# Performance of human fecal anaerobe-associated PCR-based assays in a multi-laboratory method evaluation study

*Blythe A. Layton[1], Yiping Cao[1], Darcy L. Ebentier[2], Kaitlyn Hanley[2], Elisenda Ballesté[3], João Brandão[4], Muruleedhara Byappanahalli[5], Reagan Converse[6,7], Andreas H. Farnleitner[8], Jennifer Gentry-Shields[9], Maribeth L. Gidley[10], Michèle Gourmelon[11], Chang Soo Lee[12], Jiyoung Lee[12], Solen Lozach[11], Tania Madi[13], Wim G. Meijer[3], Rachel Noble[6], Lindsay Peed[14], Georg H. Reischer[8], Raquel Rodrigues[4], Joan B. Rose[15], Alexander Schriewer[16], Chris Sinigalliano[17], Sangeetha Srinivasan[15], Jill Stewart[9], Laurie C. Van De Werfhorst[6], Dan Wang[19], Richard Whitman[5], Stefan Wuertz[16,20], Jenny Jay[2], Patricia A. Holden[18], Alexandria B. Boehm[19], Orin Shanks[14] and John F. Griffith[1]*

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

[2]*University of California, Civil and Environmental Engineering Department, Los Angeles, CA*

[3]*University College Dublin, School of Biomolecular and Biomedical Science, Dublin, Ireland*

[4]*Instituto Nacional de Saúde, Lisboa, Portugal*

[5]*US Geological Survey, Great Lakes Science Center, Lake Michigan Ecological Research Station, Porter, IN*

[6]*University of North Carolina, Chapel Hill Institute of Marine Sciences, Morehead City, NC*

[7]*Oak Ridge Institute for Science and Education, Oak Ridge, TN*

[8]*Vienna University of Technology, Institute of Chemical Engineering, Environmental Microbiology and Molecular Ecology Group and Interuniversity Cooperation Centre for Water and Health, Vienna, Austria*

[9]*University of North Carolina, Department of Environmental Sciences and Engineering, Chapel Hill, NC*

[10]*University of Miami, Cooperative Institute for Marine and Atmospheric Studies, Miami, FL*

[11]*IFREMER, Laboratoire de Microbiologie, MIC/LNR, Département Ressources Biologiques et Environnement, Unité Environnement, Microbiologie et Phycotoxines, Plouzane´, France*

[12]*Ohio State University, College of Public Health, Division of Environmental Health Science, Columbus, OH*

[13]*Source Molecular Corporation, Miami, FL*

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

[15]*Michigan State University, College of Agriculture and Natural Resources, Department of Fisheries and Wildlife, East Lansing, MI*

[16]*University of California, Department of Civil & Environmental Engineering, Davis, CA*

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

[18]*University of California, Bren School of Environmental Science & Management, Earth Research Institute, Santa Barbara, CA*

[19]*Stanford University, Department of Civil & Environmental Engineering, Stanford, CA*

[20]*Nanyang Technological University, Singapore Centre on Environmental Life Sciences Engineering, School of Biological Sciences and School of Civil and Environmental Engineering, Singapore*

## Abstract

A number of PCR-based methods for detecting human fecal material in environmental waters have been developed over the past decade, but these methods have rarely received independent comparative testing in large multi-laboratory studies. In this study, ten of these methods (BacH, BacHum-UCD, *B. thetaiotaomicron* (BtH), BsteriF1, gyrB, HF183 endpoint, HF183 SYBR, HF183 Taqman®, HumM2, and *M. smithii nifH* (Mnif)) were evaluated using 64 blind samples prepared in one laboratory. The blind samples contained either one or two fecal sources from human, wastewater or non-human sources. The assay results were assessed for presence/absence of the human markers and also quantitatively while varying the following: 1) classification of samples that were detected but not quantifiable (DNQ) as positive or negative; 2) reference fecal sample concentration unit of measure (such as culturable indicator bacteria, wet mass, total DNA, etc); and 3) human fecal source type (stool, sewage or septage). Assay performance using presence/absence metrics was found to depend on the classification of DNQ samples. The assays that performed best quantitatively varied based on the fecal concentration unit of measure and laboratory protocol. All methods were consistently more sensitive to human stools compared to sewage or septage in both the presence/absence and quantitative analysis. Overall, HF183 Taqman was found to be the most effective marker of human fecal contamination in this California-based study.

