

UCLA

UCLA Electronic Theses and Dissertations

Title

Evaluation of the repeatability and reproducibility of a suite of PCR-based microbial source tracking methods

Permalink

<https://escholarship.org/uc/item/0x5768cz>

Author

Ebentier, Darcy

Publication Date

2012

Supplemental Material

<https://escholarship.org/uc/item/0x5768cz#supplemental>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Evaluation of the repeatability and reproducibility of a suite of
PCR-based microbial source tracking methods

A thesis submitted in partial satisfaction of the requirements
for the degree Master of Science in Civil Engineering

by

Darcy Louise Ebentier

2012

ABSTRACT OF THE THESIS

Evaluation of the repeatability and reproducibility of a suite of
PCR-based microbial source tracking methods

by

Darcy Louise Ebentier

Master of Science in Civil Engineering

University of California, Los Angeles, 2012

Professor Jennifer Ayla Jay, Chair

Many PCR-based methods for microbial source tracking (MST) have been developed and validated within research laboratories, but require inter-laboratory validation before implementation. As part of the State of California Source Identification Protocol Project (SIPP), a blinded set of challenge filters were analyzed by three to five laboratories with a suite of PCR-based methods utilizing standardized methods. Repeatability (within lab agreement) and reproducibility (between lab agreement) of results were assessed by multiple metrics and compared to previously observed values for other environmental methods. Repeatability and reproducibility were found to be generally comparable to previously observed values for other methods (Median CV .001-.033 and .020-.106, respectively). Variance component analysis showed contribution of laboratory to total variability to be larger but of similar magnitude to

inherent intra-laboratory variability among laboratories using standardized methods. Results among laboratories using non-standardized protocols for the same methods were also observed to have >2 log differences at times. These findings verify the repeatability and reproducibility of these MST methods and highlight the need for standardization of protocols and consumables prior to implementation of larger scale microbial source tracking studies involving multiple laboratories.

The thesis of Darcy Louise Ebentier is approved.

Keith D. Stolzenbach

Shaily Mahendra

Jennifer Ayla Jay, Committee Chair

University of California, Los Angeles

2012

Table of Contents

Abstract	ii
Committee	iv
Table of Contents	v
Acknowledgements	vi
Introduction	1
Materials and Methods	4
Results	9
Discussion	12
Conclusions	15
Tables	17
Figures	22
References	29

Acknowledgments

I would like to thank all of the laboratories, technicians, and principal investigators that helped generate these data, including Kaitlyn Hanley and Dr. Jennifer Jay at University of California, Los Angeles, Dr. Dan Wang and Dr. Alexandria Boehm at Stanford University, Meredith Raith, Elizabeth Scott, Dr. Blythe Layton, Dr. Yiping Cao and Dr. John Griffith at Southern California Coastal Water Research Project, Laurie C. Van De Werfhorst, Dr. Jared Ervin, and Dr. Patricia Holden of University of California, Santa Barbara, Dr. Kelly Goodwin and Dr. Chris Sinigalliano of the National Oceanic and Atmospheric Administration, Solen Lozach and Dr. Michéle Gourmelon of Ifremer, Cathy A. Kelty, Lindsay Peed, Dr. Jorge Santodomingo and Dr. Orin C. Shanks from the Environmental Protection Agency, Charlie McGee formerly of Orange County Sanitation District, and Dr. Stefan Wuertz from University of California, Davis. This study was funded by a Clean Beach Initiative Grant from the California State Water Resources Control Board. I would also like to give special thanks to my mentor, Dr. John Griffith, my advisor, Dr. Jennifer Jay, and my family for always allowing me to pursue my dreams.

Introduction

In the United States alone, there are over 88,000 miles of coastline and thousands of rivers and lakes that the general public uses for swimming, boating, or other activities (NOAA 1975). All of these areas are considered recreational waters and are monitored for contamination by state and municipal entities. Fecal indicator bacteria (FIB) such as *Enterococcus spp.* and *Escherichia coli* are currently used to detect fecal contamination. FIB are enumerated using methods certified by the Environmental Protection Agency (EPA) and the concentrations found in recreational waters are correlated to public health risk (Cabelli et al. 1983). For example, the California single sample exceedance threshold for *Enterococcus spp.* in marine waters is one hundred and four colony-forming units per hundred milliliters of sample (CFU/100ml). This level is associated with an additional nineteen illnesses per thousand individuals and above this level, California monitoring agencies deem it necessary to post health warnings or close any location where an exceedance occurs (EPA 2003)

Although the current monitoring methods used by US agencies can detect and quantify fecal contamination, they give no indication of its source. Fecal contamination in recreational waters and other water resources can come from a variety of sources, including sewage, domestic animals (dogs, cattle, chickens), native wildlife (seagulls, deer) and diffuse sources such as urban runoff. Input from any of these sources will likely result in quantifiable FIB, but the guilty party will remain unknown. In addition to being non-specific in terms of source, it has also been shown that FIB originating from a contamination event in the distant past can remain and even propagate in natural reservoirs such as sand and kelp (Lee et al. 2006). The problems associated with current FIB methods and their lack of specificity have highlighted the need for new approaches.

Microbial source tracking (MST) is one of several fields that has recently emerged due to the need for better, more specific tests for fecal contamination in recreational waters. Water quality managers, public health officials, and researchers have a common interest in MST, and although the field is young, powerful partnerships have resulted in rapid technological development over the past decade. Many PCR-based methods for microbial source tracking (MST) have been developed in an effort to characterize the sources of fecal pollution in recreational waters. They have the potential to be implemented by water quality managers for identification, source allocation, and remediation of chronic contamination problems as well as for quantitative microbial risk assessment (QMRA) applications (Soller et al. 2010). MST methods must be sensitive and specific to their target source as well as repeatable (demonstrate good intra-laboratory agreement) and reproducible (demonstrate good inter-laboratory agreement) for implementation in water quality management (Simpson et al. 2002, Stoeckel and Harwood 2007). Some PCR-based MST methods have been validated within research laboratories (Shanks et al. 2010a, Shanks et al. 2010b), however, inter-laboratory validation of these methods has been minimal. The few comparative studies that have been performed on the repeatability and reproducibility of MST methods have generally evaluated library dependent methods, which have since been proven to be largely unreliable (Griffith et al. 2003, Stoeckel et al. 2004).

Despite the absence of studies evaluating the reproducibility of PCR-based MST methods, there are previous studies that assess the repeatability and reproducibility of other microbiological and molecular environmental methods, offering metrics by which MST data can also be evaluated (Stoeckel et al. 2004, Shanks et al. 2012, Cao et al. 2011). Intra- and inter-laboratory coefficients of variation for related environmental water quality methods have been

observed and offer benchmark values for comparison (Shanks et al. 2012, Cao et al. 2011).

Here, we examined results from qPCR analysis by ten laboratories on replicate, blinded filter sets to evaluate repeatability and reproducibility for nine MST methods as part of the Source Identification Protocol Project (SIPP). Repeatability was defined as the ability of a method to produce the same answer for analyses of identical samples under the same conditions in the same laboratory. Reproducibility was defined as the ability to produce the same answer for analyses of identical samples under the same conditions in different laboratories. Five of the participating laboratories adhered to predetermined protocols and used the same lots of prepared reference DNA, DNA isolation kits, and amplification reagents so as to evaluate repeatability and reproducibility exclusive of protocol deviations. The remaining laboratories employed a variety of deviations from these standardized protocols and reagents for six of these eleven methods; thus we were able to evaluate the importance of protocol standardization to overall method reproducibility.

The specific goals of the study are to: 1) investigate the repeatability of each MST method evaluated by the SIPP study within laboratories when protocols and reagents are standardized 2) investigate the reproducibility of each MST method across several laboratories when protocols and reagents are standardized, 3) quantify the relative contribution of intra-laboratory and inter-laboratory variability to total variability, and 4) evaluate the effects of qPCR platform and protocol deviations on the reproducibility of MST methods. We also aim to identify the most repeatable and reproducible methods of those evaluated.

