during the 3-year implementation phase. Dashed lines indicate the upper limit of acceptable standard deviation (1 Cq).



CHAPTER 2. ASSESSMENT OF *ENTEROCOCCUS* QPCR METHODS APPLICABILITY AT THIRTY-SIX SOUTHERN CALIFORNIA BEACHES

Introduction

Quantitative polymerase chain reaction (qPCR) methods for beach water quality monitoring yield results in considerably less time than the culture-based methods they are intended to replace, providing opportunity for improved public health protection. Several demonstration projects have illustrated the feasibility of transitioning this technology from research laboratories to routine monitoring programs (Ferretti et al. 2013, Griffith and Weisberg 2011) and a number of epidemiological studies have found a strong relationship between qPCR-based sampling and health risk (Colford et al. 2012, Wade et al. 2008, Wade et al. 2006, Yau et al. 2014). As a result, the US Environmental Protection Agency (USEPA) released the revised Recreational Water Quality Criteria (RWQC) in 2012 that allowed for use of qPCR for routine beach health notification monitoring (U.S. EPA 2012b).

However, U.S.EPA recommends side-by-side testing with currently used culture-based methods and qPCR methods before the latter are adopted for implementation at a site. This parallel testing addresses three potential concerns. First, it provides a demonstration that the laboratory can produce acceptable quality data with the new method. Second, it ensures that water at the site does not consistently contain constituents that inhibit the qPCR reactions, which could cause underestimation potentially detrimental to public health protection (Cao et al. 2012, Haugland et al. 2016). Third, it provides a means for establishing qPCR thresholds through determining site-specific equivalence relationships to EPA approved culture methods. This is necessary because a single equivalence value across all sites is unlikely as qPCR and culture relationships differ among sites (Whitman et al. 2010). Using the qPCR beach action values (BAV) in the 2012 RWQC as benchmark is also not an appealing option, because the BAVs are based on epidemiology studies conducted at sites predominantly influenced by treated wastewater and appear poorly applicable to sites with mixed and diffusive sources (Raith et al. 2013).

USEPA published a technical support material (TSM) document that describes procedures for assessing sufficiency of the relationship between alternative analytical methods (e.g. an enterococci qPCR method) and benchmark EPA-approved methods (e.g. EPA method 1600), as well as for determining site-specific implementation thresholds for the alternative methods (U.S. EPA 2014). Many studies have conducted parallel processing to compare results between enterococci qPCR and culture methods (Converse et al. 2012, Haugland et al. 2005, Lavender and Kinzelman 2009, Noble et al. 2010, Raith et al. 2013), but none have tested the adequacy of the relationship using the new USEPA TSM procedures. Here, we assess the applicability of enterococci qPCR in southern California by monitoring 36 beaches with side-by-side culture and qPCR methods employed by seven agencies that routinely monitor beach water quality. The regional dataset was used to assess prevalence of matrix interference and regional applicability of qPCR by using EPA-approved culture methods as benchmark following the procedure prescribed in the TSM.

Methods

Sites, sampling, and culture-based analyses

Approximately fifty samples per season were collected from thirty-six southern California beach sites (Figure 1) sampled under summer and winter dry-weather conditions between 2013 and 2015. Beach sites were selected based on importance of the sites to routine recreational water quality monitoring, and included both embayment and open coastal beaches. Sites were also categorized as high vs. low organic material sites observationally depending on whether fresh or decaying vegetation (e.g. kelp, eelgrass) was present at the beach during more than half of the sampling events. Summer season was customarily defined as the main recreation season (April 1- October 31) specified in California statute (Assembly Bill No. 411), and November 1 to March 31 was considered winter season. No sample was collected if there was >0.1 inches rain within three days preceding the scheduled sampling.

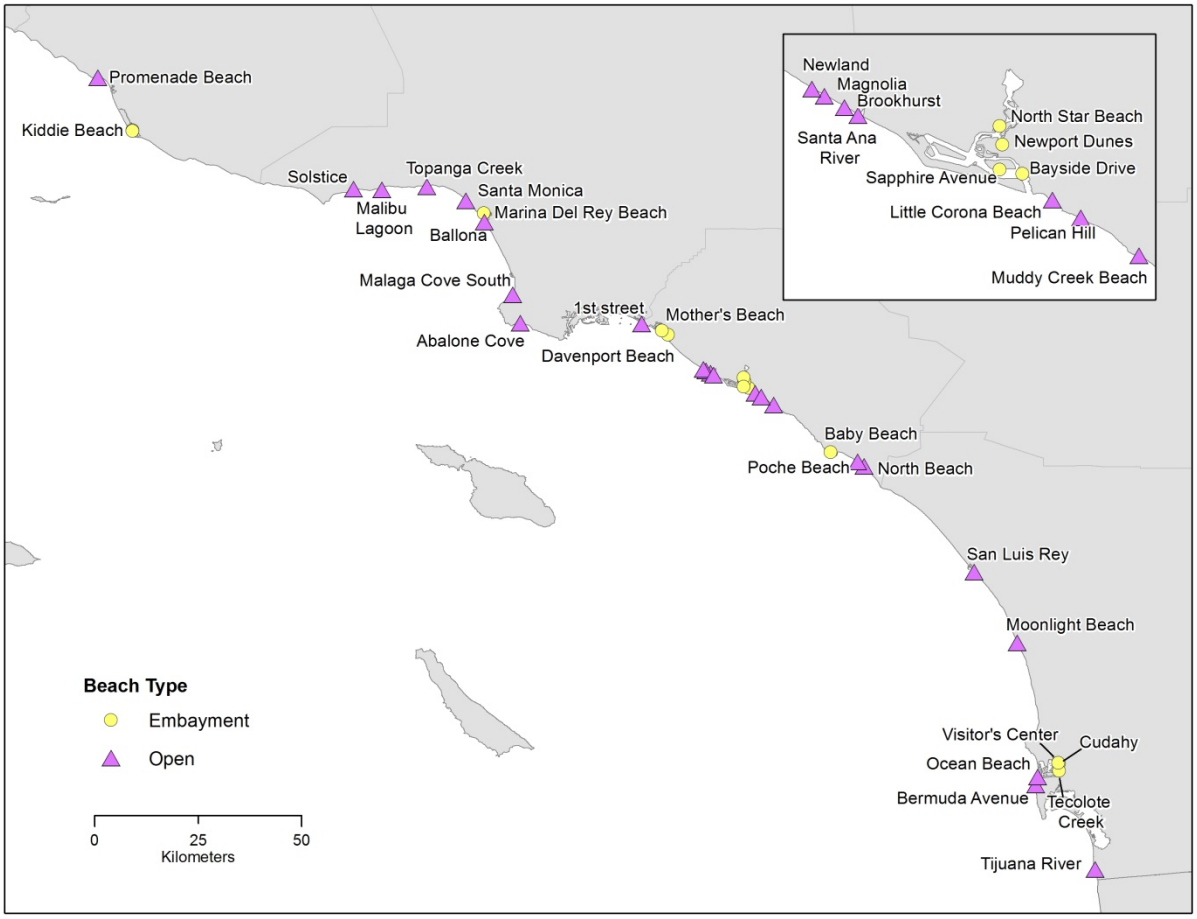


Figure 1. Sampling sites. Circles and triangles indicate embayment and open coast beach sites, respectively.

All samples were collected in the early morning to mirror current monitoring protocol and schedule (generally weekly) for each site, and transported on ice to the laboratory for enterococci analysis by culture-based methods and for filtration for qPCR analysis within 6 hours. Culture-based methods for enterococci were EPA Method 1600 or Enterolert (IDEXX, Westbrook, ME), depending on the routine procedure employed by the participating labs. The lower limit of quantification (LLOQ) and upper limit of quantification (ULOQ) were reported by each lab based on sample dilution level and method range of quantification, which is determined by colony countable range (1-200 colonies per plate per routine procedure employed by the participating labs) or the most probable number (MPN) range (as defined in the Enterolert MPN table).

Filtration, DNA extraction, and qPCR analyses

Samples were filtered (0.45µM polycarbonate filters, 100 ml per filter or until clogging) within 6 hours of sample collection and stored at -80°C until crude DNA extraction using extraction buffer spiked with a known amount of salmon testes DNA, following the published protocol (U.S. EPA 2012a). DNA extraction was performed in batches of 16 samples, one *Enterococcus* whole cell calibrator, and 1 negative filtration/extraction control. The crude DNA extracts were analyzed by qPCR undiluted.

The enterococci qPCR method consists of a simplex assay for the detection of *Enterococcus* spp., TaqEnviron (Cao et al. 2012), and a simplex Sketa assay (i.e., sketa22, (U.S. EPA 2012a)). The TaqEnviron assay follows the EPA 1611 method (U.S. EPA 2012a) targeting the 23S rRNA gene in *Enterococcus,* except the former uses a superior master mix more robust against matrix interference (Cao et al. 2012). The Sketa assay measures the spiked salmon testes DNA to assess the extent of matrix interference (U.S. EPA 2012a).

Seven laboratories were involved in sampling and sample processing and all followed the same standard operating procedures for qPCR, including identical qPCR 96-well plate setups, dictating precise positions for genomic DNA standard curves, calibrators, negative controls, and samples. All samples, standards, and controls were run with duplicate qPCR reactions. Each qPCR plate contained one *Enterococcus* genomic DNA standard curve (4-point, 10-fold dilution, 40000 to 40 copies per reaction), one whole cell calibrator, one negative extraction controls (NEC; one for each batch of DNA extractions consisting of 16 samples), one no-template control (in triplicate), and 16 samples. The *Enterococcus* and Sketa assays occupied the same plate with an identical setup, minus the standard curve for the latter. All reference material including *Enterococcus* genomic DNA and whole cell calibrator (10^5 cells/filter), both from *E. faecalis* (ATCC 29212), and salmon testes DNA (Sigma Aldrich cat 3 D1626) were prepared as described previously (U.S. EPA 2012a) into single use aliquots at a central facility (Southern California Coastal Water Research Project Authority, i.e. SCCWRP), and stored at -80°C before use at the individual laboratories.

Matrix interference was assessed based on the shift of cycle of quantification (Cq) values in samples compared to controls for the Sketa assay following the procedure described in EPA method 1611 (U.S. EPA 2012a). Matrix interference was deemed present in a sample if the sample Sketa Cq in any qPCR replicate was 3 cycles greater than the average Sketa Cq in negative extraction blanks from the same DNA extraction batch (i.e. on the same qPCR plate).

Since different qPCR quantification models would produce different qPCR estimates of *Enterococcus* density from the same qPCR raw data, potentially affecting assessment of culture and qPCR agreement, the three most common quantification models (Cao et al. 2013, Shanks et al. 2012) were included to allow a comprehensive evaluation. Briefly, master standard curves (Cq vs. $\log_{10}$ concentration) were calculated for each laboratory using linear regression with an outlier removal procedure (a data point is deemed an outlier if its studentized residual > 3) (Ebentier et al. 2013). In model 1 (reg), the resulting regression equation was used directly to calculate the number of 23S gene copies for a sample, which was then converted to cell equivalents per 100ml water assuming 4 copies of 23S rRNA gene per *Enterococcus* cell. Models 2 (commonly referred as ΔCt) and 3 (commonly referred as ΔΔCt) used the whole cell calibrator (predetermined to have 10^5 cells by microscopy) to convert gene copies to cell equivalents, without (ΔCt) or with (ΔΔCt) adjustment, respectively, for potential sample recovery and matrix interference using the sample Sketa Cq values. With increasing numbers of model input parameters from models 1 to 3, increasing variability in qPCR results may be expected (Cao et al. 2013, Shanks et al. 2012). For all three quantification models, the lower limit of quantification (LLOQ) was set at the lowest concentration where all qPCR replicates amplified in all labs (40 copies per reaction) and expressed as Cq + 2×Standard deviation for each lab (Stewart et al. 2013).

Assessment of culture and qPCR method agreement

Method agreement was assessed following the procedure in the USEPA TSM that involves calculation of two metrics: Index of Agreement (IA), and regression coefficient of determination (R^2) (U.S. EPA 2014). R^2 was based on regression of enterococci concentration by qPCR (tENT) vs. enterococci concentration by culture (cENT). IA was calculated as:

$$IA = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2}{\frac{1}{N}\sum_{i=1}^{N}(|x_i - \bar{x}| + |y_i - \bar{x}|)^2}$$

Where *x* represents cENT, and *y* represents tENT, $\bar{x}$ is the average cENT at the site, *i* is a counter and *N* is the total number of culture-qPCR paired data points (N ≥30 required).

In conducting these calculations, all sample data below the LLOQ or above upper limit of quantification (ULOQ) or with qPCR matrix interference were excluded per the prescribed procedure. Only sites with ≥30 samples remaining were eligible for assessment of agreement using the above metrics. Both IA and R^2 were calculated using $\log_{10}$-transformed data.

An IA≥0.7 or R^2>0.6 indicates sufficient agreement between qPCR and culture to allow replacing culture with qPCR methods. If IA≥0.7, the same numeric criteria used by culture-based methods can be applied directly to qPCR data. If IA<0.7 and R^2>0.6, the culture criteria can be plugged into the qPCR-culture regression equation to directly calculate the corresponding qPCR criteria.

Results

Paired culture (cENT) and qPCR (tENT) data were generated for 1826 and 1526 samples at 36 sites during summer and winter sampling seasons, respectively. Numbers of samples per site ranged from 39 to 60 in summer, and from 24 to 50 in winter (Table 1). All 36 sites in summer and 32 sites in winter had at least 30 samples with paired cENT and tENT results.

Table 1. Number of samples with paired culture and qPCR data (# paired samples), percentage of samples with cENT (Enterococcus density by culture-based methods) greater than California Ocean Plan single sample maximum of 104/100ml (%>104), percentage of samples that failed the matrix interference control (%skFail), and number of samples with both cENT and tENT (Enterococcus density by qPCR) within ROQ (# paired in ROQ), for summer and winter. Sites are listed approximately from north to south as shown in the map (Figure 1). Note that # paired in ROQ may include samples that failed the matrix interference control, which were excluded prior to subsequent qPCR-culture agreement analysis.