## Introduction

The search for highly specific, sensitive, and cost effective human fecal-associated PCR-based assays has been a major focus of microbial source tracking (MST) research over the last decade. Many new methods have emerged as a result of that effort (Field and Samadpour 2007, Roslev and Bukh 2011). It is essential that MST methods be able to confirm the presence of human fecal contamination in environmental waters because of the ubiquity of fecal indicator bacteria (FIB) in non-human sources, knowledge gaps regarding illness risk from recreational exposure to non-human fecal sources (Boehm and Soller 2012), and the need to prioritize investment in wastewater infrastructure. Previously, library-based methods were in common use, but these were largely supplanted by PCR-based methods following a 2003 MST method evaluation study (Griffith *et al.* 2003). Until now, a large-scale multiple-laboratory MST method evaluation study has not been conducted since PCR-based methods came to the fore.

The need for confidence in the performance of human fecal-associated MST assays has recently become more urgent. The new United States Environmental Protection Agency (USEPA) criteria for recreational water quality offer beach managers the possibility of using quantitative microbial risk assessment (QMRA) to set site-specific criteria at beaches where the presence of human fecal pollution has been found sufficiently small through approved MST studies. Therefore it is crucial to robustly characterize the performance of MST methods that may be used to determine whether a beach is contaminated with human fecal pollution.

The most comprehensive, multiple-laboratory PCR-based MST method evaluation study to date is described in Boehm *et al.* (2013). Several important issues from this study remain open for further exploration in the present work. First, how does alternate classification of detectable but not quantifiable (DNQ) samples, versus non-detect (ND) samples, as positive change assay performance in presence/absence metrics? DNQ data handling often presents a trade-off between sensitivity and specificity, and as there is currently no consensus in the field regarding how to handle DNQ samples, it is important to consider both approaches when judging assay performance. Secondly, how did the assays perform, both qualitatively and quantitatively, under all available challenge filter sample units of measure? Characterization of challenge sample fecal concentrations in different terms (such as one milligram of wet feces or one nanogram of total DNA) may produce variable performance results, and every available characterization of the samples should be considered in order to comprehensively compare performance among assays. Third, how did the source of "target" sample (human stools, sewage, or septage) influence assay performance? Several factors may affect the performance assessment of these targets, including different states of decay and potential presence of non-human bacteria in wastewater; thus, it is important to evaluate these "target" sources separately. Lastly, how did the human-associated assays perform with mixed-source samples? The "doubleton" samples all contained a human stool, sewage, or septage "target" source plus a non-human fecal source, and thus offered the unique opportunity

to investigate human-associated assay sensitivity in the presence of non-target feces.

The present work seeks to fill the above gaps by considering each of these issues in detail. Accordingly, the goals of this paper are to evaluate the performance of ten human fecal anaerobe-associated PCR-based assays under varying characterizations of: 1) DNQ samples; 2) the challenge filter sample concentration unit of measure; and 3) the human "target" samples; and to evaluate the effect of combining human and non-human fecal sources in a sample.

## Methods

### Sample Creation and Analysis

Briefly, 64 blind challenge samples were created by mixing fresh feces (from chicken, cow, dog, deer, goose, gull, horse, human, pig, or pigeon), sewage, or septage in artificial freshwater. All fecal, sewage and septage samples were obtained from various sites across California. The filter set included 19 single-source ("singleton") and 13 mixed-source ("doubleton") samples in duplicate. Each doubleton sample contained human stools, septage, or sewage combined with one non-human fecal source. Detailed methods for the creation of the challenge filter samples can be found in Boehm *et al.* (2013).