Materials and Methods

Sample Processing and Analysis

Blinded sets of 64 filters (32 duplicate samples generated from 12 composite fecal sources) were prepared in replicate according to Boehm et al. (2012). Fresh fecal composites included multiple human, sewage, septage, cow, dog, deer, pigeon, seagull, goose, chicken, pig, and horse samples collected from several geographic locations across California. Slurries containing single composite fecal sources (singletons) were prepared using sterile artificial water¹ and diluted to attempt a concentration of approximately 1000 *Enterococcus spp.* CFU/50mL². Dual source (doubleton) slurries were prepared by combining singleton slurries in 90% and 10% by volume ratios to approximate dominant and minor sources. A total of nineteen singleton slurries and thirteen doubleton slurries of varying strengths and ratios were generated (Table 1). All slurries were filtered through a 0.45µm polycarbonate filters (Millipore, Billerica, MA). Volumes filtered depended on desired strength; 50mL was filtered for full strength singletons and doubletons, while 5mL was filtered 1:10 strength singletons (Boehm et al. 2012). Filters were immediately flash frozen in liquid nitrogen and stored at -80 degrees Celsius. Replicate sets of frozen filters were shipped overnight on dry ice to participating laboratories and extracted by each laboratory, with the exception of Laboratory 7. Laboratory 7 performed analysis on extracts from Laboratory 9 that had been shipped overnight on dry ice.

Ten laboratories contributed data to this study. They included the Boehm laboratory at Stanford University (Palo Alto, CA, USA), the Wuertz laboratory at University of California, Davis (Davis, CA, USA), the Holden laboratory at University of California, Santa Barbara (Santa Barbara, CA, USA), the Jay laboratory at University of California, Los Angeles (Los

¹ Distilled water supplemented with 0.3 mM MgCl₂, 0.6 mM CaCl₂, and 1.4 mM NaHCO₃ and filter-sterilized with a 0.2 µm-pore size filtration unit (Polycap 36 AS, Whatman, Florham Park, NJ)

² Actual concentrations ranged over several orders of magnitude (Figure 1 in Boehm et al 2012)

Angeles, CA, USA), Southern California Coastal Water Research Project (SCCWRP) (Costa Mesa, CA, USA), two participants from the National Oceanographic and Atmospheric Administration Atlantic Oceanographic and Meteorological Laboratory (La Jolla, CA, and Miami, FL, USA), two participants from the U.S. EPA National Risk Management Research Laboratory (Cincinnati, OH, USA), and the laboratoire de microbiologie, EMP, at Ifremer (Plouzane, France). Stanford, UC Santa Barbara, UC Los Angeles, SCCWRP and one of the participating laboratories from U.S. EPA were randomly assigned numbers 1 through 5. The remaining laboratories were randomly assigned numbers 6 through 10.

Nine qPCR methods that were each run by four to six different laboratories were selected for repeatability and reproducibility analysis (Table 2). Laboratories 1-5 (hereafter referred to as core laboratories) used standardized protocols and reagents for DNA extraction and PCR analysis, while laboratories 6-10 utilized variable reagents and protocols (Table 3, Supplementary Materials). Each method was run by a minimum of three core laboratories. Samples were pre-screened by the coordinating laboratory for inhibition by spiking with a known concentration of target and assaying three serial 1:10 dilutions following the procedure described in Cao et al (2012). Core laboratories ran each sample in triplicate and screened sample data using salmon testis DNA as a sample processing control (Haugland et al. 2005). Core laboratories also obtained all reagents from the same commercial source and all prepared reference DNA standards from a centralized source (Tables 4, 5).

It should be noted that the HF183SYBR method (Seurinck et al. 2005) was eliminated from this study since this method was only performed on samples by two of the core laboratories, thus did not provide a large enough data set to observe reliable relationships and results. It was intended for inclusion in this study, however faced insurmountable difficulties in a

third core laboratory and was unable to be successfully used. Analysis of data from the two core laboratories that successfully implemented HF183SYBR showed very low specificity in one laboratory and high specificity in the other, indicating discrepant results (Boehm et al. 2012).

qPCR Data Treatment

Raw qPCR cycle threshold (C_T) values from core laboratories were treated with the same QA/QC protocols. Standard curve data for each lab, for each method were pooled and regression analysis was used to qualify outliers as those with a standardized residual $>|3|$. Outliers were removed iteratively until the standard curve data set no longer contained values with standardized residuals $>|3|$. An individual laboratory's method limit of detection (LOD) was defined as the standard concentration wherein at least 10 of 12 replicates amplified. Each LOD was then applied to the corresponding standard curve with outliers removed. Three core laboratories (1, 3, and 5) performed qPCR using a StepOnePlus platform (Applied Biosystems, Carlsbad, CA) while the remaining two (2, 4) used a CFX-96 platform (Bio-Rad, Hercules, CA). Platform-specific master standard curves were then generated by pooling standard curve data from the respective laboratories and again iteratively removing outliers with standardized residuals $>|3|$. Platform-specific master standard curves were used to determine the lower limit of quantification (LLOQ) for each platform-method, which was defined as the C_T value corresponding to 40 target copies per reaction. Upper limit of quantification (ULOQ) was defined as C_T value of 15. Limits of quantification (LOQs) were then applied as a qualifier to sample data. Samples with no amplification were given a C_T value of 40. Target concentrations for triplicate qPCR reactions of each filter were calculated for each method and adjusted for

crude extract recovery³ in the extraction step to approximate final copies/filter and then log transformed for later analysis.

Repeatability Among Laboratories Using Standardized Protocols and Reagents

Repeatability, defined as the ability of a method to produce the same answer for analyses of identical samples under the same conditions and in the same laboratory, was assessed for qPCR methods by using log transformed target concentration data from core laboratories. Coefficient of variation (CV) was calculated between triplicate measured concentrations for each filter containing target source material for each method in each laboratory. CV distributions were observed. Mean filter concentrations were then calculated from triplicate concentrations and the regression correlation coefficient (R^2) was observed between duplicate filters (n=32) for each method to corroborate challenge filter set integrity.

Reproducibility Among Laboratories Using Standardized Protocols and Reagents

Reproducibility, defined as the ability of a method to produce the same answer for analyses of identical samples under the same conditions in different laboratories, was assessed by using log transformed target concentration data from core laboratories. Mean filter concentrations for each laboratory were used to calculate CV across all core laboratories for each method. CV distributions were again observed.

Relative Contribution of Intra- and Inter-laboratory Variability to Total Variability

To determine the relative contribution of intra-laboratory and inter-laboratory variability to total variability in final target concentrations, nested ANOVA with variance component analysis was performed in SAS for all nine qPCR methods (SAS Institute, Cary, NC).

Laboratory and sample type were treated as random effects so as to infer the general inter-

³ For core laboratories, filters were lysed in 500 μ l buffer solution but only 350 μ l were able to be recovered for purification. Target concentrations were thus multiplied by a factor of 1.43 to approximate true filter contents.

laboratory reproducibility of these methods. The influence of sample type⁴ on inter-laboratory variability was assessed through inclusion of an interaction term of laboratory and sample type in the ANOVA model.

Reproducibility Among Laboratories Using Variable Equipment, Protocols, and Reagents

Core laboratories utilized two thermal cycler models, allowing for the observation of the effect of platform on reproducibility exclusive of any other protocol or reagent deviations. Initial regression analysis between mean target concentrations from each platform showed frequent bias high by the ABI StepOnePlus model. Thus, statistical analysis was performed to determine whether these regression lines had slopes significantly different than 1 or y-intercepts significantly different than 0.

The effects of protocol and reagent deviations on the reproducibility of MST methods were assessed using log transformed target concentration data from the six qPCR methods for which data from both core (1-5) and non-core (6-10) laboratories were available. Target concentrations for source filters from all participating laboratories were first visualized on a scatter plot for each method. Subsequently, coefficients of variation (CV) were again calculated from mean filter concentrations for each laboratory, this time including data from all laboratories 1-10. CV distributions were observed and compared to previously observed CVs among core laboratories.