	Summer				Winter			
Site	# paired samples	%>104	%skFail	# paired in ROQ	# paired samples	%>104	%skFail	# paired in ROQ
Promenade Beach	39	3	10	3	29	3	15	2
Hobie Beach	43	2	0	12	36	6	18	9
Kiddie Beach	39	3	8	18	36	17	8	22
Solstice	50	2	2	12	50	2	0	13
Malibu Lagoon	50	0	2	9	50	20	2	30
Topanga	50	6	0	4	50	12	0	20
Santa Monica	50	10	2	33	50	20	2	41
Marina del Rey Beach	50	20	0	34	50	42	2	49
Ballona	50	4	2	26	50	14	0	30
Malaga Cove	57	0	2	10	40	0	5	12
Abalone Cove	58	0	19	6	40	0	10	3
1st street	50	26	4	38	43	19	2	30
Mother's Beach	50	2	4	25	49	18	0	42
Davenport Beach	50	2	4	17	50	8	2	27
Newland	49	6	2	5	42	2	2	2
Magnolia	50	8	0	5	44	11	0	4
Brookhurst	50	14	2	6	43	16	0	10
Santa Ana River	50	6	2	10	43	7	5	7
North Star Beach	50	0	6	10	50	14	4	35
Newport Dunes	50	6	12	22	48	17	4	35
Sapphire Avenue	50	6	4	26	50	10	0	31
Bayside Drive	50	10	6	38	50	12	2	38
Little Corona Beach	49	14	0	11	42	5	2	10
Pelican Hill	50	4	0	3	43	9	2	2
Muddy Creek Beach	49	0	0	0	43	5	0	4
Baby Beach	50	26	4	40	48	13	0	45
Poche Beach	50	2	2	17	49	2	10	21
North Beach	50	2	0	10	50	4	0	23
San Luis Rey	44	0	8	1	38	0	8	1
Moonlight Beach	43	12	5	20	39	8	9	18
Visitor's Center	55	45	22	51	33	42	15	32
Cudahy	56	48	13	55	31	71	13	30
Tecolote	56	11	16	50	31	35	23	31
Ocean Beach	60	3	20	34	24	0	25	13
Bermuda Avenue	57	2	28	18	26	4	19	9
Tijuana River	55	5	7	42	27	7	15	22

General water quality in cENT varied greatly among sites, with percentages of samples exceeding California Ocean Plan single sample maximum (SSM) of 104 *Enterococcus* per 100ml ranging from 0 to 48% in summer, and from 0-71% in winter. Limits of quantification for cENT were as reported by each lab, and were generally 1-99 and 400-12000 *Enterococcus*/100ml for LLOQ and ULOQ, respectively, depending on the dilution conducted prior to analysis. Overall, 72% and 81% of the samples were within ROQ for cENT, in summer and winter, respectively.

qPCR standard curves showed satisfactory performance across all labs, with R^2 and amplification efficiency ranging from 0.81 to 0.99 and from 0.93 to 1.00, respectively (Supplementary Information, Table S1). All negative controls (n=1398) met the quality control requirements (U.S. EPA 2012a), showing no issue of cross-contamination. LLOQ ranged from 34 to 37 Cq, and no ULOQ was applied as qPCR ULOQ is generally orders of magnitude higher than necessary for environmental samples. Overall, 46% and 54% of the samples were within ROQ for tENT in summer and winter, respectively.

Observed frequencies of matrix interference were generally low (6% and 5% overall in summer and winter, respectively) albeit differing greatly across sites (Table 1). Among 36 sites, 8 and 10 sites showed no sign of matrix interference while 7 and 8 sites had more than 10% samples failing the matrix interference control, for summer and winter, respectively. While sampling season (summer vs. winter, p-value=0.9) and beach type (embayment vs. open coast, p-value=0.3) did not influence how often matrix interference was observed, sites with high organic material were more likely to have matrix interference than sites with low organic material (p=0.02).

Among 36 sites, only 10 and 9 sites, in summer and winter, respectively, had sufficient data (number of samples with paired qPCR and culture data within ROQ $\geq$ 30, qPCR data with matrix interference excluded) to allow assessment of agreement with the two metrics (IA, R2) set in the TSM. Among these qualified sites, only one site in summer and three sites in winter demonstrated sufficient agreement (IA$\geq$0.7 or R^2>0.6) between culture and qPCR enterococci results, regardless of qPCR quantification model (Figures 2 and 3, Supplementary Information Table S2).

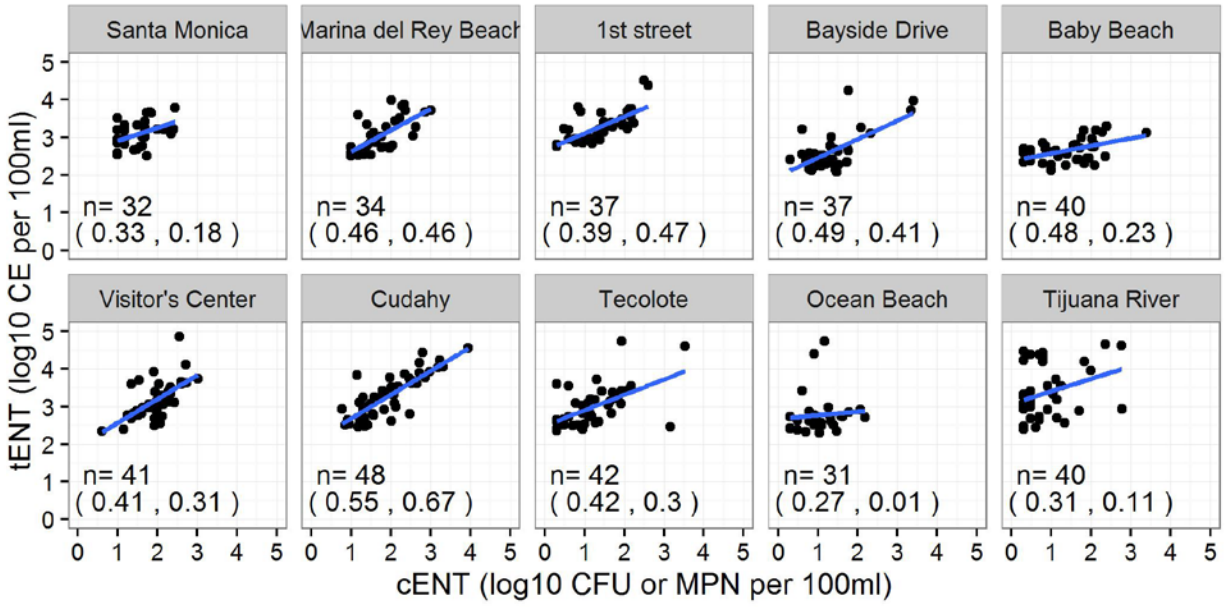


Figure 2. qPCR (tENT) vs. culture (cENT) results for the 10 sites in the summer that had sufficient data (number of samples as displayed for each site) for agreement assessment. Index of agreement (no sites with IA≥0.7) and R2 (one site with R2>0.6) are shown in parenthesis in the order given separated by a comma. Only qPCR quantification model 1 (reg, as described in Methods) was shown here, but the other quantification models show similar trends (Supplementary Information Table S2).

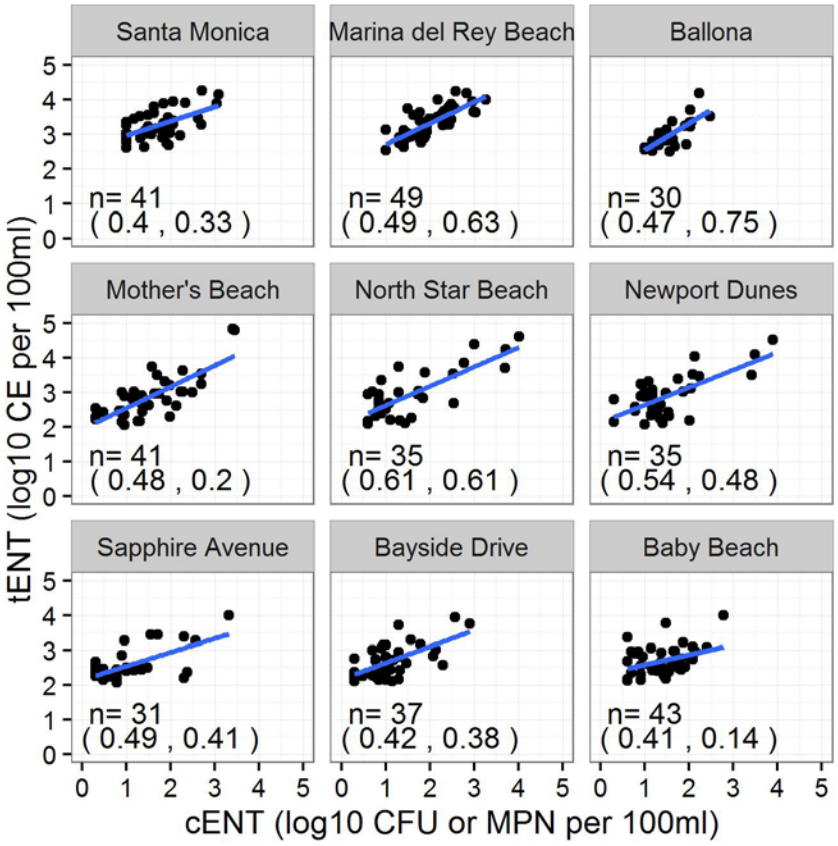


Figure 3. qPCR (tENT) vs. culture (cENT) results for the 9 sites in the winter that had sufficient data (number of samples as displayed for each site) for agreement assessment. Index of agreement (no sites with IA≥0.7) and R^2 (three sites with R^2>0.6) are shown in parenthesis in the order given separated by a comma. Only qPCR quantification model 1 (reg, as described in Methods) was shown here, but the other quantification models show similar trends (Supplementary Information Table S3).

Discussion

There was insufficient agreement between qPCR and culture results following the TSM assessment procedures for application of qPCR at most southern California beaches. There are several possible reasons for this finding. First, qPCR is a relatively new technology for the seven labs in this study and it is possible that their execution of the qPCR methods was ineffective. However, this explanation appears unlikely. SCCWRP, a lab highly experienced in molecular methods, repeated analysis for eight of the sites and found almost no increase in agreement with culture results compared to original qPCR results produced by less experienced labs (Supplementary Information, Table S4). Even inherent qPCR method variability did not seem an important contributor to the failure, as there was not a big difference in agreement with culture methods among the three qPCR quantification models that differ in method variability (Supplementary Information, Tables S2 and S3). Additionally, SCCWRP also ran samples from the same eight sites using droplet digital PCR (ddPCR). While ddPCR has superior accuracy and reproducibility compared to qPCR (Cao et al. 2015, Hindson et al. 2013), it did not show substantially increased agreement with culture results compared to qPCR vs. culture (Supplementary Information, Table S4).

Matrix interference was also not a major obstacle to achieving the agreement between the two methods. Approximately 80% of the sites had less than 10% samples failing the matrix interference control, and one third of these sites had no sample failing at all. Excluding samples with matrix interference (resulting in 5 fewer sites eligible for TSM assessment in winter) did not reduce the number of sites with sufficient sample size ($N \geq 30$) to allow TSM agreement assessment in summer. Additionally, the $\Delta\Delta$Ct qPCR quantification model that aims to adjust for potential slight matrix interference also did not improve agreement with culture results compared to the other two quantification models.

A third possible explanation is that qPCR and culture methods measure different endpoints and those relationships were too variable to support culture-qPCR agreement at each site. Culture-based methods measure the number of enterococci cells that can grow within a prescribed incubation timeframe (e.g. 18-24 hours) on selective media, whereas qPCR measures the genetic presence of cells regardless of their physiological state. Any changes in the ratio of culturable vs. total cells at a site over the course of the study would contribute considerable variability to the qPCR vs. culture relationship. Some of the watersheds in this study are known to use recycled water for irrigation, which is typically disinfected in a manner that eliminates or reduces bacteria cultivability, but might allow retention of unculturable or dead but intact cells containing genetic material that would be measured by qPCR. However, use of recycled water only occurs in a few of the watersheds and would not be responsible for the widespread lack of culture-qPCR agreement across the wide range of watersheds studied.

The most likely explanation for the failure of meeting the EPA TSM requirements at so many beaches is that the enterococci concentrations were at the low end of the quantification range for these methods. This leads to two problems for conducting EPA TSM evaluation. First, even with two years of weekly sample collection, we were not able to achieve the requisite number of samples ($N \geq 30$) in the quantifiable range for many sites. Second, with measurements clustered at the low end of the quantifiable range, this led to higher method variability and shorter range for regression analysis, and thus a lower correlation between methods. This hypothesis is reinforced by the IA and R^2 values having a high correlation with the percentage of samples exceeding the

California Ocean Plan single sample maximum of 104 *Enterococcus* per 100ml ($p<0.05$) and with the geometric mean of culturable enterococci concentration ($p<0.05$) (Supplementary Information, Figure S1). Moreover, the few sites that met culture-qPCR agreement criteria were those with the highest exceedance rates and average concentrations (Figures 2 and 3).

This suggests that some modification of the EPA TSM protocols is necessary if qPCR is to be adoptable at sites that are typically clean with sporadic pollution events. These sites with transient pollution events are where the speed advantage of qPCR is most valuable for improved public health protection and recreational resource utilization. Several studies have found that water quality often changes with respect to thresholds for health warning by the time culture results become available (Kim and Grant 2004, Leecaster and Weisberg 2001, Rabinovici et al. 2004). While epidemiological analyses demonstrated indicator-illness relationship with culture-based methods, such analyses were conducted as if the culture results were available on the same day as the swimmers were exposed. In practice, culture results from the previous day were used for beach posting although no significant association between prior-day culture results and illness from current day exposure was found (Colford et al. 2012).

With the current TSM protocols, it is challenging to obtain enough (minimum requirement: 30 samples) paired culture-qPCR data that fall within the range of quantification at such beaches, and it is even more difficult to collect samples with a sufficiently large portion of high enterococci concentrations to produce IA and R^2 values high enough to meet the thresholds set in the TSM. However, the underlying goals of the USEPA requirements to establish consistent and predictable relationships between the alternative and EPA-approved benchmark methods (e.g. qPCR vs. culture in this study) at each site are still warranted (U.S. EPA 2014).

There are a few possible approaches that might be available to assess that relationship at beaches with generally low enterococci concentrations. One of these is to select a benchmark method (i.e. EPA1611 or EPA1609) that measures the same endpoint as the alternative method. There is some evidence that methods measuring the same endpoint are more likely to tolerate data clustering at the low end of the quantifiable range and pass the TSM criteria than methods measuring different endpoints (data not shown). However, this approach may not be feasible because the numeric thresholds for the benchmark methods (EPA1611 or EPA1609) applied poorly to sites with diffusive sources (Nevers et al. 2013, Raith et al. 2013) presumably because these thresholds were developed at beaches impacted by treated wastewater (U.S. EPA 2012b, Wade et al. 2006).

Another possibility is to use culture-based benchmark methods but assess the relationship in the source water to the beach, rather than at the beach itself. In southern California, the water quality is often defined by nearby creek or drainage systems that serve as the primary source of contamination to the beach, and typical dilution rates are at least 1:1000 between the creek and the beach during dry weather conditions. Using creek water is not as ideal as measuring beach water, since the ratio of culturable vs. total *Enterococcus* could change as the water moves from the freshwater creek to saltwater, but it would allow a more robust evaluation than relying on a relationship between extremely low values. Alternatively, the beach water could be spiked with a small amount of creek water inoculant so that the test occurs in a high salinity sample. At some beaches, it may also be possible to move monitoring locations to creek-ocean confluence (a.k.a. point zero) instead of a fixed distance from the discharge, as is the practice in some counties in southern California.

A third possibility would be to shift the agreement evaluation from a continuous correlative evaluation (i.e. as defined in TSM) to a categorical analysis that concerns only whether beach posting decisions agree when using culture vs. qPCR methods. The rationale for such a shift is that, if both culture and qPCR methods agree which samples are high enough to issue a health warning, the lack of continuous correlative relationship between culture and qPCR results might be of little concern. While a continuous correlation would be preferable, the categorical approach presents an option for evaluation when too few samples are available to create a reliable continuous correlative assessment. In the absence of an alternative to the present EPA TSM, the only beaches likely to be approved for qPCR application are those that are so frequently above management thresholds that the need for rapid methods is limited, while managers would be prevented from taking advantage of qPCR methods at the infrequently contaminated beaches where it is most needed.