Seventeen laboratories from the United States and the European Union contributed data to the study. The assay naming conventions used here generally follow the original publications (Table 1). The number of laboratories that performed each method is as follows: BacH (1), BacHum-UCD (6), BsteriF1 (4), BtH (1), gyrB (1), HF183 endpoint (7), HF183 SYBR (4), HF183 Taqman (5), HumM2 (6), and Mnif (5). The laboratories used six different DNA extraction methods: GeneRite DNA-EZ (12), Qiagen DNeasy® (1), Qiagen QIAamp® (1), MP Biomedicals FastDNA™ SPIN (1), MoBio PowerWater® (1), and phenol:chloroform extraction (1). Five laboratories involved in planning the study agreed to standardize their methods of DNA extraction (GeneRite DNA-EZ ST) and quantification (NanoDrop), q/PCR chemistries (Applied Biosystems TaqMan® Universal PCR Master Mix or TaKaRa Ex Taq® with original authors' primer/probe concentrations), and data processing (described in detail in Ebentier *et al.* 2013). These standardized operating protocols (SOPs) were made available to all participating laboratories, but adherence to the protocols was not required. Details of the laboratory SOPs and supply vendors are provided in the Supplemental Information (SI) Section 1 and Tables SI-1 and SI-2 (ftp://ftp.sccwrp.org/pub/download/DOCUMENTS/AnnualReports/2013AnnualReport/ar13_445_459SI.pdf). All data analyses in the present work were performed in R (v 2.14.0) with RStudio (v 0.96). Details of each analysis are described below.

### Classification of DNQ

The presence/absence sensitivity and specificity metrics were calculated in two ways: once with DNQ (detected but not quantifiable) samples considered positive and a second time with DNQ considered negative. All laboratories' data were analyzed together as one dataset and only the singleton (single-source) samples were included (every doubleton contained both a human and non-human fecal source, so it was not possible to independently evaluate sensitivity and specificity in the doubletons). All positive results for the endpoint assay were considered DNQ. The lower limit of quantification (LLOQ) for qPCR assays was defined for each laboratory as the lowest concentration on the standard curve where amplification was observed in at least 50% of qPCR replicates. The LLOQ values for each laboratory and assay are listed in Table SI-3. For samples within the range of quantification, the reported copy numbers were used. Samples with a quantification threshold cycle (Cq) greater than the laboratory-specific LLOQ were classified as DNQ regardless of how these samples were originally reported. DNQ samples were assigned a value of 150 copies/filter for quantitative analyses. This value was based on three assumptions: 1) a theoretical minimum detection limit of 3 copies per reaction (Bustin *et al.* 2009); 2) 2-µl template total DNA per reaction; and 3) 100 µl of DNA extract per filter. Assumptions 2 and 3 were valid for most laboratories and assays. Sensitivity and specificity metrics were calculated using the same equations and benchmarks described by Boehm *et al.* (2013).

### Challenge Filter Sample Units of Measure

The following units of measure were used to normalize the singleton qPCR data: wet mass, total DNA, *Enterococcus* CFU, *E. coli* CFU, *Enterococcus* qPCR (Haugland *et al.* 2005), *E. coli* 23S qPCR (Chern *et al.* 2011), and "general" Bacteroidales qPCR by GenBac3 (Siefring *et al.* 2008), AllBac (Layton *et al.* 2006), BacUni-UCD (Kildare *et al.*