⁴ Sample type includes both fecal source and slurry strength

Results

Repeatability Among Laboratories Using Standardized Protocols and Reagents

Median intra-laboratory CV ranged from .001 to .033, indicating minimal variation in replicate qPCR measurements of the same filter (Figures 1, 2). Animal-associated methods, with the exception of Laboratory 1 for the Gull2Taqman method, had maximum intra-laboratory CV values below .100 (Figure 2). DogBact, Pig2Bac, CowM2 and BacCow were found to have similar levels of repeatability, while Gull2Taqman showed poorer repeatability. Human-associated methods showed higher maximum intra-laboratory CV values than animal-associated methods (up to .307) (Figure 1). Upon further examination, the increased maximum CV values observed for human-associated methods were only found to be associated with a few filters, as shown by the CV distributions among these methods (Figure 1). In fact, with the exception of HumM2, all methods were observed to have a 90th percentile CV below .100 (Figure 1, 2). BacHum, HF183Taqman, and Bsteri were found to have similar levels of repeatability, while HumM2 showed poorer repeatability.

Evaluation of the replicate challenge filters produced coefficients of determination (R^2) between .88-1.00. An example regression is shown in Figure 3. These coefficients indicate good agreement between replicate filters within laboratories, as well as corroborating the integrity of the challenge filter set.

Reproducibility Among Laboratories Using The Same Protocols and Reagents

Median inter-laboratory CV for each method ranged from .020 to .106, indicating low variation between laboratories analyzing the same filter (Figure 4). Animal-associated methods, with the exception of Gull2Taqman, had maximum inter-laboratory CV values below .100, while human methods showed higher maximum inter-laboratory CV values (up to .277). Again, these

increased CV values were only associated with a few filters, as shown by the CV distributions among these methods (Figure 4). Among the human-associated methods, BacHum appeared to be the most reproducible, followed closely by HF183Taqman and Bsteri, and finally by HumM2. Among the animal-associated methods, BacCow appeared to be the most reproducible, followed by Pig2Bac, DogBact, CowM2 and Gull2Taqman.

Relative Contribution of Intra- and Inter-laboratory Variability to Total Variability

Nested ANOVA with variance component analysis produced relative contributions of singular factors to total error for each method (Figure 5). The sample type factor, which includes both fecal source and slurry strength, was found to contribute the most to total variability for each of these methods, as would be expected. Inter-laboratory and filter-to-filter variability were often found to be higher but of the same order of magnitude as inherent intra-laboratory variability for each method. Animal methods generally had lower total variability than human methods.

Reproducibility Among Laboratories Using Variable Protocols and Reagents

Statistical analysis of the slopes and y-intercepts of regression lines between platforms showed no significant differences from slope of 1 or y-intercept of 0 among animal-associated methods. However, the human-associated BacHum and HF183Taqman methods showed slopes significantly different than 1, and HF183Taqman, HumM2, and Bsteri methods showed y-intercepts significantly different than 0, indicating platform differences for these methods.

As expected, scatter plots showed widely varying differences between laboratories utilizing standardized protocols and laboratories deviating from standardized protocols. Example scatter plots are shown in Figures 6 and 7. Observed differences between core and non-core laboratories ranged from indiscernible to greater than two log copies/filter. The

HumM2 method appeared to show the most similar final target concentrations despite protocol and reagent deviations, whereas the BacHum method showed conspicuous differences for one non-core laboratory (Figure 6).

Median inter-laboratory CV for each of the six qPCR methods performed by core (1-5) and non-core (6-10) laboratories ranged from .041-.169, indicating lower reproducibility among target concentrations measured by laboratories deviating from standardized protocols (Figure 7). The Pig2Bac, BacCow, and DogBact methods showed maximum CVs $<.100$, however BacHum, HumM2, and Gull2Taqman showed maximum CVs between .573-.669. The highest CV values (CV $>.25$) again seemed to be associated with a small number of filters, as shown by the CV distributions among these methods (Figure 7). HumM2, despite much a much higher maximum CV value among all laboratories 1-10, showed the most similar reproducibility compared to reproducibility among core laboratories. Gull2Taqman and BacHum showed the poorest reproducibility compared to reproducibility among core laboratories.

Discussion

The present study provides promising information with regards to the future of PCR-based MST methods and their eventual implementation. Intra- and inter-laboratory coefficients of variation were generally low, with median CVs for all MST methods analyzed ranging from .001-.033 and .020-.106, respectively. Animal-associated methods showed maximum inter-laboratory CVs up to .101. These values are very similar to those observed during inter-laboratory validation of qPCR methods for enumeration of fecal indicator bacteria (FIB), as well as those found during inter-laboratory validation of methods for enumeration of FIB from sand (Shanks et al. 2012, Cao et al. 2011). Human-associated methods in this study showed maximum inter-laboratory CVs up to .307. These higher coefficients of variation, as mentioned above, were often associated with a small number of filter and were often doubleton filters containing sewage and/or septage as the minor contributor or 1:10 strength singletons (Table 1). This result is not surprising, though: upon reexamination of the data, these filters in particular contained relatively low amounts of target DNA: thus, higher variability as a function of high C_T values would be expected (Cao et al. 2012). Additionally, the HumM2 method, which observed target concentrations 1-2 log copies/filter lower than the other human-associated methods, fared the worst in terms of repeatability and reproducibility. These observations further support the finding that sample type contributes the most to total variability by the nested ANOVA with variance component analysis and suggest that more sensitive methods are more reproducible. Given the fact that sample type varies in both fecal source type as well as concentration (full or 1:10 strength), the relationship between fecal source and target concentration variability warrants further investigation.

It should also be considered that this study was not exclusively designed with repeatability and reproducibility analyses in mind. Given the serious logistical implications of generating duplicate challenge sets consisting of 64 filters for more than 25 participating laboratories, decisions had to be made regarding priority sources. The design of the challenge filter set was heavily weighted towards human sources: 38 of 64 filters contained solid human feces, sewage or septage, while animal sources were present on between 2 and 12 filters of the 64 (Table 1). The smaller number of source filters analyzed by animal-associated MST methods may explain the lower CVs and lower total variability observed compared to human-associated methods. Additionally, although efforts were made to provide varying concentration levels, the samples analyzed in this study were made up of relatively highly concentrated fecal source material with host-associated sequences sometimes existing at concentrations upwards of 10^9 gene copies/filter. Concentrations this high, though not impossible in some source waters, are not likely to occur in the natural environment and thus evaluation of repeatability and reproducibility at lower, more environmentally realistic concentrations is necessary. Future multi-laboratory studies of MST methods reproducibility should emphasize the inclusion of samples at concentrations more likely to be observed in the environment and in a more balanced design.

The findings of this study indicate that deviations from a standardized protocol can have widely varying impacts on the reproducibility of PCR-based MST methods. In the evaluation of the use of different thermal cycling platforms among laboratories employing otherwise standardized protocols and reagents, some methods were observed to have significantly different slopes and y-intercepts than 1 and 0, respectively. However, the variability associated with platform is grouped with variability associated with laboratory in the nested ANOVA with

variance component analysis. This actually suggests that platform has a very minor impact on total target concentration variability, considering that this grouped laboratory-plus-platform variability is still on the same order of magnitude as intra-laboratory variability, and that the contribution of laboratory to total target concentration variability may actually be lower than reported here.