Acknowledgements

Staff from many agencies contributed to the study: Ventura County Public Health Lab, City of Los Angeles, Los Angeles County Sanitation District, Orange County Sanitation District, Orange County Public Health Lab, Southern California Coastal Water Research Project Authority, City of Oceanside, City of Encinitas, City of San Diego, Weston Solutions, National Oceanic and Atmosphere Administration.

References

Cao, Y., J.F. Griffith, S. Dorevitch, and S.B. Weisberg. 2012. Effectiveness of qPCR permutations, internal controls and dilution as means for minimizing the impact of inhibition while measuring *Enterococcus* in environmental waters. Journal of Applied Microbiology 113(1), 66-75.

Cao, Y., M.R. Raith and J.F. Griffith. 2015. Droplet digital PCR for simultaneous quantification of general and human-associated fecal indicators for water quality assessment. Water Research 70, 337-349.

Cao, Y., M. Sivaganesan, J. Kinzelman, A.D. Blackwood, R.T. Noble, R.A. Haugland, J.F. Griffith and S.B. Weisberg. 2013. Effect of platform, reference material, and quantification model on enumeration of Enterococcus by quantitative PCR methods. Water Research 47(1), 233-241.

Colford, J.M., K.C. Schiff, J.F. Griffith, V. Yau, B.F. Arnold, C.C. Wright, J.S. Gruber, T.J. Wade, S. Burns, J. Hayes, C. McGee, M. Gold, Y. Cao, R.T. Noble, R. Haugland and S.B. Weisberg. 2012. Using rapid indicators for *Enterococcus* to assess the risk of illness after exposure to urban runoff contaminated marine water. Water Research 46(7), 2176-2186.

Converse, R.R., J.F. Griffith, R.T. Noble, R.A. Haugland, K.C. Schiff and S.B. Weisberg. 2012. Correlation between quantitative polymerase chain reaction and culture-based methods for measuring *Enterococcus* over various temporal scales and three California marine beaches. Applied & Environmental Microbiology 78(4), 1237-1242.

Ebentier, D.L., K.T. Hanley, Y. Cao, B. Badgley, A. Boehm, J. Ervin, K.D. Goodwin, M. Gourmelon, J. Griffith, P. Holden, C.A. Kelty, S. Lozach, C. McGee, L. Peed, M. Raith, M.J. Sadowsky, E. Scott, J. Santodomingo, C. Sinigalliano, O.C. Shanks, L.C.V.D. Werfhorst, D. Wang, S. Wuertz and J. Jay. 2013. Evaluation of the repeatability and reproducibility of a suite of PCR-based microbial source tracking methods. Water Research 47(18), 6839-6848.

Ferretti, J.A., H.V. Tran, S.J. Peterson and V. Loftin. 2013. Rapid method demonstration project at four New Jersey marine beaches using real time quantitative Polymerase Chain Reaction (qPCR). Marine Pollution Bulletin 71(1), 51-63.

Griffith, J.F. and S.B. Weisberg. 2011. Challenges in implementing new technology for beach water quality monitoring: Lessons from a California demonstration project. Marine Technology Society Journal 45, 65-73.

Haugland, R.A., S. Siefring, M. Varma, K.H. Oshima, M. Sivaganesan, Y. Cao, M. Raith, J. Griffith, S.B. Weisberg, R.T. Noble, A.D. Blackwood, J. Kinzelman, T. Anan'eva, R.N. Bushon, E.A. Stelzer, V.J. Harwood, K.V. Gordon and C. Sinigalliano. 2016. Multi-laboratory survey of qPCR enterococci analysis method performance in U.S. Coastal and inland surface waters. Journal of Microbiological Methods.

Haugland, R.A., S.C. Siefring, L.J. Wymer, K.P. Brenner and A.P. Dufour. 2005. Comparison of *Enterococcus* measurements in freshwater at two recreational beaches by quantitative polymerase chain reaction and membrane filter culture analysis. Water Research 39, 559–568.

Hindson, C.M., J.R. Chevillet, H.A. Briggs, E.N. Gallichotte, I.K. Ruf, B.J. Hindson, R.L. Vessella and M. Tewari. 2013. Absolute quantification by droplet digital PCR versus analog real-time PCR. Nat Methods 10(10), 1003-1005.

Kim, J.H. and S.B. Grant. 2004. Public Mis-Notification of Coastal Water Quality: A Probabilistic Evaluation of Posting Errors at Huntington Beach, California. Environmental Science & Technology 38(9), 2497-2504.

Lavender, J.S. and J.L. Kinzelman. 2009. A cross comparison of QPCR to agar-based or defined substrate test methods for the determination of Escherichia coli and enterococci in municipal water quality monitoring programs. Water Research 43(19), 4967-4979.

Leecaster, M.K. and S.B. Weisberg. 2001. Effect of sampling frequency on shoreline microbiology assessment. Marine Pollution Bulletin 42(11), 1150-1154.

Nevers, M.B., M.N. Byappanahalli and R.L. Whitman. 2013. Choices in recreational water quality monitoring: new opportunities and health risk trade-offs. Environmental Science and Technology 47(7), 3073-3081.

Noble, R.T., A.D. Blackwood, J.F. Griffith, C.D. McGee, and S.B. Weisberg. 2010. Comparison of rapid QPCR-based and conventional culture-based methods for enumeration of *Enterococcus* spp. and *Escherichia coli* in recreational waters. Applied and Environmental Microbiology 76(22), 7437-7443.

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.

Raith, M.R., D.L. Ebentier, Y Cao, J.F. Griffith and S.B. Weisberg. 2013. Factors affecting the relationship between quantitative polymerase chain reaction (qPCR) and culture-based enumeration of Enterococcus in environmental waters. Journal of Applied Microbiology.

Shanks, O.C., M. Sivaganesan, L. Peed, C.A. Kelty, A.D. Blackwood, M.R. Greene, R.T. Noble, R.N. Bushon, E.A. Stelzer, J. Kinzelman, T. Anan'eva, C. Sinigalliano, D. Wanless, J. Griffith, Y. Cao, S. Weisberg, V.J. Harwood, C. Staley, K.H. Oshima, M. Varma and R.A. Haugland, 2012. Inter-laboratory general fecal indicator quantitative real-time PCR methods comparison study. Environmental Science & Technology 46(2), 945–953.

Stewart, J.R., A.B. Boehm, E.A. Dubinsky, T.T. Fong, K.D. Goodwin, J.F. Griffith, R.T.Noble, O.C. Shanks, K. Vijayavel and S.B. Weisberg. 2013. Recommendations following a multi-

laboratory comparison of microbial source tracking methods. Water Research 47(18), 6829-6838.

U.S. EPA 2012a. Method 1611: Enterococci in Water by TaqMan® Quantitative Polymerase Chain Reaction (qPCR) Assay. EPA-821-R-12-008, Office of Water, Washington, DC.

U.S. EPA 2012b. Recreational Water Quality Criteria, Office of Water 820-F-12-058, Washington, DC.

U.S. EPA 2014. Site-specific alternative recreational criteria technical support materials for alternative indicators and methods. EPA-820-R-14-011., Office of Water, Office of Science and Technology, Washington, DC.

Wade, T.J., R.L. Calderon, K.P. Brenner, E. Sams, M. Beach, R. Haugland, L. Wymer and A.P. Dufour. 2008. High sensitivity of children to swim ming-associated gastrointestinal illness: Results using a rapid assay of recreational water quality. Epidemiology 19(3), 375-383.

Wade, T.J., R.L. Calderon, E. Sams, M. Beach, K.P. Brenner, A.H. Williams and A.P. Dufour. 2006. Rapidly measured indicators of recreational water quality are predictive of swimming-associated gastrointestional illness. Environmental Health Perspectives 114(1), 24-28.

Whitman, R.L., Z.F. Ge, M.B. Nevers,A.B. Boehm, E.C. Chern, R.A. Haugland, A.M. Lukasik, M. Molina, K. Przybyla-Kelly, D.A. Shively, E.M. White, R.G. Zepp and M.N. Byappanahalli. 2010. Relationship and variation of qPCR and culturable Enterococci estimates in ambient surface waters are predictable. Environmental Science & Technology 44(13), 5049-5054.

Yau, V.M., K.C. Schiff, B.F. Arnold, J.F. Griffith, J.S. Gruber, C.C. Wright, T.J. Wade, S. Burns, J.M. Hayes, C. McGee, M. Gold, Y. Cao, A.B. Boehm, S.B. Weisberg and J.M. Colford, Jr. 2014. Effect of submarine groundwater discharge on bacterial indicators and swimmer health at Avalon Beach, CA, USA. Water Research 59, 23-36.

Chapter 2. Supplementary Information - Assessment of Enterococcus qPCR methods applicability at thirty-six southern California beaches

Discussion

Investigation on lack of agreement between culture and qPCR methods

There was insufficient agreement between qPCR and culture results following the TSM assessment procedures for application of qPCR at most southern California beaches. One potential reason for the finding is that qPCR is a relatively new technology for the seven labs in this study and it is possible that their execution of the qPCR methods was ineffective. Additionally, inherent qPCR method variability may also contribute to the lack of qPCR-culture agreement.

To test these hypotheses, filters from eight summer sites that had sufficient amounts of data to allow agreement assessments following the TSM were reanalyzed by SCCWRP (a research lab highly experienced in qPCR and digital PCR) using qPCR following the exact same qPCR protocol (as described in the Methods section) and digital PCR (shown to be more accurate and reproducible than qPCR) following protocols previously described (Cao et al. 2015, Cao et al. 2016). The level of agreement between new qPCR data (generated by SCCWRP) and culture, and between digital PCR and culture were then assessed following TSM procedures, and compared to the level of agreement between the original qPCR and culture results.

High proficiency of the research lab was confirmed by the qPCR data quality. The qPCR standard curve showed high lab proficiency (amplification efficiency=1.00 with R^2=0.99). This high proficiency was also reflected in low Cq standard deviation, leading to a relatively high LLOQ (Cq=33.9), which effectively eliminated all sites from being eligible for TSM analysis due to the limited number of samples (<30 samples) with qPCR values within ROQ. To proceed with the comparative analysis, a relaxed data inclusion rule was applied letting all DNQ (i.e. detected but below LLOQ) be counted towards the minimum requirement of 30 paired qPCR-culture samples. Note that while qPCR used three quantification models (reg, ΔCt, $\Delta\Delta$Ct as described in main text) to convert *Enterococcus* 23S copies per 100ml to *Enterococcus* cell equivalents per 100ml, ddPCR used an approach similar to that used in qPCR reg by assuming each *Enterococcus* cell contains 4 copies of the targeted 23S gene.

However, the new qPCR data showed no increase in agreement with culture results compared to the original qPCR results produced by less experienced labs (Supplementary Information, Table S4). The ddPCR data also did not show substantially increased agreement with culture results compared to qPCR vs. culture (Supplementary Information, Table S4). Two sites (Cudahy, Visitor's Center) and one site (Cudahy) met IA and R^2 criteria, respectively, for SCCWRP ddPCR-culture, while zero and one site (Cudahy) met IA and R^2 criteria, respectively, for SCCWRP qPCR-culture (Supplementary Information Table S4). Results indicate that lab execution of the qPCR method and its method variability are unlikely reasons for the lack of agreement between qPCR and culture results.

References

Cao, Y., Raith, M.R. and Griffith, J.F. 2015. Droplet digital PCR for simultaneous quantification of general and human-associated fecal indicators for water quality assessment. Water Research 70, 337-349.

Cao, Y., Raith, M.R. and Griffith, J.F. 2016. A duplex digital PCR assay for simultaneous quantification of the *Enterococcus* spp. and the human fecal-associated HF183 marker in waters. Journal of Visualized Experiment e53611, doi:10.3791/53611.

Hindson, C.M., Chevillet, J.R., Briggs, H.A., Gallichotte, E.N., Ruf, I.K., Hindson, B.J., Vessella, R.L. and Tewari, M. 2013. Absolute quantification by droplet digital PCR versus analog real-time PCR. Nat Methods 10(10), 1003-1005.

Raith, M.R., Ebentier, D.L., Cao, Y., Griffith, J.F. and Weisberg, S.B. 2013. Factors affecting the relationship between quantitative polymerase chain reaction (qPCR) and culture-based enumeration of *Enterococcus* in environmental waters. Journal of Applied Microbiology.

Supplementary Tables

Table S1. The *Enterococcus* qPCR standard curve parameters and lower limit of quantification in Cq. Lab H (SCCWRP) standard curves were generated during post hoc analysis of the replicate filters from a subset of sites (8 sites). LLOQ: lower limit of quantification.

Lab	n	Slope	Y intercept	R^2	E	LLOQ in Cq
A	311	-3.40	40.0	0.956	0.97	36.3
B	267	-3.47	38.7	0.951	0.94	35.7
C	181	-3.32	37.8	0.940	1.00	34.5
D	438	-3.50	39.0	0.875	0.93	36.7
E	372	-3.39	38.8	0.988	0.97	34.2
F	144	-3.34	35.7	0.812	0.99	33.9
G	95	-3.31	39.4	0.945	1.00	36.3
H	205	-3.34	38.7	0.994	0.99	33.9

Table S2. Index of agreement (IA) and R^2 for assessing culture and qPCR method agreement, for the sites that had sufficient number of samples (n>30) with paired culture and qPCR data within ROQ (qPCR data with matrix interference excluded) in summer sampling. Three qPCR quantification models (ΔCt, ΔΔCt, reg as described in Methods) were used and assessed for agreement with culture results independently. IA and R^2 values in bold indicate sufficient agreement as required by the TSM.

Site	# paired samples	IA			R^2		
		reg	ΔCt	ΔΔCt	reg	ΔCt	ΔΔCt
1st street	37	0.386	0.453	0.456	0.469	0.409	0.428
Baby Beach	40	0.484	0.577	0.570	0.226	0.208	0.178
Bayside Drive	37	0.493	0.653	NA[a]	0.415	0.404	0.339
Cudahy	48	0.553	**0.704**	0.666	**0.675**	**0.655**	0.517
Marina del Rey Beach	34	0.464	0.580	0.559	0.465	0.476	0.433
Ocean Beach	31	0.272	0.319	0.318	0.007	0.001	0.004
Santa Monica	32	0.330	0.384	0.388	0.184	0.057	0.073
Tecolote	42	0.417	0.493	0.489	0.303	0.223	0.242
Tijuana River	40	0.305	0.365	0.364	0.105	0.092	0.104
Visitor's Center	41	0.405	0.562	0.543	0.311	0.307	0.299

[a]IA could not be calculated due to missing data.

Table S3. Index of agreement (IA) and R^2 for assessing culture and qPCR method agreement, for the sites that had sufficient number of samples (n>30) with paired culture and qPCR data within ROQ (qPCR data with matrix interference excluded) in winter sampling. Three qPCR quantification models (ΔCt, ΔΔCt, reg as described in Methods) were used and assessed for agreement with culture results independently. IA and R^2 values in bold indicate sufficient agreement as required by the TSM.