**Table 1. Summary of original assay developer's publications.**

| Assay | Reference | Target Gene | Oligo Names | Reference Material (standards) | Test Material (target) | Challenge Material (non-target) | Challenge (quantity/reaction) | LLOQ (quantity/reaction) | Sensitivity (%) | Specificity (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| BacH | Reischer *et al.* (2007) | *Bacteroides* 16S | BacHf, BacH-pT, BacH-pC, BacHr | Plasmid | Human, Sewage, Cesspits | Cattle, Deer , Chamois, Roe Deer, Sheep, Goat, Horse, Fox , Dog, Cat, Pig, Chicken, Turkey, Swan, Duck, Black Grouse | 1 mg wet wt | 30 copies | 98 | 98 |
| BacHum-UCD | Kildare *et al.* (2007) | *Bacteroides* 16S | BacHum160f, BacHum193p, BacHum241r | Plasmid | Human, Sewage | Cow, Horse, Dog, Cat, Seagull | 5000 copies BacUni-UCD | 30 copies | 100 | 87 |
| BsteriF1 | Haugland *et al.* (2010) | *B. stericoris* 16S | BsteriF1DE, BthetP1, BthetR1 | Plasmid | Human, Sewage | Cattle, Pig, Chicken, Dog, Cat | 1 ng fecal DNA | 10 copies | 100 | NR[a] |
| BtH | Yampara-Iquise *et al.* (2008) | *B. thetaiota-omiron* a-1-6 mannanase | BtH-F, BtH-P, BtH-R | Genomic *B. thetaiota-omicron* DNA | Human, Sewage | Dogs, Beef Cattle, Dairy Cattle, Horses, Swine, Goose, Chickens, Turkeys | 1 ng fecal DNA | 9.3 copies | 100 | 100 |
| gyrB | Lee *et al.* (2010) | *B. fragilis* gyrB | Bf904F, Bf923MGB, Bf958R | genomic *B. fragilis* DNA | Human | Cow, Dog, Pig | 10 ng fecal DNA | $1.1 * 10^2$ copies | 100 | 97 |
| HF183 Endpoint | Bernhard and Field (2000) | *Bacteroides* 16S | HF183F, Bac708R | Plasmid | Human, Sewage | Cat, Cow, Deer, Dog, Duck, Elk, Goat, Llama, Pig, Seagull, Sheep | 2 to 4 ng fecal DNA | $1.4 * 10^{-6}$ g/L dry sewage | 88 | 100 |
| HF183 SYBR | Seurinck *et al.* (2005) | *Bacteroides* 16S | HF183F, HFsybR | Plasmid | Human, Sewage | Chicken, Cow, Dog, Horse, Pig | 2.2 mg wet wt | $2.8 * 10^2$ copies | 91 | NR[b] |
| HF183 Taqman | Haugland *et al.* (2010) | *Bacteroides* 16S | HF183F, BthetP1, BthetR1 | Plasmid | Human, Sewage | Cattle, Pig, Chicken, Dog, Cat | 1 ng fecal DNA | 10 copies | 100 | NR[c] |
| HumM2 | Shanks *et al.* (2009) | *B. fragilis* hypothetical protein BF3236 | HumM2F, HumM2P, HumM2R | Plasmid | Human, Sewage | Alpaca, Cow, Goat, Sheep, Horse, Pig, Antelope, Whitetail Deer, Mule deer, Moose, Elk, Canadian Goose, Duck, Pelican, Gull, Turkey, Chicken, Marine Dolphin, California Sea Lion, Cat, Dog | 1 ng fecal DNA | 10 copies | 100 | 99 |
| Mnif | Johnston *et al.* (2010) | *Methanobrevi-bacter smithii nifH* | Mnif202F, MnifP, Mnif353R | Genomic *M. smithii* DNA | Sewage | Gull, Ambient Seawater | 15 mg wet wt | 5 genome equiva-lents | 100 | 72 |

[a]Not reported; strong cross-reaction with cat, dog. [b]Not reported; cross-reaction with one chicken. [c]Not reported; weak cross-reaction with chicken, dog.

2007), *B. fragilis* group (Matsuki *et al.* 2002), and fecal *Bacteroides* (Converse *et al.* 2009). The fecal source characterizations presented in Ervin *et al.* (2013) were used for wet mass, *Enterococcus* CFU, *E. coli* CFU, and *E. coli* 23S qPCR. Total DNA mass data were obtained from the laboratories: a majority (13 of 17) measured total DNA concentrations on each filter with a NanoDrop spectrophotometer. When values for total DNA yield were reported as negative, "too low" or some other indication of data below the detection limit, a value of 1 ng/filter was substituted. Paired measurements (of human assay targets and DNA mass) per filter were used to normalize the data. Some laboratories measured and reported "general" qPCR assay characterizations of the samples, including Entero1A (5), GenBac3 (3), AllBac (1), BacUni-UCD (1), *B. fragilis* group (1) and fecal *Bacteroides* (1). The general assay data were used to normalize the human-associated qPCR data from those laboratories using paired measurements (of human and general assay targets) per filter.

For each of the above units of measure, the assay with the highest gene copy abundance among target samples (human stools, sewage and septage taken together) was considered the most sensitive, and the assay with the largest difference in median gene copy abundance between target and non-target samples was considered the most specific. Note that these performance metrics differ from those used in Boehm *et al.* (2013).

Because performance outcomes can change under different characterizations of fecal concentration, it was necessary to select a primary unit of measure by which to judge quantitative assay performance. This study focused on total DNA mass on each challenge filter as measured by NanoDrop spectrophotometry. Quantitative benchmarks for sensitivity and specificity were defined based on copies per nanogram of total DNA: an assay was quantitatively sensitive if the median abundance in every target source (human stools, sewage and septage considered separately) was greater than 10 copies/ng, and an assay was quantitatively specific if the interquartile ranges of copies/ng did not overlap between target and non-target sources.