Though analyses by laboratories 6-10 were not conducted in a way that allowed this study to quantify the contribution of other protocol deviations such as DNA extraction to sample concentration variability, the findings of increased inter-laboratory variability among laboratories deviating from standardized protocols certainly affirm the need for establishment of standardized protocols and centralized sources of reference materials before successful implementation and technology transfer. For example, analysis of challenge filters by the BacHum method showed minimal differences between Laboratory 9 and core laboratories, however Laboratory 10 observed target concentrations as much as two logs higher than core laboratories (Figure 5). Laboratory 9 deviated from the core laboratory protocol only in their choice of qPCR supermix and used the same prepared reference plasmid material as core laboratories (Tables 2, 3, Supplementary Materials). Laboratory 10 deviated from core laboratories in minor aspects of the extraction protocol and their choice of qPCR supermix and used the same reference plasmid as core laboratories, however their reference material was extracted in house (Table 3, Supplementary Materials). The effects of these minor differences highlight the need for complete standardization of methods both in terms of reagents and sources of reference materials prior to implementation for large-scale microbial source tracking efforts.

Conclusions

In summary, the present study was able to determine that

- Intra- and inter-laboratory coefficients of variation for PCR-based MST methods in laboratories employing standardized protocols are generally comparable to published values for other environmental methods, with the exception of some values produced by human-associated methods on filters containing low levels of sewage and/or septage.
- Animal-associated MST methods showed lower total variability than human-associated MST methods, and variability was largely associated with sample type. However, this lower variability is likely associated with sample size and warrants reevaluation in a future study with a balanced challenge sample design
- Among human-associated methods, BacHum, HF183Taqman, and Bsteri were equally repeatable. BacHum was the most reproducible, followed closely by HF183Taqman and Bsteri. HumM2 was the least repeatable/reproducible and had higher coefficients of variation than other human-associated methods, but showed lower total variability. Lower observed target concentrations for HumM2 compared to other human-associated methods are suspected to be the reason for this result.
- Among animal-associated methods, Pig2Bac, CowM2, BacCow, and DogBact were more repeatable than Gull2Taqman. BacCow was more reproducible than CowM2. DogBact, Pig2Bac, and Gull2Taqman were of similar reproducibility to CowM2. Like HumM2, lower observed target concentrations for CowM2 compared to BacCow are may be the reason for poorer CowM2 reproducibility than BacCow.
- Inter-laboratory variability was found to be higher than intra-laboratory variability for most methods, however relative contribution to total variability was of the same order of

magnitude as intra-laboratory and was much less than the relative contribution of sample type. It is also suspected that a portion of the observed inter-laboratory variability may be attributable to thermal cycler platform differences.

- Observed differences between laboratories performing different SOPs varied widely and direct relationships between the number or magnitude of deviations and differences were unable to be established with this study design.
- These findings reaffirm the need for further investigation of sample type and concentration relationships to MST method reproducibility as well as standardization of protocols, equipment and consumables prior to implementation of larger scale microbial source tracking studies involving multiple laboratories.

Tables

Table 1. List of all singleton and doubleton slurries of various strengths and volumetric ratios (Boehm et al 2012)

	Fecal Source	Strength or Ratio
Singletons	Chicken	Full
	Deer	Full
	Dog	Full, 1:10
	Goose	Full
	Gull	Full, 1:10
	Horse	Full
	Pig	Full, 1:10
	Pigeon	Full
	Cow	Full, 1:10
	Human	Full, 1:10
	Septage	Full, 1:10
	Sewage	Full, 1:10
	Doubletons	Sewage/Chicken
Sewage/Gull		10/90, 90/10
Sewage/Pig		10/90, 90/10
Human/Cow		10/90, 90/10
Human/Dog		10/90, 90/10
Human/Goose		10/90
Human/Gull		10/90, 90/10
Septage/Horse		10/90

Table 2. Laboratories and Methods

Method Nickname	Target	Analyzing Laboratories	Reference
BacCow	Cow	1, 2, 4, 5, 10	(Kildare et al. 2007)
BacHum	Human	1, 2, 3, 5, 9, 10	(Kildare et al. 2007)
BSteriF1	Human, Dog	1, 2, 4, 5	(Haugland et al. 2010)
CowM2	Cow	1, 2, 3, 4, 5	(Shanks et al. 2008)
DogBact	Dog	3, 4, 5, 7, 9	(Shibata et al. 2010)
Gull2Taqman	Gull	1, 3, 4, 6, 9	(Shibata et al 2010)
HF183Taqman	Human	1, 2, 3, 4, 5	(Haugland et al. 2010)
HumM2	Human	1, 2, 3, 4, 5, 9	(Shanks et al. 2009)
Pig2Bac	Pig	1, 2, 4, 5, 8	(Mieszkin et al. 2009)

Table 3. Extraction Methods Utilized By Participating Laboratories

Laboratory	Physical Lysis Method	Extraction Kit	DNA Extract Vol.
Laboratories 1-5	Mini-Beadbeater-16, 2 min (BioSpec, Bartlesville, OK)	DNA-EZ ST1 ⁵ (Generite, Kendall Park, NJ)	100µl
Laboratory 6	Mini-Beadbeater-8, 40 sec ⁸ (BioSpec, Bartlesville, OK)	MoBio PowerLyzer PowerSoil DNA Isolation Kit (MoBio, Carlsbad, CA)	100µl
Laboratory 7	FastPrep FP120, 2x30s ⁷ (Qbiogene, Carlsbad, CA)	DNA-EZ ST1 ⁵ (Generite, Kendall Park, NJ)	100µl
Laboratory 8	None.	QIAamp DNA Mini Kit ⁶ (Qiagen, Valencia, CA)	100µl
Laboratory 9	FastPrep FP120, 2x30s ⁷ (Qbiogene, Carlsbad, CA)	DNA-EZ ST1 ⁵ (Generite, Kendall Park, NJ)	100µl
Laboratory 10	Mini-Beadbeater-8, 1 min ⁸ (BioSpec, Bartlesville, OK)	DNA-EZ ST1 (Generite, Kendall Park, NJ)	200µl

⁵ Utilized modified kit protocol (Boehm et al 2012)

⁶ Utilized modified kit protocol (Mieszkin et al 2009)

⁷ Speed setting 5.0, 2 rounds of 30 second agitation

⁸ Max speed- 2800 oscillations/min

Table 4: Reagent Sources for Laboratories 1-5

Item	Part number	Source	Specifications
Primers	NA	Operon	Salt Free
Taqman Probes	NA	Operon	HPLC Purified
Universal Master Mix	4318157	ABI	
QuantiTect Master Mix	204435	Qiagen	
Bovine Serum Albumen	15260-037	Invitrogen	
DNase free water	17012-5200	MO BIO	
AE Buffer	19077	Qiagen	

Table 5: DNA Standards for Laboratories 1-5

Method Nickname	Standard Source	Extracted/Prepared By	Notes
BacCow	Wuertz Lab, University of California, Davis	Stanford	Circular plasmid
BacHum	Wuertz Lab, University of California, Davis	Stanford	Circular plasmid
BSteriF1	IDT Technologies	SCCWRP	Linearized plasmid
CowM2	IDT Technologies	SCCWRP	Linearized plasmid
DogBact	Dr. Sinigalliano, NOAA	SCCWRP	Circular plasmid
Gull2Taqman	Dr. Sinigalliano, NOAA	SCCWRP	Circular Plasmid
HF183Taqman	IDT Technologies	SCCWRP	Linearized plasmid
HumM2	IDT Technologies	SCCWRP	Linearized plasmid
Pig2Bac	IDT Technologies	SCCWRP	Linearized plasmid

Figures

Figure 1. Intra-laboratory coefficients of variation (CV) for human-associated methods. CVs were calculated between qPCR triplicate reactions. Boxes represent the 25th to 75th percentile range, whiskers represent the 10th and 90th percentiles, and dots represent outliers.