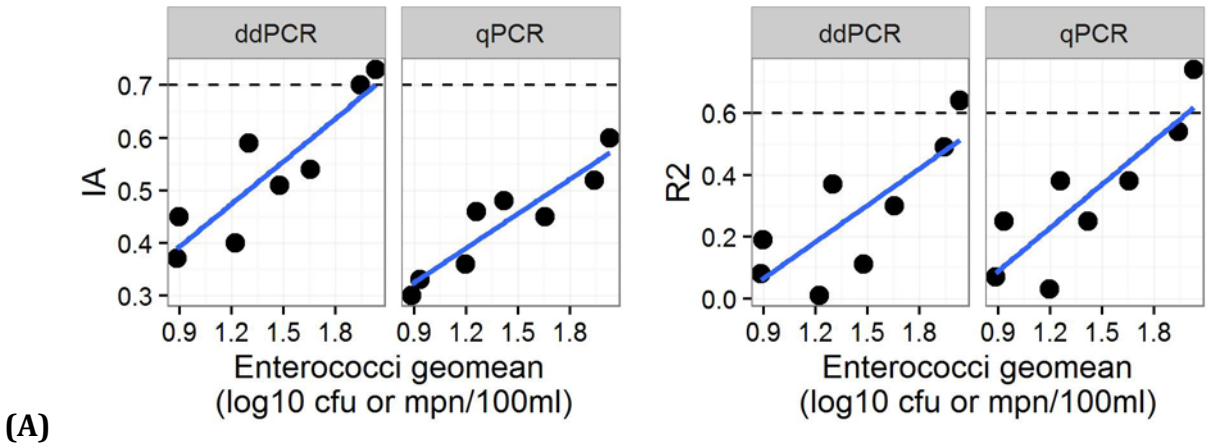
	# paired	IA			R^2		
Site	samples	reg	ΔCt	ΔΔCt	reg	ΔCt	ΔΔCt
Baby Beach	43	0.410	0.474	0.479	0.139	0.076	0.077
Ballona	30	0.470	0.662	NA[a]	**0.752**	**0.702**	**0.689**
Bayside Drive	37	0.421	0.487	0.509	0.379	0.220	0.280
Marina del Rey Beach	49	0.490	0.649	0.639	**0.635**	0.501	0.483
Mother's Beach	41	0.479	0.545	0.543	0.202	0.209	0.198
Newport Dunes	35	0.537	0.643	0.668	0.480	0.396	0.481
North Star Beach	35	0.610	**0.700**	**0.721**	**0.614**	0.561	**0.617**
Santa Monica	41	0.404	0.567	0.574	0.326	0.338	0.350
Sapphire Avenue	31	0.491	0.589	0.613	0.406	0.318	0.404

[a]IA could not be calculated due to missing data.

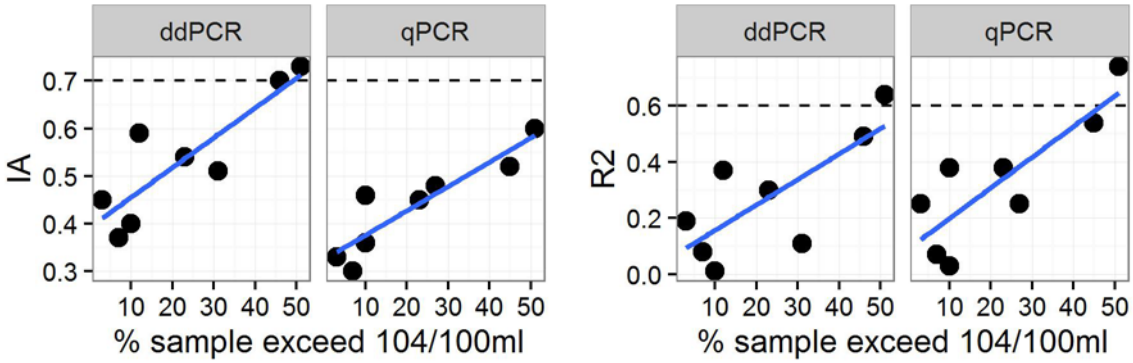
Table S4. Comparison of number of samples (n.spl) with paired culture-qPCR data within range of quantification (qPCR data with matrix interference excluded), index of agreement (IA) and R^2 for assessing culture and qPCR method agreement, among SCCWRP qPCR-culture, SCCWRP ddPCR-culture, and Lab qPCR-culture, for the eight summer sites selected for reanalysis. Note that the same culture results were used in all pairs of culture-qPCR and culture-ddPCR assessments. Only qPCR quantification model 1 (reg, as described in Methods in the main text) is shown here, but the other two qPCR quantification models (ΔCt, ΔΔCt) show a similar trend. SCCWRP qPCR and ddPCR data refer to new data generated by SCCWRP which is a research lab highly experienced in qPCR and digital PCR analysis, while Lab qPCR refers to the original qPCR data generated by the participating labs that have relatively less experience with qPCR. IA and R2 values in bold indicate sufficient agreement as required by the TSM.

	Lab qPCR			SCCWRP qPCR			SCCWRP ddPCR		
Site	n.spl	IA	R^2	n.spl	IA	R^2	n.spl	IA	R^2
Baby Beach	40	0.48	0.23	33	0.48	0.25	39	0.51	0.11
Bayside Drive	37	0.49	0.41	39	0.46	0.38	41	0.59	0.37
Cudahy	48	0.55	**0.67**	53	0.60	**0.74**	53	**0.73**	**0.64**
Marina del Rey Beach	34	0.46	0.46	43	0.45	0.38	43	0.54	0.3
Ocean Beach	31	0.27	0.01	31	0.33	0.25	36	0.45	0.19
Tecolote	42	0.42	0.3	51	0.36	0.03	52	0.40	0.01
Tijuana River	40	0.31	0.11	41	0.30	0.07	41	0.37	0.08
Visitor's Center	41	0.41	0.31	49	0.52	0.54	52	**0.70**	0.49

Supplementary Figures



(A)



(B)

Figure S1. Relationships of agreement metrics (IA: index of agreement, R^2) with A) enterococci concentration or with B) frequency of enterococci concentration exceeding California Ocean Plan single sample maximum of 104/100ml, for the subset of eight sites where SCCWRP reanalyzed the replicate filters. All relationships are significant (p-values<0.05). Horizontal dashed lines indicate thresholds IA and R^2 values for sufficient culture-qPCR agreement as set in the TSM. Blue lines indicate the regression lines.

CHAPTER 3. REGIONAL ASSESSMENT OF HUMAN FECAL CONTAMINATION IN SOUTHERN CALIFORNIA COASTAL DRAINAGES

Introduction

Recreational water quality is routinely monitored using fecal indicator bacteria (FIB), such as *Enterococcus* spp. and *E. coli*, as proxies for fecal contamination, because they can be measured cheaper and faster than pathogens (Field and Samadpour 2007). Water bodies with FIB concentrations exceeding recreational water quality criteria (U.S. EPA 2012b) are treated as a public health risk, and management actions such as beach advisories and pollution remediation are typically implemented in response.

However, FIB measurements are not diagnostic of whether fecal contamination originates from human, animal, or non-fecal sources, which is important for two reasons. First, understanding the sources of fecal contamination allows managers to more appropriately target remediation actions (Field and Samadpour 2007, Griffith et al. 2013, Stoeckel and Harwood 2007, U.S. EPA 2005). Second, human fecal material is generally considered a greater public health risk than non-human fecal material (Soller et al. 2010), making it appropriate to prioritize sites for remediation based on the extent of human fecal contamination. In recognition of different risk posed by different sources, the U.S. Environmental Protection Agency has developed a quantitative microbial risk assessment (QMRA) process for defining alternative management strategies for beaches that have high FIB counts, but with a corresponding low-level of human fecal contamination (U.S. EPA 2012b).

Host-associated genetic markers that allow for fecal source identification and that are sensitive and specific to their target hosts are available (Ahmed et al. 2016, Boehm et al. 2013, Green et al. 2014, Layton et al. 2013). These markers have been used extensively as diagnostic tools to discern fecal sources within watersheds (Ahmed et al. 2015, Ervin et al. 2014, Griffith et al. 2013), but have not been used as a monitoring tool to prioritize the need for remediation among watershed systems on a regional scale. Cao et al. (2013) identified several challenges in prioritization, such as determining whether frequency of human marker occurrence or magnitude of the signal are more important to the outcome. Here we present a study in which we conduct a regional evaluation of human marker prevalence among drainages that discharge to the Southern California ocean, evaluate the sensitivity of site rankings to some of the decisions outlined by Cao et al., and investigate how those relationships change between periods with, and without, rainfall.

Methods

Sites and sampling

Approximately fifty water samples were collected from twenty-two southern California coastal drainages which included creeks, rivers, and storm drains (Figure 1, higher resolution interactive map at http://bit.ly/2sxLHcI). Drainages were selected largely based on frequent historical *Enterococcus* exceedances at nearby swimming beaches to which they were important source water. Samples were collected at approximately weekly intervals under summer dry-weather conditions between 2013 and 2015. A target of 50 wet weather samples were also collected from

23 drainages year-round between June 26, 2013 to May 26, 2016, though the number of wet weather samples varied among drainages because of differences in rainfall patterns and available effort in some locations to respond to rainfall events.

Figure 1. Sampling sites. Majority of the sites were sampled in both storm and summer dry conditions. Higher resolution interactive map, including site photos, is available at http://bit.ly/2sxLHcl.

A wet weather event was defined as at least 0.10” rainfall at the closest rain gauge to the sampling site following an antecedent dry period of three or more days (Griffith et al. 2010). While wet weather events are generally confined to the fall and winter seasons in southern California, about 17% of the wet weather samples were collected during summer.

Dry weather samples were collected following the routine sampling schedule (generally weekly) for each site. Samples were collected in the early morning to mirror current monitoring procedures and limit degradation of the bacterial signal due to sunlight exposure. Wet weather samples were collected as soon as possible following the first 0.10” of rain, but no later than 72 hours after initiation of the rain event. All samples were taken inside the drainage, upstream of

tidal influences and any onsite disinfection facilities (e.g., UV or ozonation, if present), and transported on ice to the laboratory for processing.

Sample processing and qPCR analysis

Samples from all but four sites under dry weather, and for all but three sites under wet weather, were tested for cultivable Enterococcus by EPA Method 1600 or Enterolert (IDEXX, Westbrook, ME). In addition, samples were filtered (0.45µM polycarbonate filters, 100 ml per filter or until clogging) within 6 hours of sample collection following standard protocol (U.S. EPA 2012a). Filters were stored at -80°C until DNA extraction using the GeneRite DNA EZ Extraction kit (GeneRite, North Brunswick, NJ), followed by analysis for the HF183 human-associated fecal marker by the duplex HF183/BacR287 qPCR assay as described elsewhere (Green et al. 2014). Salmon testes DNA was spiked into the lysis buffer during DNA extraction as a sample processing control and measured by the sketa22 qPCR assay (Green et al. 2014, U.S. EPA 2012a).

Eight laboratories processed samples following the same standard operating protocols for qPCR, including identical qPCR 96-well plate setups, with precise positions indicated for standard curves, negative controls, and samples. All samples, standards, and controls were run with triplicate qPCR reactions. Each HF183 qPCR plate contained one standard curve (6-point, 10-fold dilution, 1×106 to 1 copy per reaction), two negative extraction controls (one for each batch of DNA extractions consisting of 11 samples), one no-template control, and 22 environmental samples. All qPCR standards were prepared (Green et al. 2014) into single use aliquots at a central facility (Southern California Coastal Water Research Project Authority), and stored at -80°C before use at the individual laboratories.

Data analysis

Master standard curves [cycle of quantification (Cq) vs. $\log_{10}$ concentration] were calculated for each laboratory using regression with an outlier removal procedure (a data point was deemed an outlier if its studentized residual was greater than 3), and the resulting regression equations were used for HF183 qPCR quantification as described elsewhere (Ebentier et al. 2013). The limit of detection (LOD) was set at the lowest concentration on the standard curve (1 copy per reaction, >90% replicates amplified at all labs) and expressed in Cq for each lab (Stewart et al. 2013). The lower limit of quantification (LLOQ) was set at the lowest concentration where all qPCR replicates amplified in all labs (10 copies per reaction) and expressed as Cq + 2×Standard deviation for each lab (Stewart et al. 2013).

Frequency of HF183 positive and site average HF183 concentration were calculated as metrics of human fecal contamination at the sites. Samples were defined as positives for frequency calculation if any of the qPCR replicates amplified. Site average concentration of HF183 was calculated as the arithmetic mean at $\log_{10}$ scale (i.e., the geometric mean at the normal scale) of all qPCR replicates from all samples at each site. In calculating the means, non-detect (ND) and detection below LOD (DBLOD) (Stewart et al. 2013) were not assigned a zero value but instead substituted with a more appropriate statistical estimate based on Poisson distribution (method detail in Supplementary Information).

Results

Sampling and results summary

1013 and 627 samples from dry and wet weather, respectively, were analyzed for the HF183 human fecal marker. Number of samples per site ranged from 15 to 54 under dry weather, and from 3 to 50 under wet weather (Supplementary Information, Table S1). Sites (none in dry weather, but five in wet weather) with ten or fewer samples were excluded from analyses.

Standard curves showed satisfactory performance across all labs, with R^2 and amplification efficiency ranging from 0.95 to 0.99 and from 0.89 to 0.99, respectively (Supplementary Information, Table S2). Limit of detection and lower limit of quantification were 1 and 10 gene copies per reaction, respectively, corresponding to 79 and 789 copies per 100ml, respectively. Negative controls showed little sign of cross-contamination. Among the 936 no template control reactions, only 2 reactions showed a low level of amplification (Cq>37). Among the 1179 negative extraction control reactions, only 2 amplified (Cq>36). No drainage water samples showed signs of inhibition or sample processing failure.

The HF183 human fecal marker was detected at all except two sites in dry weather and at all sites in wet weather. Overall, HF183 was detected in 21% and 52% of samples in dry and wet weather, respectively. Among the detections, 22% and 44% in dry and wet weather, respectively, were at high enough concentrations to be quantifiable (i.e. above LLOQ). The median HF183 concentration was below LOD in dry weather and at LOD in wet weather. Nevertheless, the highest concentrations detected were $1.5x10^7$ and $3.2x10^6$ copies per 100ml in dry and wet weather, respectively. Summaries by site are provided as (Supplementary Information, Figures S1, S2).

Across all sites, the average *Enterococcus* concentration (geomean) was 237 and 1265 per 100ml in dry and wet weather, respectively. *Enterococcus* concentration in drainages discharging to the ocean exceeded California's single sample maximum (SSM) of 104/100ml for ocean samples in 67% and 88% of samples, and was 100 times higher than the SSM in 3% and 19% of samples, in dry and wet weather, respectively (Supplementary Information, Table S1).

Site prioritization

Site prioritization was conducted using both frequency and site average concentration of HF183 detection, in both dry and wet weathers. There were clear differences in the occurrence of the HF183 marker across southern California creeks (Figure 2). Among the 22 sites in dry weather, the marker was not detected in any sample at two sites (Santa Ana River, Topanga Creek), but was detected in 100% of samples at one site (Malaga Cove East). The frequency of HF183 detection gradually increased over fifteen sites from 7% to 30%, then jumped to 48% to 100% for five sites. Site average concentration, ranging from 1 to 154 copies per 100ml, also showed large site differences (Supplementary Information, Figure S2).