To study the effect of challenge filter sample units of measure on the presence/absence performance metrics, a balanced subset of the data was selected, and an *in silico* dilution experiment was performed. This subset consisted of assays performed by the method developer's laboratory (BacH, BacHum-UCD, BsteriF1, gyrB, HumM2, HF183 Taqman and Mnif). In this subset, the assays were performed under optimal conditions (in the hands of their developer's laboratory), and the *n* for all assays was the same. Presence/absence method performance in this subset was evaluated using the same challenge filter sample units of measure that the method developers used when the assays were first published (Table 1). This was done by *in silico* dilution or addition of the appropriate amount of fecal material and calculating what the copy numbers would have been based on the observed amplification with the actual challenge filter samples. For this exercise, a limit of detection (LOD) of 10 copies per reaction was applied: amplification below this level was considered negative and anything above 10 copies was considered positive.

### Doubleton Analyses

To determine the effect of mixed fecal sources on assay performance, sensitivity was evaluated in the doubleton samples with respect to the non-human source present and the estimated relative contribution of total DNA from each source. The proportion of total DNA contribution from each fecal source was estimated using a mass ratio approach. The median NanoDrop measurements on the singleton samples were multiplied by the volumetric proportions used to create the doubleton samples (see Boehm *et al.* 2013 for sample creation details), and the ratio of target:non-target DNA on each doubleton filter was estimated from those values. Presence/absence sensitivity (with DNQ values considered positive) was calculated for every assay according to doubleton type and compared to the target:non-target DNA ratios.

## Results

### Performance by DNQ Classification

None of the assays met the 80% benchmark used by Boehm *et al.* (2013) and the USEPA (2005) for both specificity and sensitivity when DNQ was considered positive (Table 2). With DNQ negative, BtH, HF183 SYBR and HF183 Taqman met the benchmark for both sensitivity and specificity metrics. Assay sensitivity was high, but specificity was low when DNQ results were regarded as positive. All assays except HF183 endpoint and Mnif were at least 80% sensitive with DNQ positive. The only assay that was at least 80% specific with DNQ

positive was HF183 endpoint; however, HF183 SYBR, BacH and HumM2 were not appreciably behind the mark at 78, 77, and 75%, respectively. When DNQ was negative, sensitivity decreased in all assays except BacHum-UCD, and all assays were considered specific except BacHum-UCD, BsteriF1 and Mnif. Note that the results presented in Table 2 use a "per filter" characterization of presence/absence in the challenge filter samples and consider all laboratories' data together as one dataset.

## Performance by Challenge Filter Sample Unit of Measure

The presence/absence specificities of HF183 Taqman and BacHum-UCD under their developer's challenge filter sample units of measure were starkly different from the "per filter" specificity results (both 96% in developers' laboratory versus 46 and 37% across all laboratories, respectively, DNQ positive). In general, the assays performed well with their developers' execution and test sample quantities (Table 3). However, in the present study, the assays often performed worse than reported in their original publications (Table 1), except for BacHum-UCD under its original challenge sample units of measure (Table 3). Interestingly, BacHum-UCD was the only assay that showed excellent sensitivity using the units of 5000 copies of BacUni-UCD per reaction, which was the benchmark used to develop the BacHum-UCD assay (Kildare *et al.* 2007).

When the challenge filter samples were characterized by total DNA mass and all laboratories' data were analyzed together, HF183 Taqman was the only assay categorized as both quantitatively sensitive and specific (Figure 1). The four assays targeting functional genes (BtH, gyrB, HumM2 and Mnif) were less sensitive than the assays targeting the *Bacteroides* 16S rRNA gene, likely due to fewer copies of the functional genes per cell. All assays were considered quantitatively sensitive except BtH, HumM2, and Mnif, while only HF183 Taqman and BtH were considered specific. Dog was a frequent source of false positives: BacH, BacHum-UCD and BsteriF1 had cross-reactivity in dog samples at levels equivalent to that of sewage/septage (BacH, BacHum-UCD) or human stools (BsteriF1). BacH cross-reacted with the fewest number of non-human sources (only dog and deer).