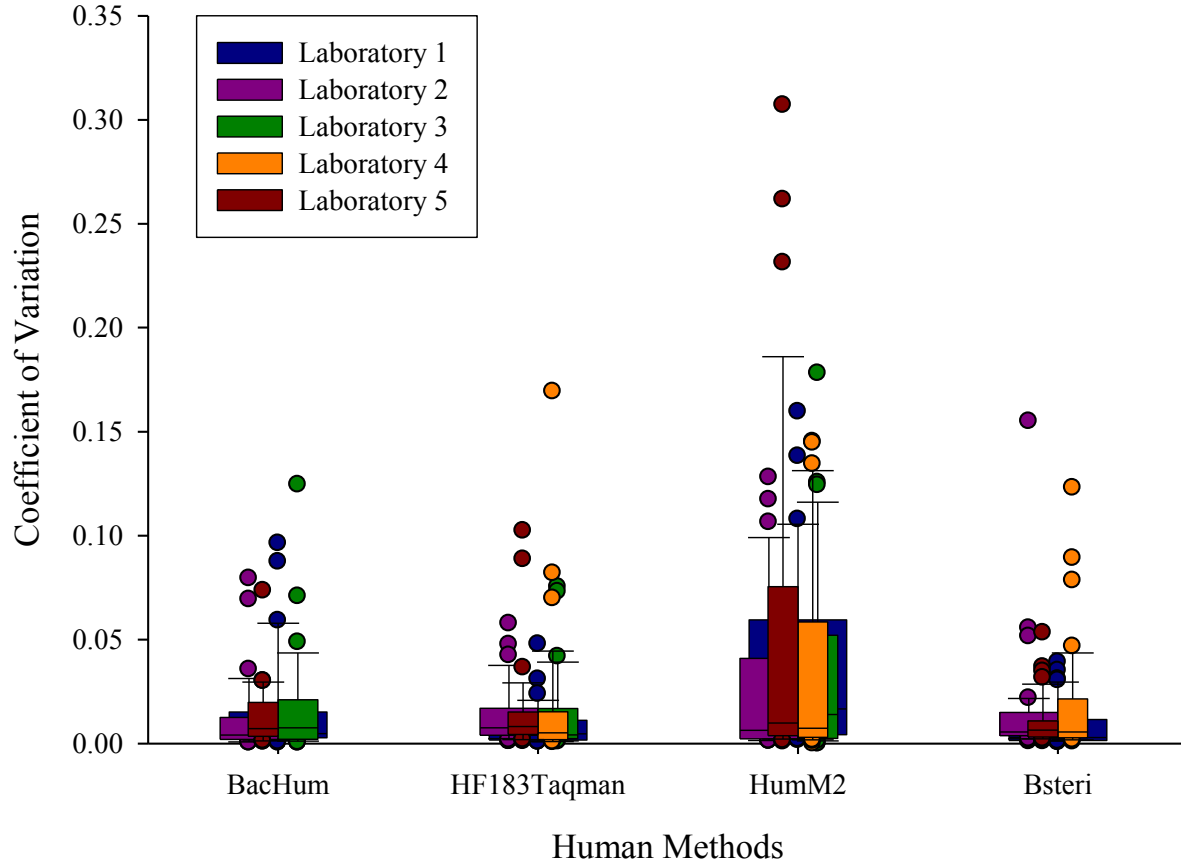


Figure 2. Intra-laboratory coefficients of variation (CV) for animal-associated methods. CVs were calculated between qPCR triplicate reactions. Boxes represent the 25th to 75th percentile range, whiskers represent the 10th and 90th percentiles, and dots represent outliers.

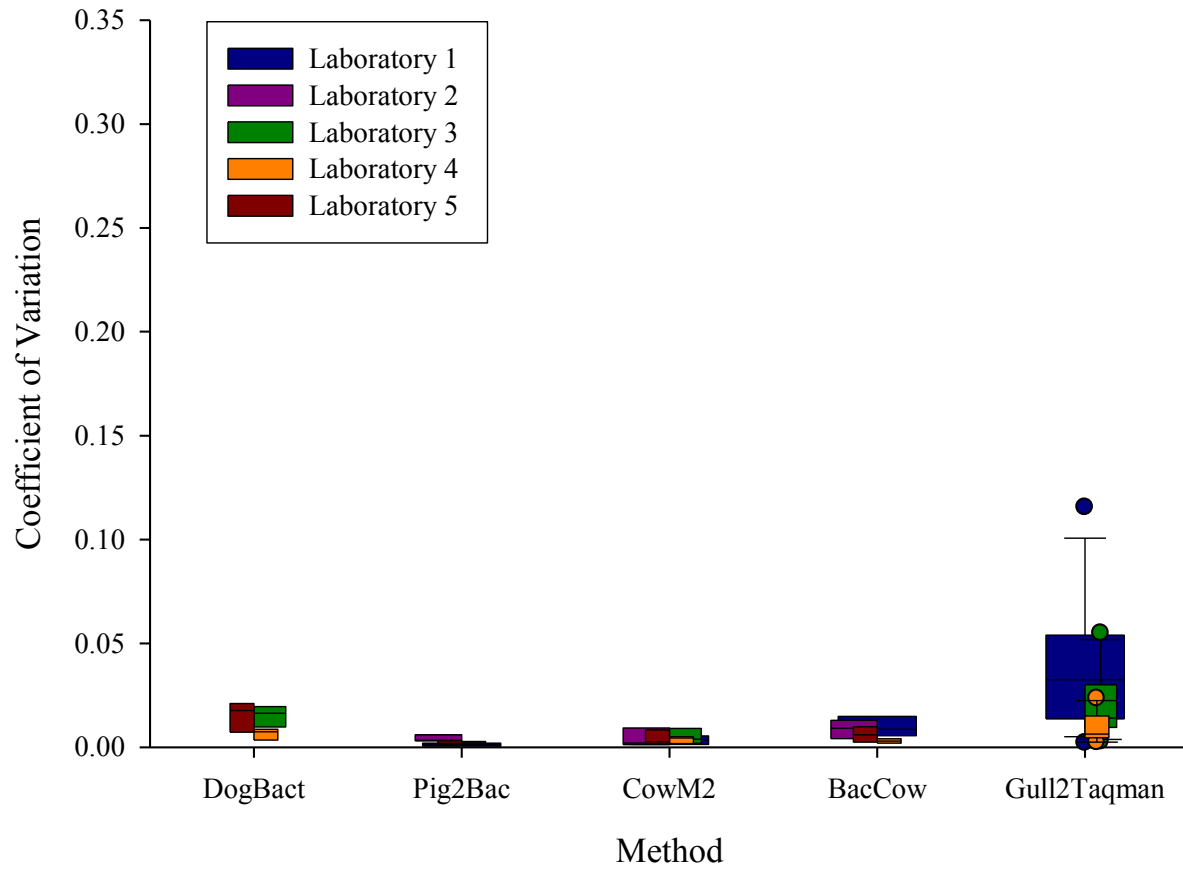


Figure 3. Sample duplicate filter regression plot for Laboratory 3, HF183Taqman method.