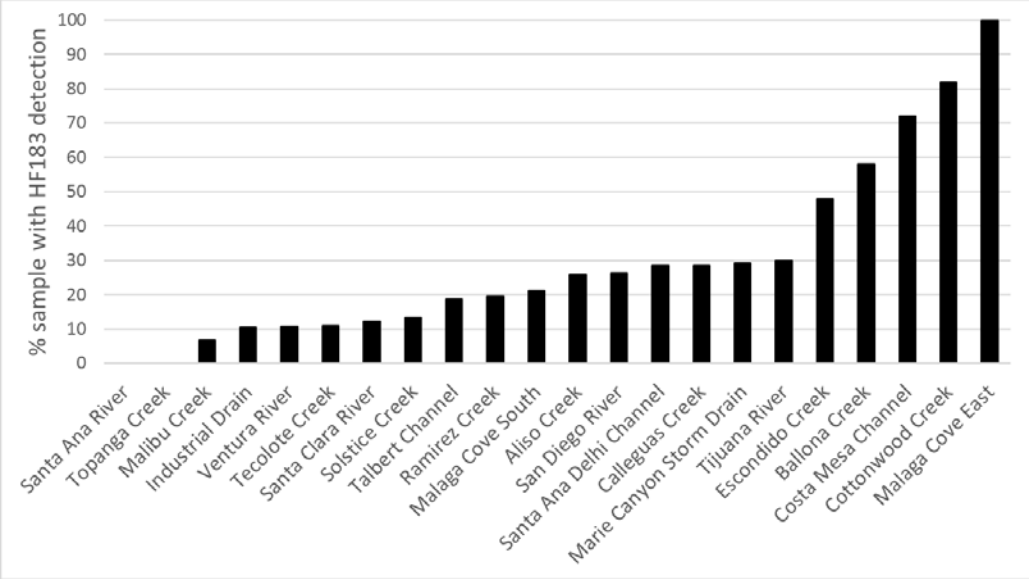


Figure 2. Frequency of HF183 detection at the 22 sites in summer dry weather. Frequency of HF183 detection is defined as % samples with HF183 detection, and HF183 is said detected in a sample if HF183 is detected in any of the three qPCR replicates.

The ranking of drainages was relatively consistent regardless of whether the ranking was based on frequency of HF183 positives or average HF183 concentration (Figure 3, Table S3). Among the 22 dry weather sites, only five sites received different positions in the ranking based on frequency vs. concentration, among which two sites differed only by one position. All five sites were also among middle positions in ranking by either metric.

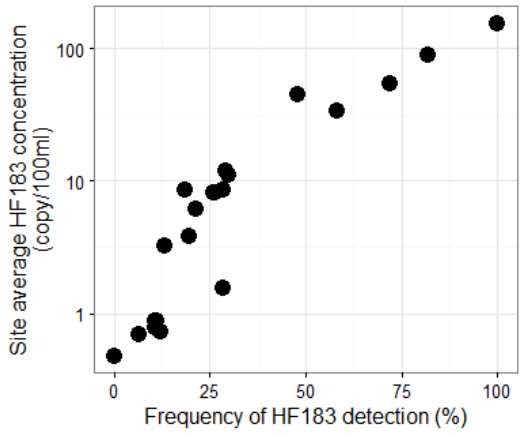


Figure 3. Site average HF183 concentration versus frequency of HF183 positive in summer dry. Frequency of HF183 positive is defined as % samples that are positive for HF183, and a sample is considered positive for HF183 if the marker is amplified in any of the three qPCR replicates. Site average concentration is calculated by the Poisson approach as described in the Method section.

Site ranking during wet weather did not correlate well with site ranking during dry weather (Figure 4, Supplementary Information, Figure S3 and Table S4). While the extent of human fecal contamination generally increased (except for 3 and 4 sites based on frequency and concentration, respectively, Figure 4 and Supplementary Information, Figure S3) from dry to wet weather, the relative extent of increase was dissimilar across sites. Among the 16 sites that were sampled in both dry and wet weather, no site had the same rank in dry as rank in wet, regardless of whether the ranking was based on frequency of HF183 positives or average site HF183 concentration (Supplementary Information, Table S4). Marie Canyon Storm Drain showed the biggest discrepancy, shifting 10 and 11 positions in ranking, based on frequency and concentration, respectively, between dry and wet.

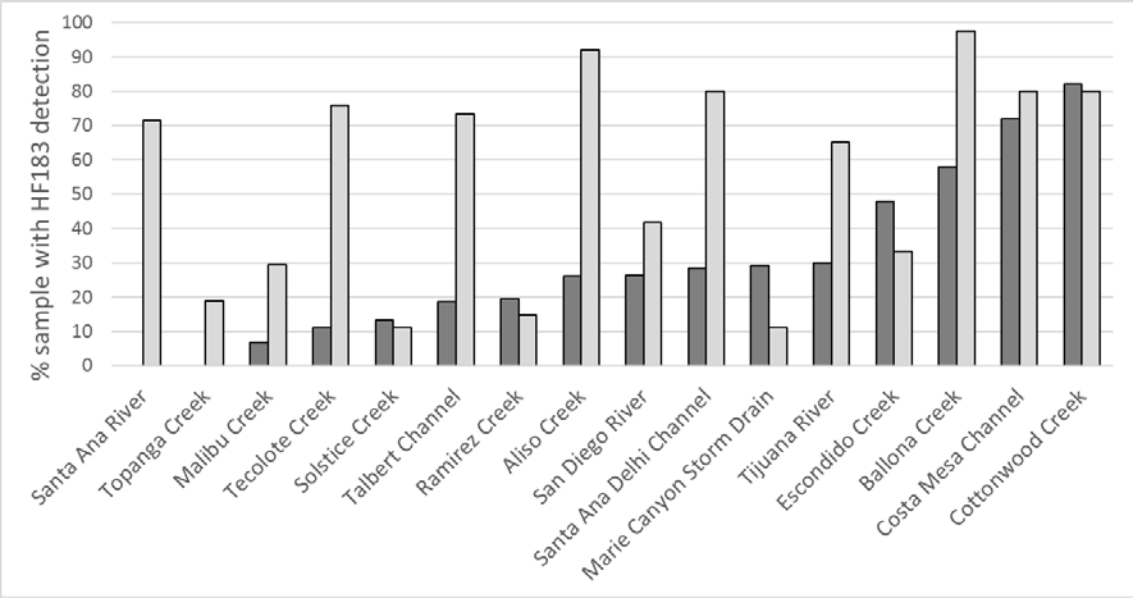


Figure 4. Frequency of HF183 detection by site in wet (light grey filled bars) versus dry (dark-filled) weather conditions. Frequency of HF183 detection is defined as % samples with HF183 detection, and HF183 is said detected in a sample if HF183 is detected in any of the three qPCR replicates. Sites are sorted from left to right by frequency of detection under dry weather conditions.

Discussion

Our study corroborates the ubiquitous presence of human fecal pollution in highly urbanized environments (Sauer et al. 2011, Sercu et al. 2009, Sidhu et al. 2013) and is also consistent with the finding of Ahmed et al. (2016) that human fecal pollution levels vary considerably across sites. These site differences don't appear to be related to simple watershed characteristics, as we observed no discernable relationship between watershed land use and extent of human fecal contamination in the drainage (Supplementary Information, Figure S5). This suggests a complexity of human fecal pollution in urban environments that relates to specific management practices and illustrates the need to conduct site-specific investigations (Sauer et al. 2011).

We also observed that the extent of human fecal pollution in drainage systems expands considerably during wet weather events, a finding consistent with previous studies (Templar et al. 2016). This is likely caused by increased human fecal input into the drainages during wet weather perhaps due to overland flow and/or subsurface sewage exfiltration. Depending on storm size and land infiltration capacity in the watershed, storms generate overland flow that brings surrounding fecal pollution into drainages (Sercu et al. 2011b). A raised groundwater table or higher subsurface soil water content can also increase sewage exfiltration as a source of human fecal pollution in drainage systems (Sercu et al. 2011a). However, a few sites showed lower frequency and concentration of HF183 in wet than in dry weather, which might be attributed to simple dilution by the large volume of storm water. Regardless, the changes in human fecal signature during rainfall certainly argue for a difference in management remediation strategies between wet and dry periods.

Site prioritization

One of the study goals was to rank sites in southern California based on the extent of human fecal contamination as a means of prioritizing remediation actions. To do so, several choices must be made, one of which is whether to base site ranking on frequency of detection or concentration of HF183. This choice did not influence ranking appreciably, probably because most sites with persistent human fecal input would be likely to have higher concentration as well, and the site average concentration was calculated by partially integrating the frequency information (i.e., via the Poisson mean substitution of ND and DBLOD). While it did not make a big difference in this study, we chose to focus on frequency, in absence of a standardized metric, for site ranking primarily because it is less sensitive to site-specific degradation. Nevertheless, we recognize that frequency of detection is less, compared to concentration of HF183, related to characterizing total human fecal contribution (i.e. mass) to FIB that are needed for risk assessment (Soller et al. 2014).

A second choice to make relates to implementation mechanics. That is, how to define a positive sample for calculating frequency. We chose to count any HF183 detection, even if that detection was unquantifiable, erring on the side of being more protective of public health. This is because, due to subsampling effect in performing the qPCR method and the wide range of degradation rate of the HF183 marker as observed in the environment, a low HF183 signal could still present a meaningful public health risk (see simulation in Supplementary Information, Figure S6). However, we also tried six other definitions of a positive sample to calculate frequency, deeming a sample positive for HF183 if at least one, or at least two, or all three qPCR replicates amplified, if at least one, or at least two, or all three qPCR replicates were DNQ or quantifiable,

or if the sample average concentration (ND and DBLOD were substituted with ½ or 1 limit of detection) was greater than limit of detection. For none of these alternative definitions did the rank order of sites change appreciably (Supplementary Information, Table S3), and the alternative frequency calculations were highly correlated (p-value<0.001 for all pair-wise correlation).

A third choice in developing site rankings is whether to use HF183 alone or in combination with *Enterococcus*, with the rationale that a HF183 signal may be of lesser concern when the *Enterococcus* levels are low. This decision made a substantial difference, as the correlation between ranking by HF183 and ranking by *Enterococcus* was low (Figure 5, Supplementary Information, Figure S7). The low correlation probably results because *Enterococcus* is a general indicator that originates from human, non-human (Ervin et al. 2013) and even non-fecal sources (Byappanahalli et al. 2012, Ferguson et al. 2016, Imamura et al. 2011, Yamahara et al. 2009) and *Enterococcus* and HF183 should correlate only when human fecal material is the dominant source of contamination (Harwood et al. 2005). As such, we suggest placing higher priority on HF183 than on *Enterococcus* level, with one important exception. HF183 is a molecular measurement and can potentially give a false reading when disinfected water is present in the watershed. It is possible that watersheds where recycled water is used for irrigation, for example on golf courses, could contribute to a high level of HF183 prevalence.

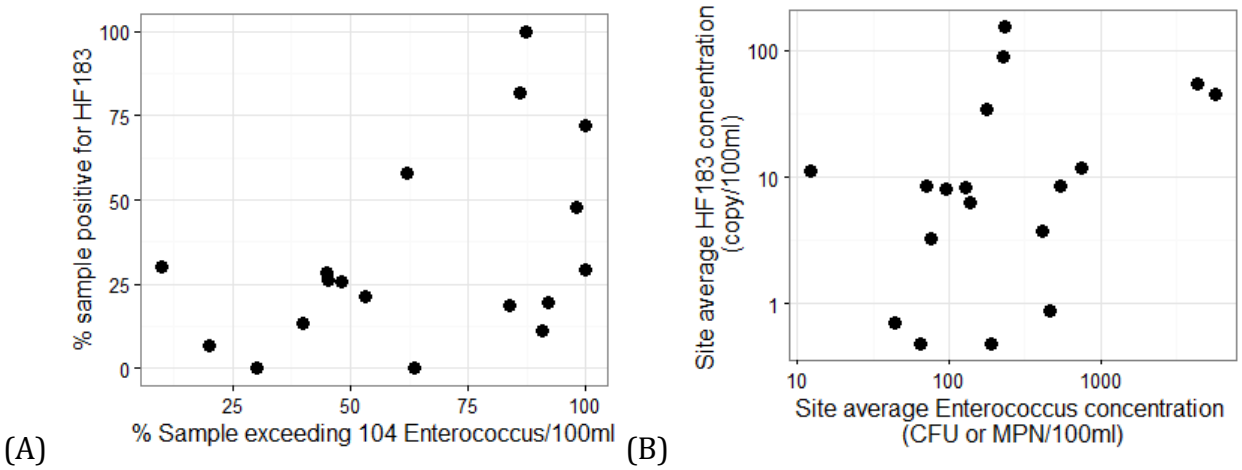


Figure 5. Ranking site by HF183- vs. Enterococcus-based metrics during summer dry weather: (a) Frequency of HF183 positive versus frequency of *Enterococcus* exceedance, (b) Site average HF183 concentration versus site average *Enterococcus* concentration. HF183-based metrics are as defined in Figure 3. Frequency of *Enterococcus* exceedance is defined as % samples with more than 104 *Enterococcus* per 100ml. Site average concentration of *Enterococcus* is defined as geomean at the site.

Implications for water quality management

The utility of the HF183 human fecal marker as a monitoring tool is both valuable and feasible. As FIB do not distinguish human, non-human or non-fecal sources, monitoring and remediation focused on FIB solely may not be cost-effective if the ultimate goal is public health protection. A monitoring framework that incorporates a more precise fecal indicator such as a human fecal marker should provide a higher level of information and greatly improve management decision making such as remediation prioritization and consideration of QMRA as alternative management strategies (this study, (Sauer et al. 2011). Additionally, the HF183 marker assay is becoming widely available for use, with EPA presently developing a nationally standardized analytical procedure (Shanks et al. 2016). Moreover, a single molecular assay that simultaneously quantifies both *Enterococcus* spp. and the HF183 marker has also been validated (Cao et al. 2016a, Cao et al. 2015) with published standard operating procedures and multi-media training materials (Cao et al. 2016b). Our study suggests that routine measurement of both enterococci and the HF183 marker will allow managers to focus remediation efforts on the highest priority sites: those with both high enterococci and HF183 (e.g. upper right corner of Figure5A).

Acknowledgments

Staff from many agencies contributed to the study: Ventura Countywide Stormwater Quality Management Program, Ventura County Public Health Lab, City of Malibu, City of Los Angeles, County of Los Angeles, Los Angeles County Sanitation District, Los Angeles County Flood Control District, Orange County Sanitation District, Orange County Department of Public Works, Orange County Public Health Lab, Southern California Coastal Water Research Project Authority, City of Oceanside, City of Encinitas, City of San Diego, Weston Solution, National Oceanic and Atmosphere Administration.

References

Ahmed, W., Hughes, B. and Harwood, V. (2016) Current Status of Marker Genes of Bacteroides and Related Taxa for Identifying Sewage Pollution in Environmental Waters. Water 8(6), 231.