When gene copy abundance of each quantitative assay in the singleton samples was normalized to all available fecal source units of measure, it was clear that which assay performed best was dependent on how the challenge samples and performance metrics were defined (Table 4). BacHum-UCD was the most sensitive assay using the total DNA mass, *E. coli* CFU, *E. coli* qPCR and GenBac3 measurements. BtH was the most sensitive assay using the *Enterococcus* qPCR copy units, but was less sensitive in other quantitative measures. BacH was the most sensitive assay only under the wet mass unit of measure, which was the same fecal unit used

**Table 2. Performance of human-associated assays in singleton samples among all labs, calculated with DNQ (detected, not quantifiable) samples as positive or negative, with presence/absence determined on a per-filter basis.**

| Assay | Sensitivity | | | Specificity | | |
|---|---|---|---|---|---|---|
| | Human *n*[a] | DNQ+ (%) | DNQ– (%) | Non-human *n*[a] | DNQ+ (%) | DNQ– (%) |
| BacH | 12 | 100 | 75 | 26 | 77 | 85 |
| BacHum-UCD | 72 | 97 | 97 | 156 | 37 | 67 |
| BsteriF1 | 48 | 100 | 96 | 104 | 44 | 61 |
| BtH | 12 | 100 | 92 | 26 | 54 | 96 |
| gyrB | 12 | 92 | 50 | 26 | 58 | 96 |
| HF183 Endpoint | 84 | 75 | NA | 182 | 96 | NA |
| HF183 SYBR | 48 | 100 | 92 | 104 | 78 | 89 |
| HF183 Taqman | 60 | 100 | 95 | 130 | 46 | 92 |
| HumM2 | 72 | 93 | 67 | 156 | 75 | 94 |
| Mnif | 60 | 78 | 60 | 130 | 68 | 76 |

[a]Values for *n* vary among assays because the methods were performed by different numbers of laboratories; see Sample Creation and Analysis in Methods section.

**Table 3. Sensitivity and specificity of human qPCR assays in singleton samples, calculated using original developer's data generated in this study and the developers' original challenge fecal sample units of measure. Developers' metrics are shown in bold.**

| Assay | Sensitivity n = 12[a] (%) | | | | | Specificity n = 26[b] (%) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 mg Wet Mass | 15 mg Wet Mass | 5000 Copies BacUni-UCD | 1 ng DNA | 10 ng DNA | 1 mg Wet Mass | 15 mg Wet Mass | 5000 Copies BacUni-UCD | 1 ng DNA | 10 ng DNA |
| BacH | **100** | 100 | 42 | 75 | 92 | **77** | 77 | 100 | 88 | 85 |
| BacHum-UCD | 100 | 100 | **100** | 100 | 100 | 62 | 54 | **96** | 65 | 65 |
| BsteriF1 | 100 | 100 | 0 | **100** | 100 | 46 | 42 | 92 | **77** | 58 |
| gyrB | 100 | 100 | 0 | 58 | **75** | 69 | 58 | 100 | 100 | **88** |
| HF183 Taqman | 100 | 100 | 17 | **100** | 100 | 62 | 42 | 96 | **96** | 73 |
| HumM2 | 100 | 100 | 0 | **58** | 83 | 92 | 81 | 100 | **100** | 92 |
| Mnif | 100 | **100** | 17 | 75 | 83 | 77 | **77** | 100 | 81 | 81 |

[a]Number of target (human stool, sewage or septage) samples in the analysis
[b]Number of non-target (non-human animal) samples in the analysis