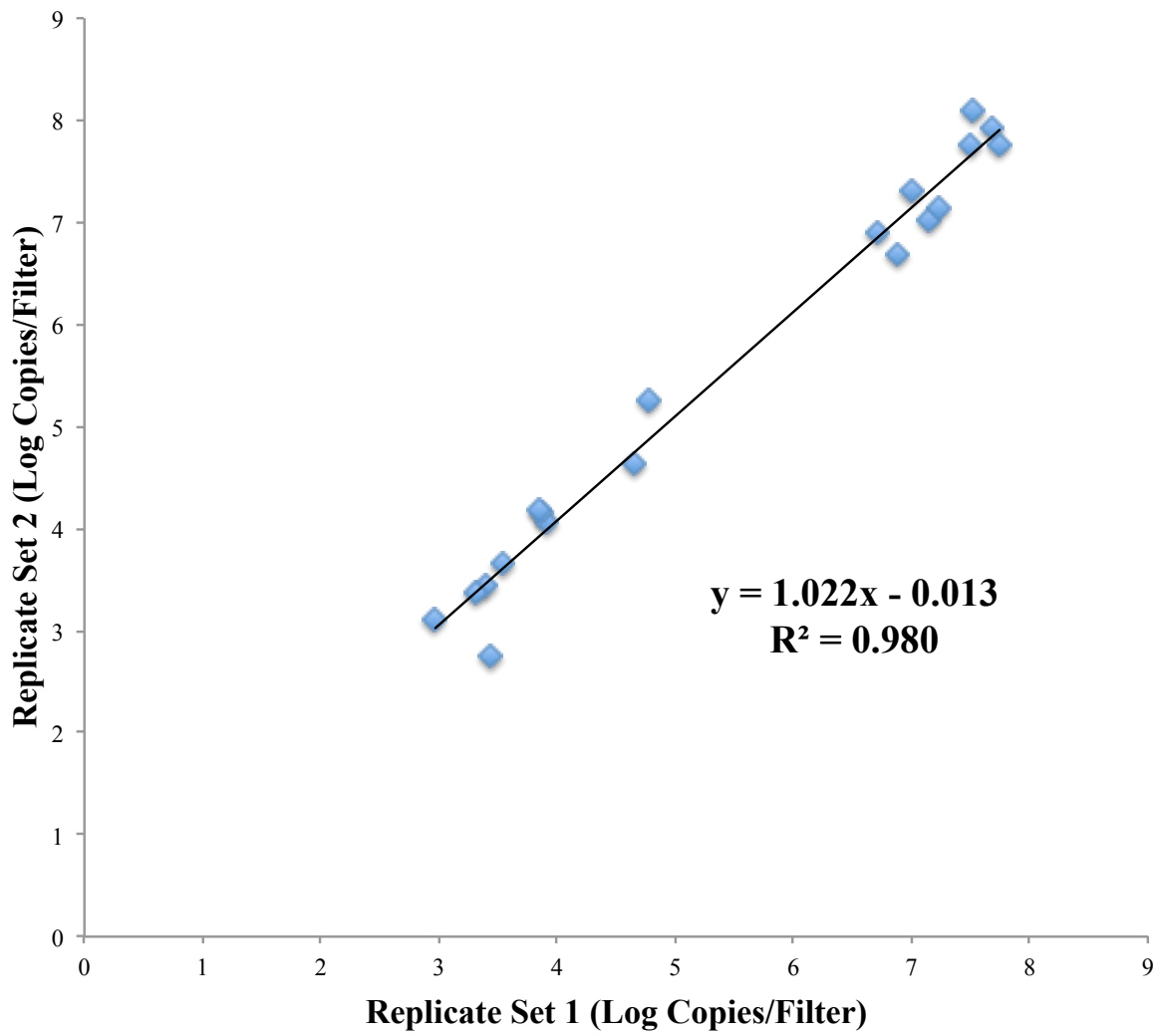


Figure 4. Inter-laboratory coefficients of variation (CV) among core laboratories (1-5) and among all laboratories (1-10). CVs were calculated between mean filter concentrations for each laboratory. With the exception of HumM2, inter-laboratory CVs were observed to be higher when laboratories employing non-standardized protocols were used.

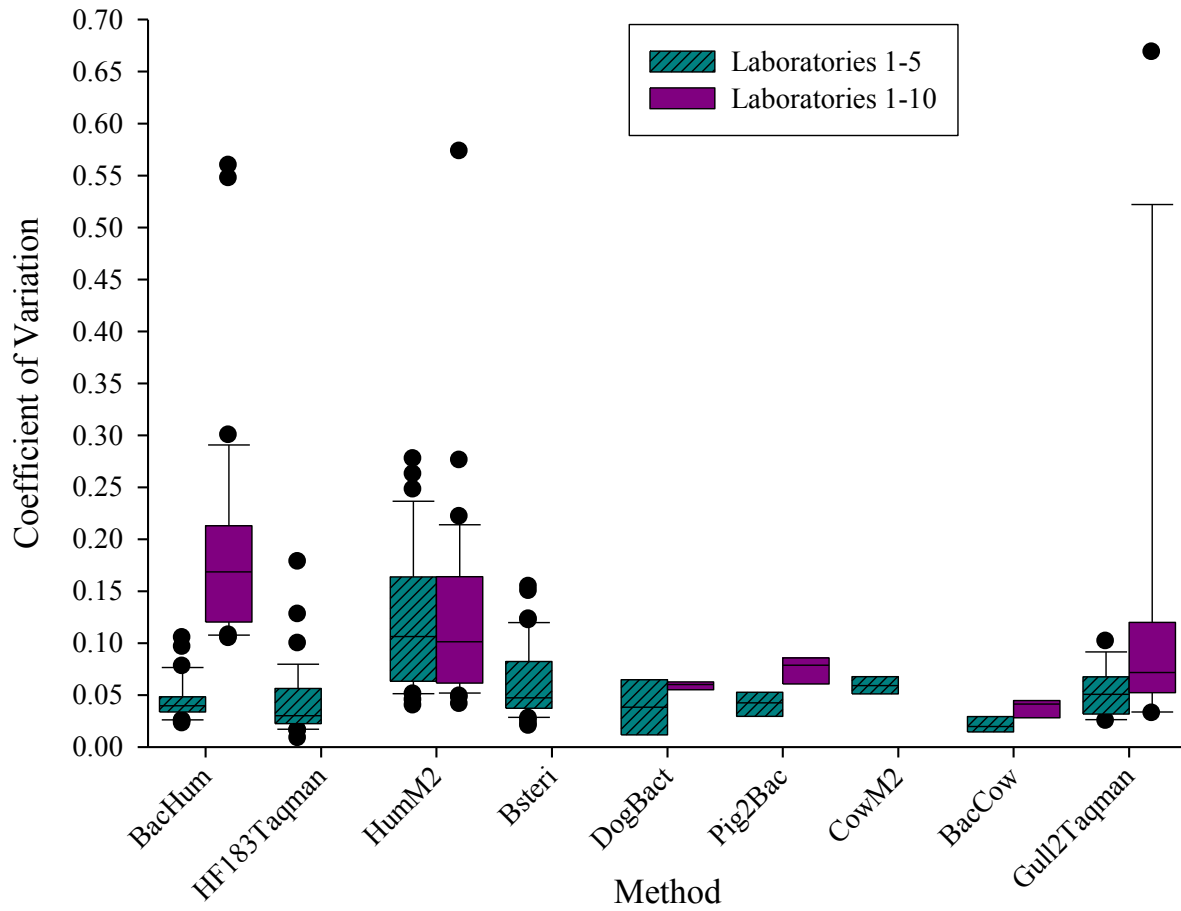


Figure 5. Relative contribution of singular factors to total variability among core laboratories (1-5) for nine qPCR MST methods