Ahmed, W., Staley, C., Sadowsky, M.J., Gyawali, P., Sidhu, J.P., Palmer, A., Beale, D.J. and Toze, S. (2015) Toolbox Approaches Using Molecular Markers and 16S rRNA Gene Amplicon Data Sets for Identification of Fecal Pollution in Surface Water. Applied & Environmental Microbiology 81(20), 7067-7077.

Boehm, A.B., Van De Werfhorst, L.C., Griffith, J.F., Holden, P.A., Jay, J.A., Shanks, O.C., Wang, D. and Weisberg, S.B. (2013) Performance of forty-one microbial source tracking methods: A twenty-seven lab evaluation study. Water Research 47(18), 6812-6828.

Byappanahalli, M.N., Roll, B.M. and Fujioka, R.S. (2012) Evidence for Occurrence, Persistence, and Growth Potential of Escherichia coli and Enterococci in Hawaii^|^rsquo;s Soil Environments. Microbes and Environments 27(2), 164-170.

Cao, Y., Griffith, J.F. and Weisberg, S.B. (2016a) Methods in Molecular Biology Series: Marine Genomics. Bourlat, S.J. (ed), Springer, New York, NY.

Cao, Y., Hagedorn, C., Shanks, O.C., Wang, D., Ervin, J., Griffith, J.F., Layton, B.A., McGee, C.D., Riedel, T.E. and Weisberg, S.B. (2013) Towards establishing a human fecal contamination index in microbial source tracking. International Journal of Environmental Science and Engineering Research 4(3), 46-58.

Cao, Y., Raith, M.R. and Griffith, J.F. (2015) Droplet digital PCR for simultaneous quantification of general and human-associated fecal indicators for water quality assessment. Water Research 70, 337-349.

Cao, Y., Raith, M.R. and Griffith, J.F. (2016b) A duplex digital PCR assay for simultaneous quantification of the Enterococcus spp. and the human fecal-associated HF183 marker in waters. Journal of Visualized Experiment e53611, doi:10.3791/53611.

Ebentier, D.L., Hanley, K.T., Cao, Y., Badgley, B., Boehm, A., Ervin, J., Goodwin, K.D., Gourmelon, M., Griffith, J., Holden, P., Kelty, C.A., Lozach, S., McGee, C., Peed, L., Raith, M., Sadowsky, M.J., Scott, E., Santodomingo, J., Sinigalliano, C., Shanks, O.C., Werfhorst, L.C.V.D., Wang, D., Wuertz, S. and Jay, J. (2013) Evaluation of the repeatability and reproducibility of a suite of PCR-based microbial source tracking methods. Water Research 47(18), 6839-6848.

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

Ervin, J.S., Van De Werfhorst, L.C., Murray, J.L. and Holden, P.A. (2014) Microbial source tracking in a coastal California watershed reveals canines as controllable sources of fecal contamination. Environmental Science and Technology 48(16), 9043-9052.

Ferguson, D.M., Weisberg, S.B., Hagedorn, C., De Leon, K., Mofidi, V., Wolfe, J., Zimmerman, M. and Jay, J.A. (2016) Enterococcus growth on eelgrass (Zostera marina); implications for water quality. FEMS Microbiology Ecology 92(4), fiw047.

Field, K.G. and Samadpour, M. (2007) Fecal source tracking, the indicator paradigm, and managing water quality. Water Research 41(16), 3517-3538.

Green, H.C., Haugland, R.A., Varma, M., Millen, H.T., Borchardt, M.A., Field, K.G., Walters, W.A., Knight, R., Sivaganesan, M., Kelty, C.A. and Shanks, O.C. (2014) Improved HF183 quantitative real-time PCR assay for characterization of human fecal pollution in ambient surface water samples. Applied & Environmental Microbiology 80(10), 3086-3094.

Griffith, J.F., Layton, B.A., Boehm, A.B., Holden, P.A., Jay, J.A., Hagedorn, C.D., McGee, C.D. and Weisberg, S.B. (2013) The California Microbial Source Identification Manual: A Tiered Approach to Identifying Fecal Pollution Sources to Beaches. , Southern California Coastal Water Research Project Commission, Costa Mesa, California, USA, Technical Report, 804, 83.

Griffith, J.F., Schiff, K.C., Lyon, G.S. and Fuhrman, J.A. (2010) Microbiological water quality at non-human influenced reference beaches in southern California during wet weather. Marine Pollution Bulletin 60(4), 500-508.

Harwood, V.J., Levine, A.D., Scott, T.M., Chivukula, V., Lukasik, J., Farrah, S.R. and Rose, J.B. (2005) Validity of the indicator organism paradigm for pathogen reduction in reclaimed water and public health protection. Applied and Environmental Microbiology 71(6), 3163-3170.

Imamura, G.J., Thompson, R.S., Boehm, A.B. and Jay, J.A. (2011) Wrack promotes the persistence of fecal indicator bacteria in marine sands and seawater. FEMS Microbiology Ecology 77(1), 40-49.

Layton, B.A., Cao, Y., Ebentier, D.L., Hanley, K., Ballesté, E., Brandão, J., Byappanahalli, M., Converse, R., Farnleitner, A.H., Gentry-Shields, J., Gidley, M.L., Gourmelon, M., Lee, C.S., Lee, J., Lozach, S., Madi, T., Meijer, W.G., Noble, R., Peed, L., Reischer, G.H., Rodrigues, R., Rose, J.B., Schriewer, A., Sinigalliano, C., Srinivasan, S., Stewart, J., Van De Werfhorst, L.C., Wang, D., Whitman, R., Wuertz, S., Jay, J., Holden, P.A., Boehm, A.B., Shanks, O. and Griffith, J.F. (2013) Performance of human fecal anaerobe-associated PCR-based assays in a multi-laboratory method evaluation study. Water Research 47(18), 6897-6908.

Sauer, E.P., Vandewalle, J.L., Bootsma, M.J. and McLellan, S.L. (2011) Detection of the human specific Bacteroides genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment. Water Research 45(14), 4081-4091.

Sercu, B., Van De Werfhorst, L.C., Murray, J. and Holden, P.A. (2009) Storm drains are sources of human fecal pollution during dry weather in three urban southern California watersheds. Environmental Science & Technology 43(2), 293-298.

Sercu, B., Van De Werfhorst, L.C., Murray, J.L. and Holden, P.A. (2011a) Sewage exfiltration as a source of storm drain contamination during dry weather in urban watersheds. Environmental Science and Technology 45(17), 7151-7157.

Sercu, B., Van De Werfhorst, L.C., Murray, J.L.S. and Holden, P.A. (2011b) Terrestrial Sources Homogenize Bacterial Water Quality During Rainfall in Two Urbanized Watersheds in Santa Barbara, CA. Microbial Ecology 62(3), 574-583.

Shanks, O.C., Kelty, C.A., Oshiro, R., Haugland, R.A., Madi, T., Brooks, L., Field, K.G. and Sivaganesan, M. (2016) Data Acceptance Criteria for Standardized Human-Associated Fecal Source Identification Quantitative Real-Time PCR Methods. Applied & Environmental Microbiology 82(9), 2773-2782.

Sidhu, J.P., Ahmed, W., Gernjak, W., Aryal, R., McCarthy, D., Palmer, A., Kolotelo, P. and Toze, S. (2013) Sewage pollution in urban stormwater runoff as evident from the widespread presence of multiple microbial and chemical source tracking markers. Science of the Total Environment 463-464, 488-496.

Soller, J.A., Schoen, M.E., Bartrand, T., Ravenscroft, J.E. and Ashbolt, N.J. (2010) Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. Water Research 44(16), 4674-4691.

Soller, J.A., Schoen, M.E., Varghese, A., Ichida, A.M., Boehm, A.B., Eftim, S., Ashbolt, N.J. and Ravenscroft, J.E. (2014) Human health risk implications of multiple sources of faecal indicator bacteria in a recreational waterbody. Water Research 66, 254-264.

Stewart, J.R., Boehm, A.B., Dubinsky, E.A., Fong, T.T., Goodwin, K.D., Griffith, J.F., Noble, R.T., Shanks, O.C., Vijayavel, K. and Weisberg, S.B. (2013) Recommendations following a multi-laboratory comparison of microbial source tracking methods. Water Research 47(18), 6829-6838.

Stoeckel, D.M. and Harwood, V.J. (2007) Performance, design, and analysis in microbial source tracking studies. Applied & Environmental Microbiology 73(8), 2405-2415.

Templar, H.A., Dila, D.K., Bootsma, M.J., Corsi, S.R. and McLellan, S.L. (2016) Quantification of human-associated fecal indicators reveal sewage from urban watersheds as a source of pollution to Lake Michigan. Water Research 100, 556-567.

U.S. EPA (2012a) Method 1611: Enterococci in Water by TaqMan® Quantitative Polymerase Chain Reaction (qPCR) Assay. EPA-821-R-12-008, Office of Water, Washington, DC.

U.S. EPA (2005) Microbial Source Tracking Guide Document, p. 150, Office of Research and Development, Cincinnatti, OH.

U.S. EPA (2012b) Recreational Water Quality Criteria, Office of Water 820-F-12-058, Washington, DC.

Yamahara, K.M., Walters, S.P. and Boehm, A.B. (2009) Growth of *Enterococc*i in unaltered, unseeded beach sands subjected to tidal wetting. Applied and Environmental Microbiology 75(6), 1517-1524.

Chapter 3. Supplementary Information - Regional assessment of human fecal contamination in southern California coastal drainages

Methods

Calculating site average HF183 concentration

Site average concentration of HF183 was calculated as the arithmetic mean at $\log_{10}$ scale (i.e. the geometric mean at the normal scale) of all qPCR replicates from all samples at each site. Depending on the target concentration, qPCR results were grouped into four categories: ND (non-detect, i.e. no amplification), DBLOD (detected below limit of detection), DNQ (detected but not quantifiable), and quantifiable (Stewart et al. 2013). In calculating the means, ND and DBLOD were considered censored data, and DNQ and quantifiable data were used as is. Two site average concentrations were calculated, differing by how the censored data were treated before averaging. Our first approach substituted ND and DBLOD with ½ or 1 limit of detection. Our second approach is more statistically based and substituted ND and DBLOD with a Poisson mean estimated by regarding ND as negative and DBLOD as positive. The percentage of positives (p) was then used to estimate a Poisson mean $\lambda = -\ln(1-p)$ representing the group of ND and DBLOD results.

Other common options include substituting censored data with zero or limit of detection which has been heavily criticized (Helsel 2012). More statistically based approaches (such as those borrowed from survival analysis) unfortunately have limited application, as the prerequisite of having a reasonable portion of uncensored data is often not met at many locations with environmental HF183 data. Therefore, they were not reported here.

Results

HF183 results in wet weather

Sixteen sites were sampled in both dry and wet weather, and a total of 18 sites sampled during wet weather had more than 10 samples. Large differences in extent of prevalence of the HF183 marker was observed across the 18 sites (Figure S4). The HF183 marker was detected in 11-97% of samples, and site average HF183 concentration ranged from 2-7551 copies per 100ml.

Discussion

Characterization of watershed land use pattern

To discern potential drivers to the observed large site differences in level of human fecal pollution, watershed land use patterns were correlated to the extent of human fecal contamination in drainages. The land use patterns were characterized by arranging land use categories (provided by responsible agencies for the studied watersheds based on analysis of GIS land use profiles) into four bins roughly based on the density of human occupancy, as follows. High human density included single family and multi-family residential areas; Medium human density included commercial, institutional, municipal, office areas; Low human density included areas with education industrial, and other similar facilities; Minimal human density included agriculture, open space, parks, transportation, vacant, water, and similar areas. The percentage area in each bin (and with bins of high and medium human density combined) was then correlated to frequency of HF183 detection and site average concentration of HF183. No significant correlation was observed (Figure S5). Although the four sites with highest frequency and average concentration of HF183 detection had more than 50% of the watershed characterized as having relatively high human occupancy (commercial and/or residential), other watersheds with similar land use had much less human fecal contamination (<30% HF183 detection).

Rationale for the definition of a HF183 positive sample

To calculate frequency of HF183 detection, a sample was deemed positive (for HF183 detection) if this marker is amplified in any of the three qPCR replicates. This recommendation is based on two lines of reasoning: 1) A low HF183 signal may still indicate meaningful risk; and 2) Not all qPCR replicates will amplify when signals are low.

In diluted fresh sewage, 4200 copies HF183 per 100ml corresponds to a U.S. EPA benchmark illness rate of public health significance (30 GI illness per 1000 swimmers) (Boehm et al. 2015). However, environmental waters mostly contain aged sewage with a wide range of reported decay rates (Ahmed et al. 2016). Recent field in-situ experiments in California measured decay rates corresponding to 0.5-3 $\log_{10}$ reduction per day (Cao et al. 2017), leading to reduction of the 4200 copy HF183 per 100ml to below or near the limit of detection of qPCR within 1-3 days. Given that the HF183 marker may decay faster than pathogens in most environmental conditions, low HF183 signals are likely of public health significance.

However, as only a small fraction (<2%) of the original water sample is analyzed based on the HF183 protocol (Green et al. 2014) (as in most qPCR protocols), simulation shows that such signal may only be present in one out of the three qPCR replicates due to subsampling effects (Figure S5). Briefly, generally 100ml of water is filtered to capture bacterial cells on the filter, which is subject to DNA extraction resulting in 100µl DNA extract, from which only 2 µl was analyzed in each qPCR reaction to measure the HF183 marker. A subsampling process therefore occurs when pipetting 2 µl out of the 100 µl total DNA volume, and this process reduces the probability of detection when HF183 concentration is low. While other factors such as DNA extraction efficiency and PCR kinetics further reduce the probability of amplification, a simulation using binomial distribution can be used to demonstrate how much subsampling (of the DNA extracts) alone reduces the probability of amplification in qPCR replicates.

A binomial distribution of B(3, p.rxn) was used to estimate the probability of HF183 amplification in at least one, or at least two, or in all of the three qPCR replicates. The

probability of having at least 1 copy of the HF183 marker present in each qPCR reaction, p.rxn, was calculated by simulating the subsampling process of pipetting 2µl (per qPCR standard operating protocol) from the 100µl DNA extracts (resulting from 100ml of water) into each qPCR reaction. The subsampling process was simulated by B(2, conc) where 2 refers to the 2µl DNA extracts added to each qPCR reaction, and conc=HF183 copies per 1 µl DNA extracts (i.e., copy of HF183 per 1 ml of water assuming 100% recovery during DNA extraction).

References

Ahmed, W., B. Hughes and V. Harwood. 2016. Current Status of Marker Genes of Bacteroides and Related Taxa for Identifying Sewage Pollution in Environmental Waters. Water 8(6), 231.

Boehm, A.B., J.A. Soller and O.C. Shanks. 2015. Human-Associated Fecal Quantitative Polymerase Chain Reaction Measurements and Simulated Risk of Gastrointestinal Illness in Recreational Waters Contaminated with Raw Sewage. Environmental Science & Technology Letters 2(10), 270-275.

Cao, Y., G.L. Andersen, A.B. Boehm, P.A. Holden, J.A. Jay and J.F. Griffith. 2017. Determination of DNA-based Fecal Marker Aging Characteristics for Use in Quantitative Microbial Source Tracking. Technical Report 978. Southern California Coastal Water Research Project Authority. Costa Mesa, CA.