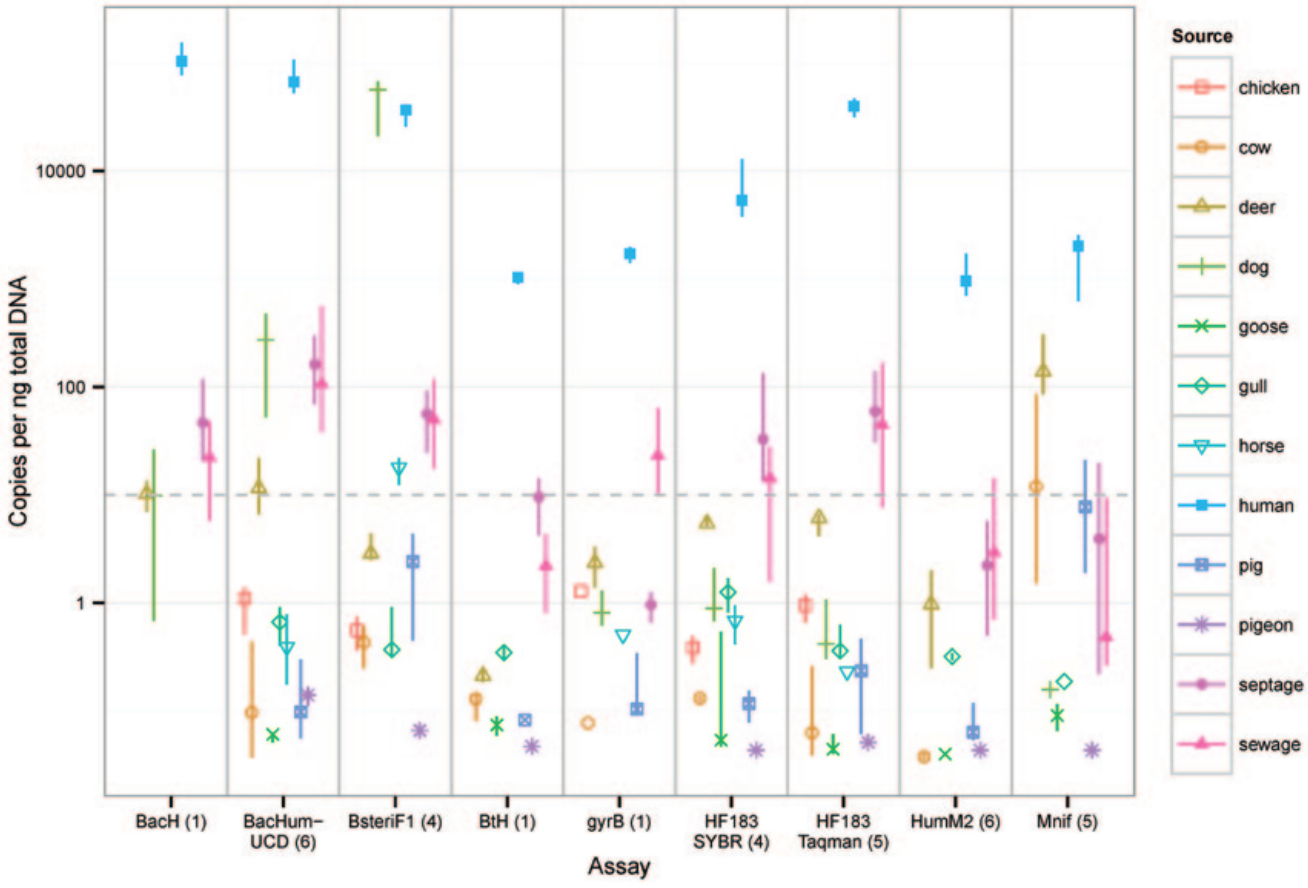


**Figure 1. Copies per nanogram total DNA in each fecal source for quantitative assays. Each point is the median value for a given source, and the bars represent the interquartile ranges (25th to 75th percentiles). Fecal sources are indicated by a unique combination of color and shape. The solid markers are "target" sources (human stools, sewage, or septage). The dashed horizontal line indicates 10 copies per nanogram, used as a benchmark of assay sensitivity. The number in parentheses after each name on the x-axis indicates the number of laboratories that performed the assay.**

to develop that assay. HF183 Taqman was the most specific assay in six of the seven units of measure where it was possible to make a comparison: milligrams of wet feces, mass of total DNA, *E. coli* CFU, *E. coli* qPCR, *Enterococcus* qPCR, and GenBac3. The only fecal source characterization for which HF183 Taqman was not the most specific assay was *Enterococcus* CFU, where BacHum-UCD excelled.

### Performance by Target Source

The sensitivity of each assay differed for each of the three "target" sources: human stools, sewage and septage. In almost every case, sensitivity was greatest in human stools, followed by septage, and least sensitive in sewage samples (the exception was gyrB, which had greater sensitivity in sewage than septage; Table 5). The presence/absence metrics were also greatly dependent on DNQ classification. With DNQ positive, five assays were perfectly sensitive (100%) to all three target sources: BacH, BsteriF1, BtH, HF183 SYBR and HF183 Taqman. No assay was 100% sensitive to all three targets with DNQ negative, though BacHum-UCD, BsteriF1, HF183 SYBR and HF183 Taqman met the 80% benchmark. In quantitative terms, every assay was orders of magnitude more sensitive (by copies per nanogram of total DNA) to human stools than to septage or sewage (Figure 1). BacH was the most sensitive of all assays to human stools, and BacHum-UCD was most sensitive of all assays to sewage and septage.