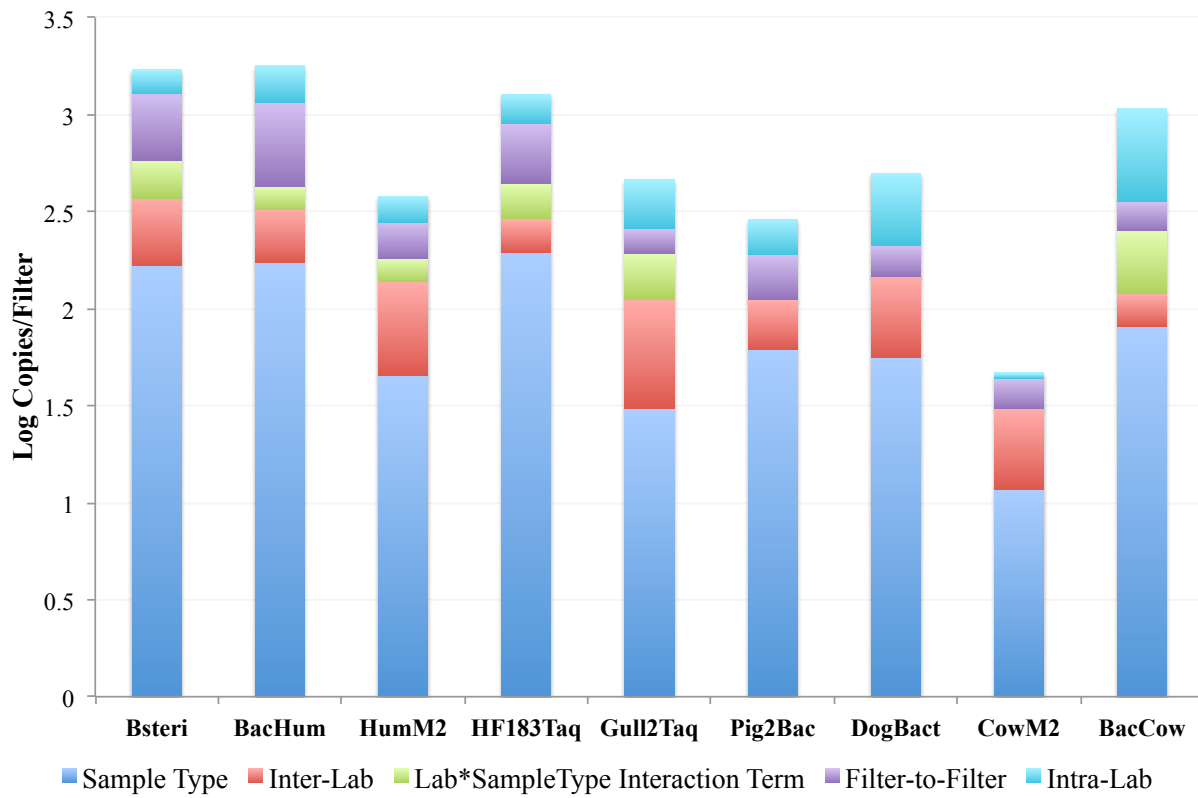


Figure 6. Observed BacHum target concentrations for laboratories 1-10. Laboratories 1-5 utilized standardized protocols and reagents, while Laboratories 6-10 employed a variety of deviations from said protocols.

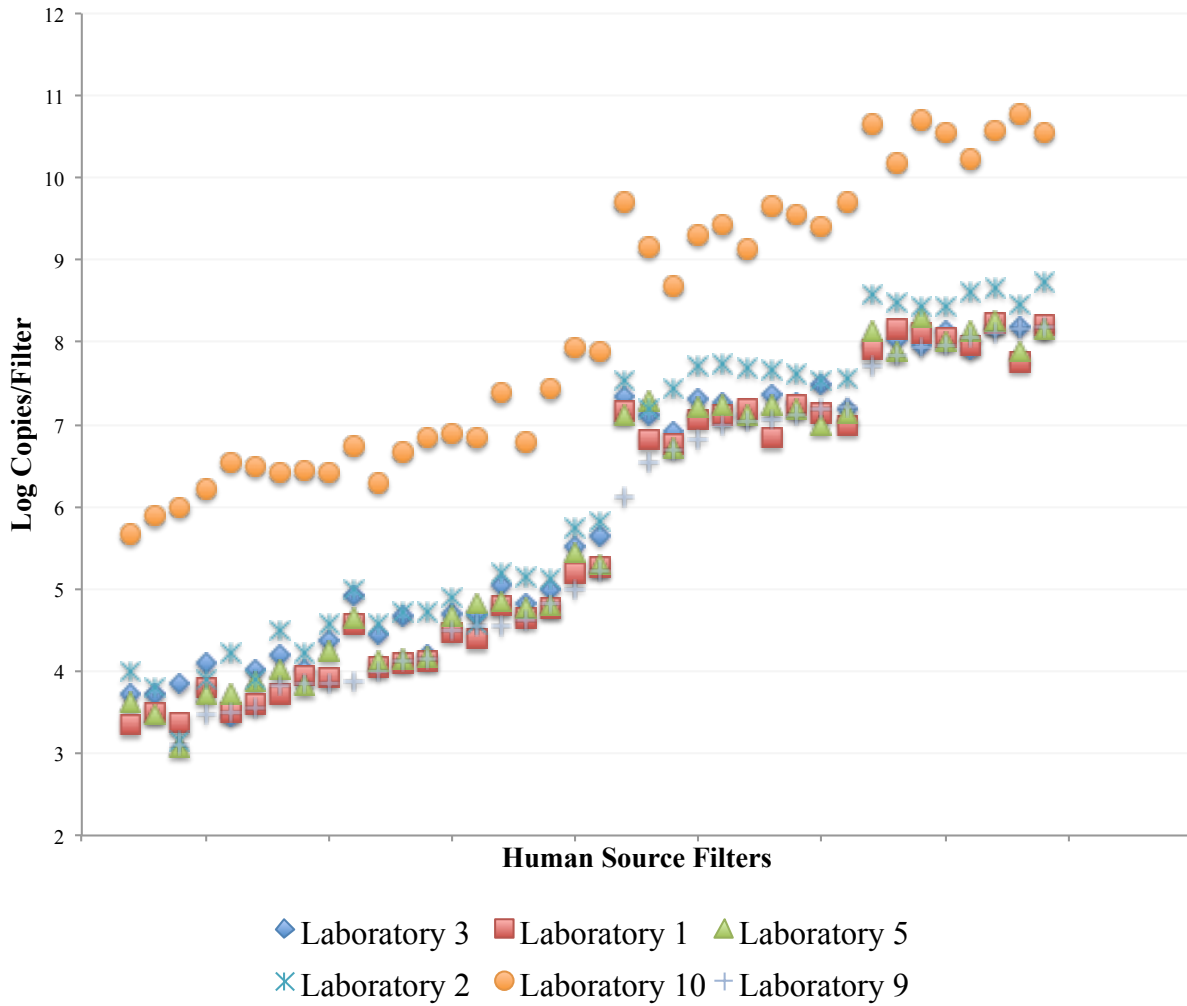
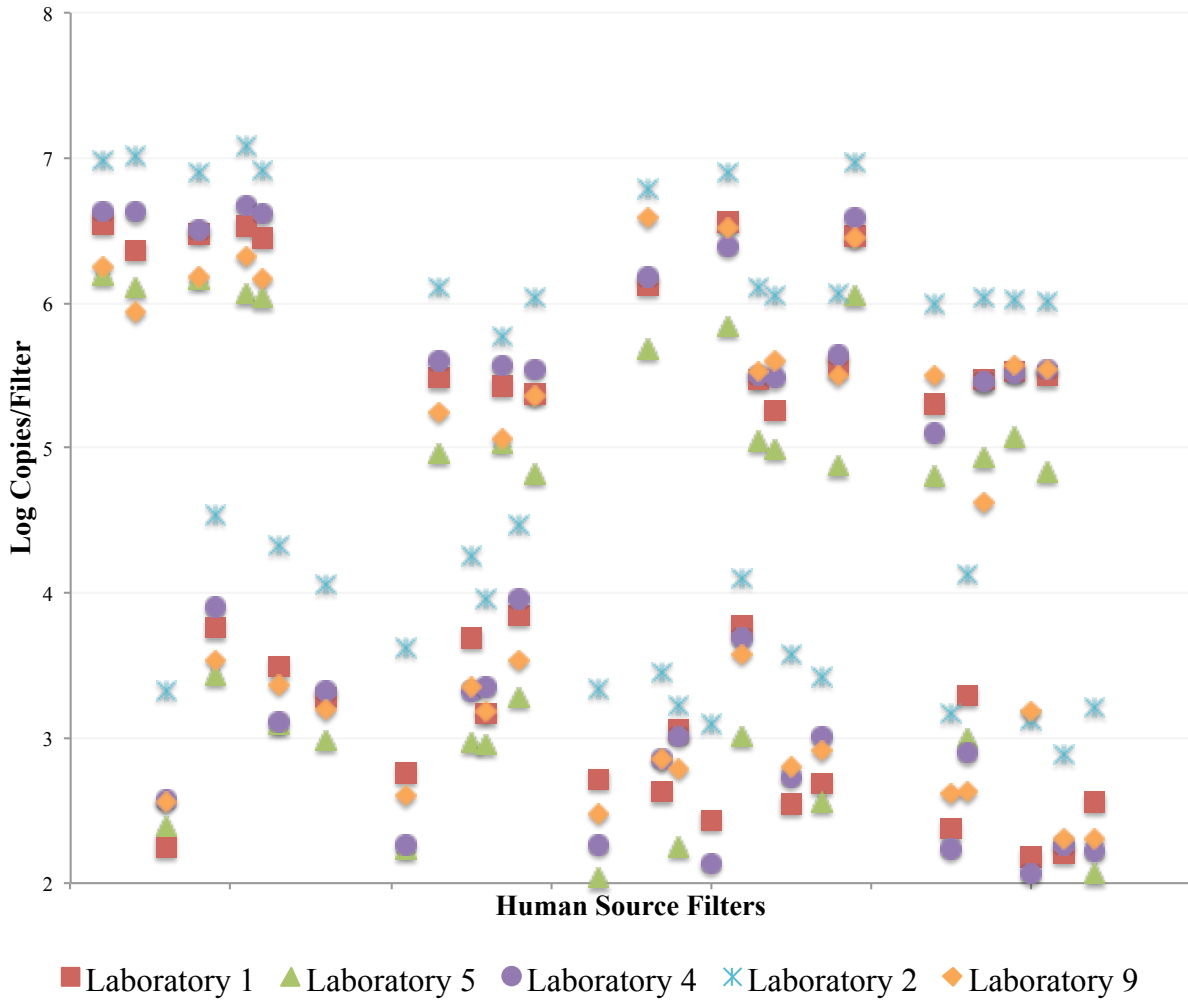