Green, H.C., R.A. Haugland, M. Varma, H.T. Millen, M.A. Borchardt, K.G. Field, W.A. Walters, R. Knight, M. Sivaganesan, C.A. Kelty and O.C. Shanks. 2014. Improved HF183 quantitative real-time PCR assay for characterization of human fecal pollution in ambient surface water samples. Applied & Environmental Microbiology 80(10), 3086-3094.

Helsel, D.R. 2012. Statistics for censored environmental data using Minitab and R, John Wiley and Sons, New York.

Stewart, J.R., A.B. Boehm, E.A. Dubinsky, T.T. Fong, K.D. Goodwin, J.F. Griffith, R.T.Noble, O.C. Shanks, K. Vijayavel and S.B. Weisberg. 2013. Recommendations following a multi-laboratory comparison of microbial source tracking methods. Water Research 47(18), 6829-6838.

Supplementary Tables

Table S1. Sampling summary. Sites are listed approximately from north to south as shown in the map. Column heading n.HF183, n.Ent, %>104, %>10400 refer to number of samples analyzed for HF183, number of samples analyzed for *Enterococcus* spp., % of samples with greater than 104 *Enterococcus* per 100ml, and % of samples with greater than 10400 *Enterococcus* per 100ml. Some samples were not analyzed for enterococci due to lab errors.

	summer dry				wet			
Site	n.HF183	n.Ent	% >104	% >10400	n.HF183	n.Ent	% >104	% >10400
Upper Ventura River	-	-	-	-	10	-	-	-
Ventura River	46	-	-	-	-	-	-	-
Upper Santa Clara River	-	-	-	-	7	-	-	-
Santa Clara River	49	-	-	-	-	-	-	-
Industrial Drain	47	-	-	-	-	-	-	-
Upper Calleguas Creek	-	-	-	-	8	-	-	-
Calleguas Creek	49	-	-	-	-	-	-	-
Ramirez Creek	51	51	92	0	34	34	94	0
Escondido Creek	48	48	98	29	36	36	97	25
Solstice Creek	30	30	40	0	27	27	70	0
Marie Canyon Storm Drain	48	48	100	2	36	36	100	0
Malibu Creek	15	15	20	0	34	34	65	0
Topanga Creek	43	43	30	0	32	32	69	0
Santa Monica Canyon	-	-	-	-	37	37	100	24
Ballona Creek	50	50	62	0	39	39	74	13
Malaga Cove South	47	47	53	0	3	3	100	0
Malaga Cove East	47	47	87	0	4	4	100	0
Talbert Channel	43	43	84	2	15	15	100	33
Santa Ana River	44	44	64	0	14	14	93	64
Santa Ana Delhi Channel	49	49	45	0	50	50	98	38
Costa Mesa Channel	50	50	100	10	50	50	86	32
Aliso Creek	50	50	48	0	50	50	98	18
San Juan Creek	-	-	-	-	50	50	94	32
Cottonwood Creek	50	50	86	2	15	15	87	27
Tecolote Creek	54	54	91	2	29	27	100	33
San Diego River	53	53	45	4	24	24	92	8
Tijuana River	50	50	10	2	23	22	55	5
Overall	1013	822	67	3	627	599	88	19

Table S2. The HF183 qPCR standard curve parameters and lower limit of quantification (LLOQ) in Cq. E is the qPCR amplification efficiency $E=10^{-1/slope}-1$.

Lab	Slope	Y intercept	R^2	n	E	LLOQ in Cq
1	-3.52	38.98	0.993	172	0.92	35.5
2	-3.42	36.32	0.983	231	0.96	32.9
3	-3.38	35.06	0.994	190	0.97	31.7
4	-3.49	36.64	0.987	290	0.94	33.2
5	-3.35	36.76	0.966	107	0.99	33.4
6	-3.51	38.43	0.988	372	0.93	34.9
7	-3.57	39.23	0.952	178	0.91	35.7
8	-3.63	38.89	0.988	83	0.89	35.3

Table S3. Comparison of site ranking based on seven different definitions of frequency of HF183 positive and two different calculations of site average HF183 concentration, in dry (A) and wet (B) weather.

Frequency of HF183 positives was defined as the percentage of samples that were positive for HF183 at each site. Seven different definitions of a positive sample were used to calculate frequency. A sample is deemed positive for HF183 if at least one (amp1), or at least two (amp2), or all three (amp3) qPCR replicates amplified, if at least one (lod1), or at least two (lod2), or all three (lod3) qPCR replicates were DNQ or quantifiable, or if the sample average concentration was greater than limit of detection (lodAvg). In calculating the sample average, ND and DBLOD were substituted with ½ or 1 limit of detection.

Site average HF183 concentration was calculated by substituting ND and NBLOD with a statistically based Poisson estimate (AvgPois) or with ½ or 1 limit of detection, respectively (AvgLod).

Number of samples at each site is denoted by n.spl, and sites with 10 or fewer samples are excluded. Sites are sorted from lowest to highest ranking position (i.e. having least to having most human fecal contamination) based on the "amp1" frequency definition.

A. Dry weather

		Site ranking position								
Site	n.spl	amp1	amp2	amp3	lod1	lod2	lod3	lodAvg	AvgPois	AvgLod
Santa Ana River	44	1.5	1.5	2	1.5	1.5	2	1.5	1.5	1.5
Topanga Creek	43	1.5	1.5	2	1.5	1.5	2	1.5	1.5	1.5
Malibu Creek	15	3	9	13	6	13	16	10	3	3
Industrial Drain	47	4	5	9	9	7.5	10.5	7	5	6
Ventura River	46	5	11	10	10	14	12	13	6	9
Tecolote Creek	54	6	6	4	8	10	4	12	7	11
Santa Clara River	49	7	4	2	11	5.5	2	5.5	4	5
Solstice Creek	30	8	3	6	3	3	6	3	9	4
Talbert Channel	43	9	14	14	7	9	13	11	14	10
Ramirez Creek	51	10	13	11	16	11	7	8	10	12
Malaga Cove South	47	11	8	12	12	7.5	10.5	9	11	13
Aliso Creek	50	12	10	7	4	4	8	4	12	7
San Diego River	53	13	16	15	15	16	14	16	13	16
Santa Ana Delhi Channel	49	14.5	7	8	5	5.5	9	5.5	15	8
Calleguas Creek	49	14.5	12	5	17	15	5	15	8	15
Marie Canyon Storm Drain	48	16	17	17	14	17	17	17	17	17
Tijuana River	50	17	15	16	13	12	15	14	16	14
Escondido Creek	48	18	18	18	18	19	19	18	19	20
Ballona Creek	50	19	19	19.5	20	20	20	20	18	19
Costa Mesa Channel	50	20	20	19.5	19	18	18	19	20	18
Cottonwood Creek	50	21	21	21	21	21	21	21	21	21
Malaga Cove East	47	22	22	22	22	22	22	22	22	22

B. Wet weather

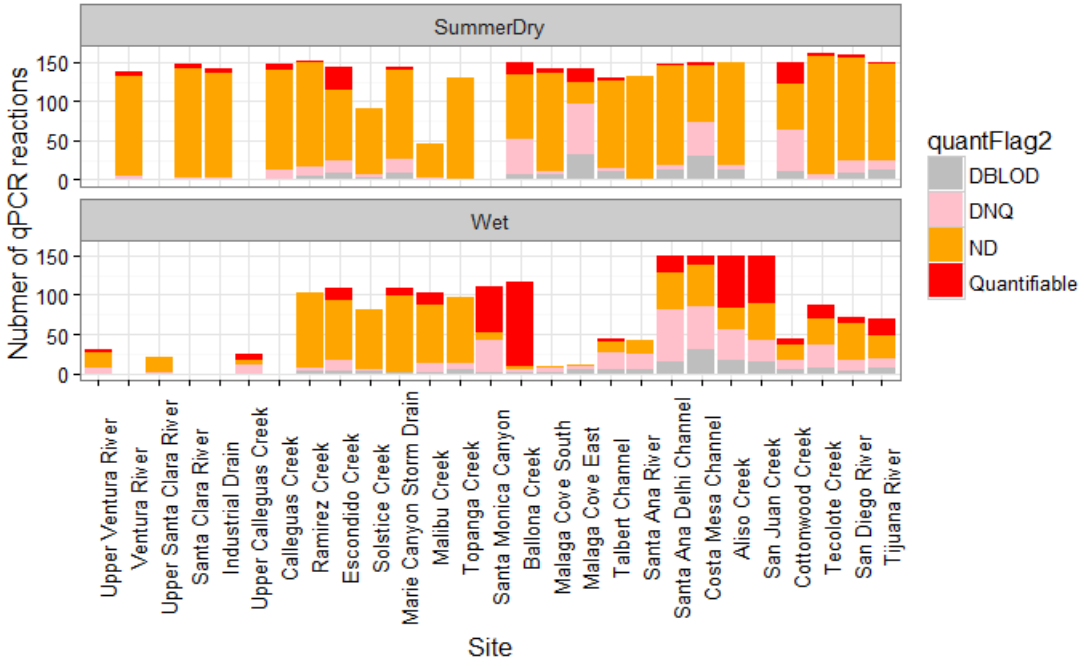
		Site ranking position								
Site	n.spl	amp1	amp2	amp3	lod1	lod2	lod3	lodAvg	AvgPois	AvgLod
Marie Canyon Storm Drain	36	1.5	2	4	2	4	4	4	1	4
Solstice Creek	27	1.5	1	1.5	1	1	1.5	1.5	3	1
Ramirez Creek	34	3	3	1.5	3	2	1.5	1.5	2	2
Topanga Creek	32	4	4	3	4	3	3	3	4	3
Malibu Creek	34	5	5	6	5	6	5	6	5	5
Escondido Creek	36	6	6	5	6	5	7	5	6	7
San Diego River	24	7	7	7	7	7	6	7	7	6
Tijuana River	23	8	10	10	9	10	11	11	14	14
Santa Ana River	14	9	11	9	10	11	8	8	8	8
Talbert Channel	15	10	13	15	14	13	14	15	10	10
San Juan Creek	50	11	14.5	14	13	14.5	15	14	15	15
Tecolote Creek	29	12	9	12	12	12	13	12	12	13
Costa Mesa Channel	50	14	12	11	11	8	10	9	9	9
Santa Ana Delhi Channel	50	14	14.5	13	15	14.5	12	13	13	12
Cottonwood Creek	15	14	8	8	8	9	9	10	11	11
Santa Monica Canyon	37	16	17	17	17	17	17	17	17	17
Aliso Creek	50	17	16	16	16	16	16	16	16	16
Ballona Creek	39	18	18	18	18	18	18	18	18	18

Table S4. Site rank contrast under wet vs. dry weather, as ranked by frequency or site average concentration. Sites are sorted from left to right by frequency of detection under dry weather conditions. Frequency of HF183 detection is defined as % samples with HF183 detection, and HF183 is said detected in a sample if HF183 is detected in any of the three qPCR replicates. Sites are sorted by ranking under dry weather.

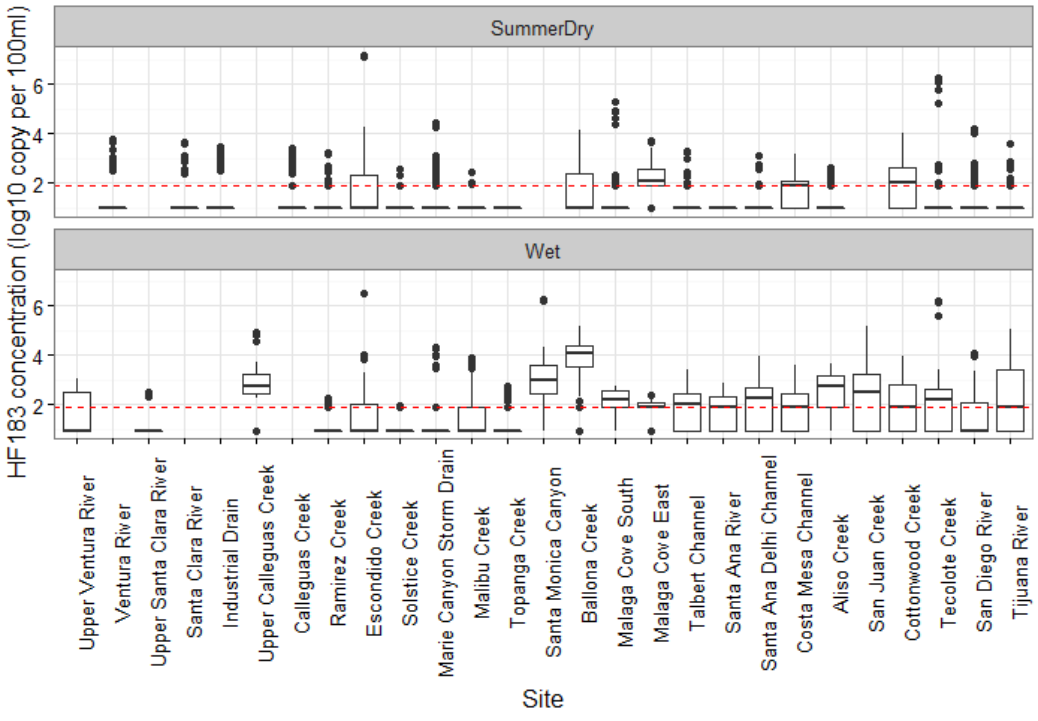
	Rank by frequency	
Site	Dry	Wet
Santa Ana River	1.5	9
Topanga Creek	1.5	4
Malibu Creek	3	5
Tecolote Creek	4	11
Solstice Creek	5	1.5
Talbert Channel	6	10
Ramirez Creek	7	3
Aliso Creek	8	15
San Diego River	9	7
Santa Ana Delhi Channel	10	13
Marie Canyon Storm Drain	11	1.5
Tijuana River	12	8
Escondido Creek	13	6
Ballona Creek	14	16
Costa Mesa Channel	15	13
Cottonwood Creek	16	13

	Rank by Site average concentration	
Site	Dry	Wet
Santa Ana River	1.5	8
Topanga Creek	1.5	4
Malibu Creek	3	5
Tecolote Creek	4	12
Solstice Creek	5	3
Ramirez Creek	6	2
Aliso Creek	7	15
San Diego River	8	7
Talbert Channel	9	10
Santa Ana Delhi Channel	10	13
Tijuana River	11	14
Marie Canyon Storm Drain	12	1
Ballona Creek	13	16
Escondido Creek	14	6
Costa Mesa Channel	15	9
Cottonwood Creek	16	11

Supplementary Figures



(A)



(B)

Figure S1. HF183 results distribution for dry and wet weather: (A) number of qPCR results in each quantification category, (B) HF183 concentrations (ND and DBLOD substitute with ½ and 1 of limit of detection). Dotted red line in (B) indicates the limit of detection. Note that empty space along the x-axis indicate that the site was not sampled during the given weather condition.

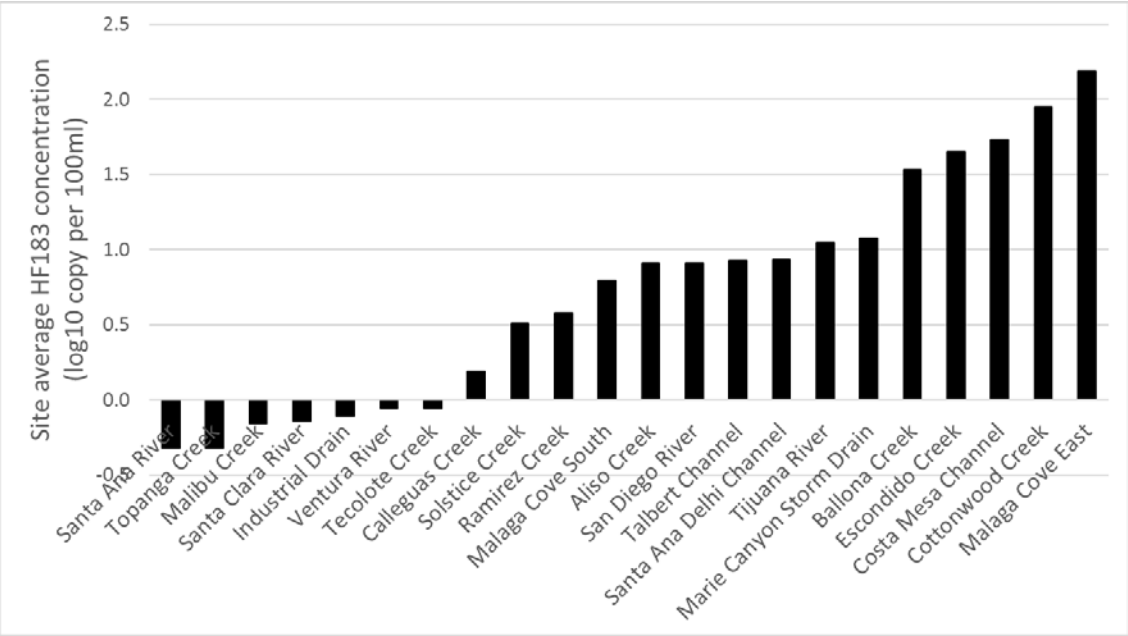


Figure S2. Site average concentration in log_{10} copies per 100ml by site for summer dry weather. Site average is calculated by the Poisson approach as described in the Methods section.

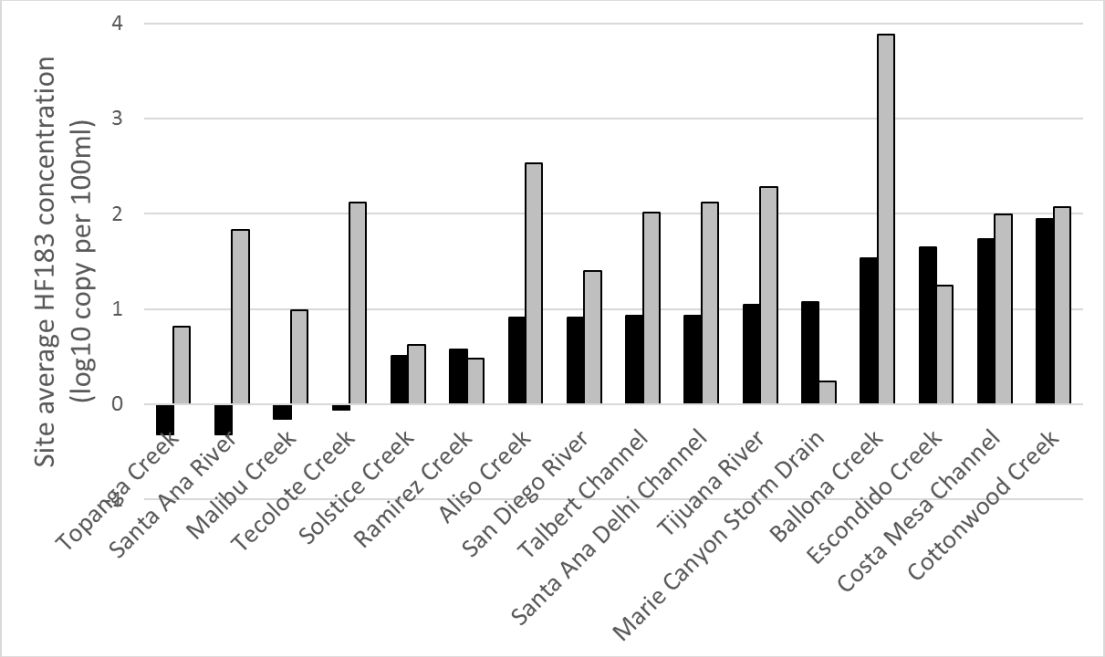


Figure S3. Site average HF183 concentration in log_{10} copies per 100ml by site in wet (light grey filled bars) versus dry (dark-filled) weather conditions. Site average is calculated by the Poisson approach as described in the Methods section. Site are sorted from left to right by site average concentration under dry weather conditions.

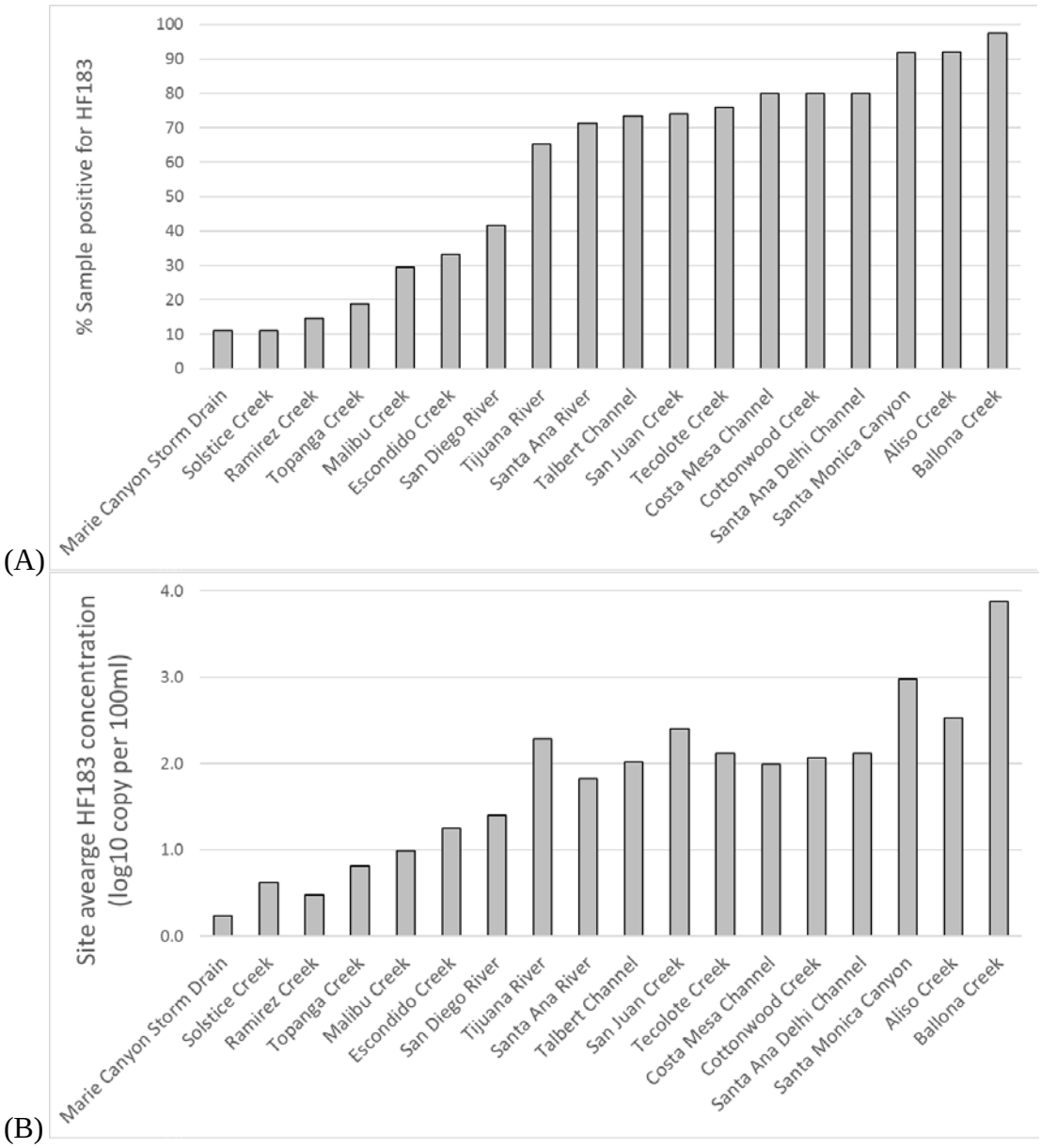


Figure S4. Frequency of HF183 positives (A) and site average HF183 concentration (B) at the 18 sites (number of samples per site >10) during wet weather. Frequency of HF183 positives is defined as % samples with positive HF183, and a sample is deemed HF183 positive if HF183 is amplified in any of the three qPCR replicates. Site average is calculated by the Poisson approach as described in the Method section. Sites are sorted from left to right by frequency of HF183 positives.

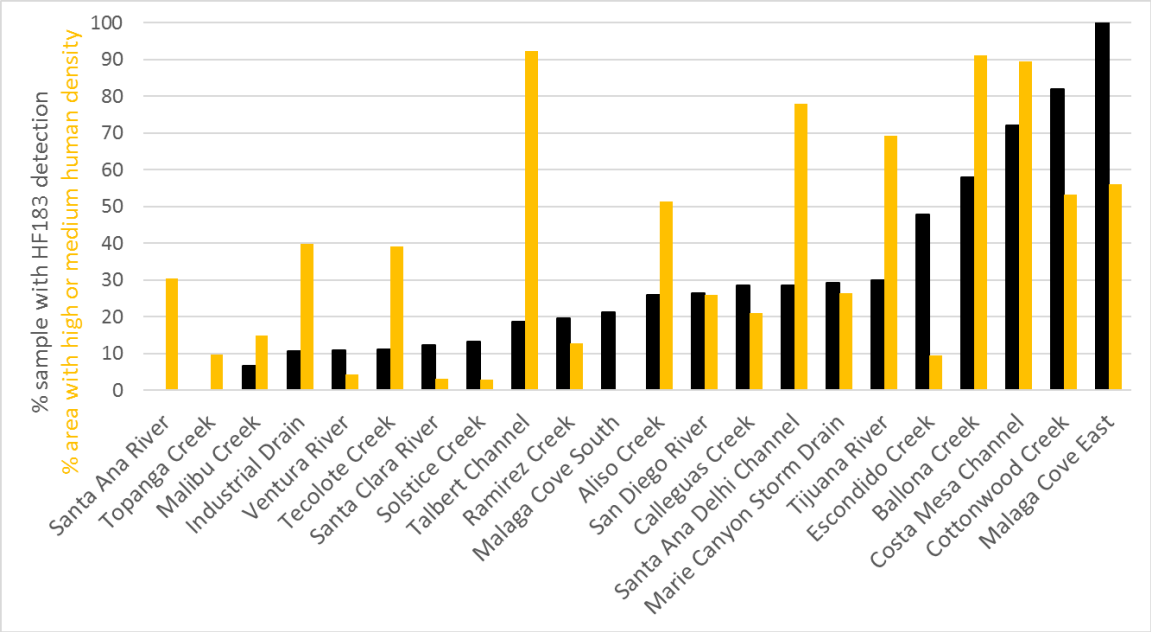


Figure S5. Frequency of HF183 positives in the drainages during dry weather conditions (black bars) and percentage area with high or medium human density in the watershed where the drainage is located (yellow bars). Sites are sorted from left to right by frequency of HF183 positives.

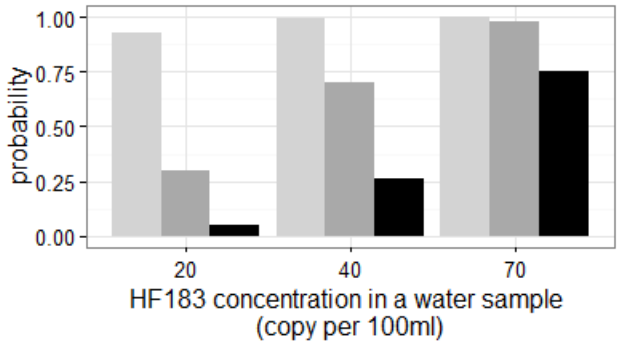
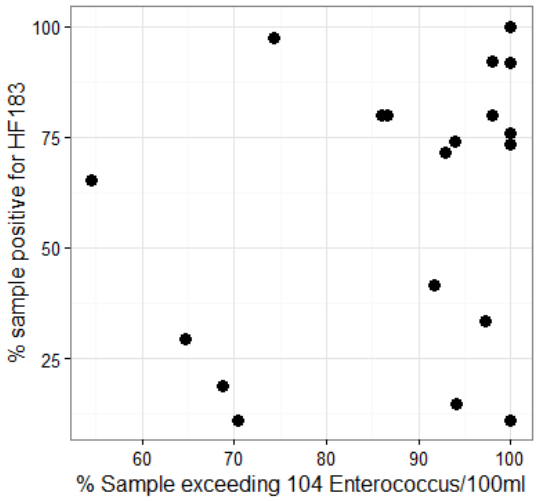
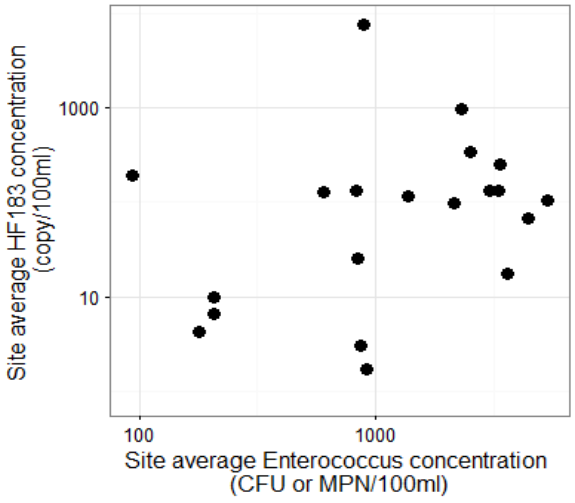


Figure S6. Probability of HF183 amplifying in at least 1 (light grey), at least 2 (dark grey), or all 3 (black) qPCR replicates at low HF183 concentration (x-axis). This simulation only accounts for subsampling effects as only 2 µl of the 100 µl total DNA extract from the water sample is analyzed in each qPCR reaction, while assuming 100% recovery during DNA extraction and 100% amplification for all reactions having at least 1 copy of the target present. Such assumption is likely untrue and the real probability of amplification is lower than shown here.



(A)



(B)

Figure S7. Ranking site by HF183- vs. Enterococcus-based metrics during wet weather: (a) Frequency of HF183 positive versus frequency of *Enterococcus* exceedance, (b) Site average HF183 concentration versus site average *Enterococcus* concentration. HF183-based metrics are as defined in Figure 3. Frequency of *Enterococcus* exceedance is defined as % samples with more than 104 *Enterococcus* per 100ml. Site average concentration of *Enterococcus* is defined as geomean at the site.