Figure 7. HumM2 target concentrations for laboratories 1-10. Laboratories 1-5 utilized standardized protocols and reagents, while Laboratories 6-10 employed a variety of deviations from said protocols.



References

- NOAA (1975) Coastline of the US. US Dept of Commerce, N.O.a.A.A. (ed).
- Cabelli, V., Dufour, A.P., McCabe, L. and Levin, M. (1983) A Marine Recreational Water Quality Criterion Consistent with Indicator Concepts and Risk Analysis. *Water Pollution Control Federation* 55(10), 1306-1314.
- EPA (2003) Bacterial Water Quality Standards For Recreational Waters, Environmental Protection Agency.
- Lee, C.M., Lin, T., Lin, C.-C., Kohbodi, G., Bhatt, A., Lee, R. and Jay, J.A. (2006) Persistence of fecal indicator bacteria in Santa Monica Bay beach sediments. *Water Research* (40), 2593-2602.
- 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.
- Simpson, J.M., Santo Domingo, J.W. and Reasoner, D.J. (2002) Microbial Source Tracking: State of the Science. *Environmental Science & Technology* 36(24), 5279-5288.
- Stoeckel, D.M. and Harwood, V.J. (2007) Performance, design, and analysis in microbial source tracking studies. *Applied and Environmental Microbiology* 73(8), 2405-2415.
- Shanks, O.C., White, K., Kelty, C., Sivaganesan, M., Blannon, J., Meckes, M., Varma, M. and Haugland, R. (2010a) Performance of PCR-Based Assays Targeting *Bacteroidales* Genetic Markers of Human Fecal Pollution in Sewage and Fecal Samples. *Environmental Science & Technology* 44(16), 6281-6288.
- Shanks, O.C., White, K., Kelty, C.A., Hayes, S., Sivaganesan, M., Jenkins, M., Varma, M. and Haugland, R.A. (2010b) Performance Assessment PCR-Based Assays Targeting *Bacteroidales* Genetic Markers of Bovine Fecal Pollution. *Applied and Environmental Microbiology* 76(5), 1359-1366.
- Griffith, J.F., Weisberg, S.B. and McGee, C.D. (2003) Evaluation of microbial source tracking methods using mixed fecal sources in aqueous test samples. *Journal of Water and Health* 1(3), 141-151.
- Stoeckel, D., Mathes, M., Hyer, K., Hagedorn, C., Kator, H., Lukasik, J., O'Brien, T., Fenger, T., Samadpour, M., Strickler, K. and Wiggins, B. (2004) Comparison of Seven Protocols To Identify Fecal Contamination Sources Using *Escherichia coli*. *Environmental Science & Technology* 38, 6109-6117.
- Shanks, O.C., Sivaganesan, M., Peed, L., Kelty, C.A., Blackwood, A.D., Greene, M.R., Noble, R.T., Bushon, R.N., Stelzer, E.A., Kinzelman, J., Anan'eva, T., Sinigalliano, C., Wanless, D., Griffith, J., Cao, Y., Weisberg, S., Harwood, V.J., Staley, C., Oshima, K.H., Varma, M. and Haugland, R.A. (2012) Interlaboratory Comparison of Real-Time PCR Protocols for

Quantification of General Fecal Indicator Bacteria. *Environmental Science & Technology* 46(2), 945-953.

Cao, Y., McGee, C.D., Griffith, J.F. and Weisberg, S.B. (2011) Method repeatability for measuring *Enterococcus* in southern California beach sands. *Letters in Applied Microbiology* 53(6), 656-659.

Boehm, A.B., Griffith, J.F., Holden, P., Jay, J.A., Shanks, O.C., Van De Werfhorst, L., Wang, D. and Weisberg, S.B. (2012) Specificity and Sensitivity of Forty-Two Microbial Source Tracking Assays: A Twenty-Five Lab Evaluation Study. *Water Research*.

Cao, Y., Griffith, J.F., Dorevitch, S. and Weisberg, S.B. (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* Accepted.

Haugland, R.A., Siefring, S.C., Wymer, L.J., Brenner, K.P. and Dufour, A.P. (2005) Comparison of *Enterococcus* measurements in freshwater at two recreational beaches by quantitative polymerase chain reaction and membrane filter culture analysis. *Water Research* 39(4), 559-568.

Seurinck, S., Defoirdt, T., Verstraete, W. and Siciliano, S.D. (2005) Detection and quantification of the human-specific HF183 *Bacteroides* 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater. *Environmental Microbiology* 7(2), 249-259.

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

Haugland, R.A., Varma, M., Sivaganesan, M., Kelty, C., Peed, L. and Shanks, O.C. (2010) Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected *Bacteroidales* species and human fecal waste by qPCR. *Systematic and Applied Microbiology* 33(6), 348-357.

Shanks, O.C., Atikovic, E., Blackwood, A.D., Lu, J., Noble, R.T., Domingo, J.S., Seifring, S., Sivaganesan, M. and Haugland, R.A. (2008) Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. *Applied and Environmental Microbiology* 74(3), 745-752.

Shibata, T., Solo-Gabriele, H., Sinigalliano, C., Gidley, M., Plano, L., Fleisher, J., Wang, J., Elmir, S., He, G., Wright, M., Abdelzaher, A., Ortega, C., Wanless, D., Garza, A., Kish, J., Scott, T., Hollenbeck, J., Backer, L. and Fleming, L. (2010) Evaluation of Conventional and Alternative Monitoring Methods for a Recreational Marine Beach with Nonpoint Source of Fecal Contamination. *Environmental Science & Technology* 44(21), 8175-8181.

Shanks, O.C., Kelty, C.A., Sivaganesan, M., Varma, M. and Haugland, R.A. (2009) Quantitative PCR for Genetic Markers of Human Fecal Pollution. *Applied and Environmental Microbiology* 75(17), 5507-5513.

Mieszkin, S., Furet, J.P., Corthier, G. and Gourmelon, M. (2009) Estimation of Pig Fecal Contamination in a River Catchment by Real-Time PCR Using Two Pig-Specific Bacteroidales 16S rRNA Genetic Markers. *Applied and Environmental Microbiology* 75(10), 3045-3054.