### Performance in Doubleton Samples

In the doubletons containing sewage, a decrease in target:non-target DNA ratio appeared to decrease sensitivity in the samples containing gull feces (bottom panel, Figure 2). This change was especially noticeable for Mnif, HumM2, gyrB and HF183 SYBR. Sensitivity was also low in the chicken/sewage samples for Mnif and HF183 endpoint. For the doubleton samples containing human stools (top panel), sensitivity remained consistently high; only HF183 endpoint showed a slight decrease in sensitivity at lower target:non-target DNA ratios. The doubletons containing septage consisted of only one sample type: 10% septage:90% horse. For these samples, all assays were considered sensitive except Mnif.

## Discussion

### DNQ Classification

The classification of DNQ samples as positive or negative dramatically affected the performance of the assays in presence/absence metrics, and this has important implications for local beach management applications. Changing the DNQ classification from positive to negative decreased sensitivity slightly, but improved specificity substantially for all assays (Table 2). The assays that were judged as acceptably sensitive or specific with the presence/absence metrics differed slightly from Boehm *et al.* (2013) because the present study considered only singletons and pooled all laboratories' data together. However, variable performance was observed among laboratories and this variability may skew the overall DNQ classification results (Figure SI-1). The variable results among laboratories may be a product of the different LLOQ values obtained by using different types and quantities of standard reference material, Cq threshold settings, and other variations in method protocols.

The treatment of DNQ samples in the MST literature is mixed. For example, some studies have defined DNQ values as negative (Stapleton *et al.* 2009), while others have regarded DNQ amplification as a positive detection (Kelty *et al.* 2012); still others established a lower limit of detection for the qPCR but did not differentiate between LLOQ and LOD (Sauer *et al.* 2011). For SYBR assays, most groups consider DNQ samples negative, due to the difficulty in validating melt curves for such low amplification. Information on DNQ handling is often not reported at all. To date, the present study and associated manuscripts (Raith *et al.* 2013, Sinigalliano *et al.* 2013) are the first to comprehensively examine the effects of varying DNQ classification on MST assay performance.

The present study found that assay specificity was superior when DNQ results were treated as negative. One possible explanation is that most human-associated genetic markers are not strictly found in human sources; instead, they are typically found at a higher abundance in human sources (Shanks *et al.* 2010). Thus, the more sensitive the method is, the more likely it is that cross-reactivity will be observed in the DNQ range. In environmental samples, DNQ measurements may result from dilution or degradation of a human fecal source or from cross-reactivity. Experts in the field have not yet reached consensus

**Table 4. Human-associated marker abundance under all available fecal source characterizations. Values are median (standard deviation) of $log_{10}$-transformed copy numbers across all labs. Bold indicates the most sensitive assay (largest copy abundance in target samples), and shading indicates the most the specific assay (largest difference in median copy abundance between target and non-target) for each unit of measure.**

| Assay | Source | *n* | MgWet[c] | DNA[d] | ENT Mf[e] | ENT qPCR[f] | *E. coli MF*[g] | *E. coli* qPCR[h] | GenBac3[i] | AllBac[j] | BacUni-UCD[k] | Bfrag[l] | FecalB[m] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BacH | target[a] | 12 | **7.5 (0.2)** | 1.9 (2) | 2.1 (1.8) | – | 1.5 (0.9) | 0.6 (0.9) | – | – | – | – | – |
| | non-target[b] | 26 | 2.8 (0.9) | 0.9 (0.9) | -1.5 (1.8) | – | -1.5 (1.2) | -2.2 (0.9) | – | – | – | – | – |
| BacHum-UCD | target | 72 | 7.1 (1) | **2.7 (1.7)** | **3 (1.3)** | -0.4 (1.5) | **1.9 (1.1)** | **1 (0.9)** | **-1.2 (0.4)** | – | -0.4 (0.2) | – | – |
| | non-target | 156 | 2.4 (1.7) | 0 (1.7) | -0.6 (1.3) | -2.9 (1.3) | -1.3 (1.8) | -2.2 (1.7) | -3.9 (1.9) | – | -3.8 (2.3) | – | – |
| BsteriF1 | target | 48 | 6.8 (0.2) | 2 (1.5) | 2.3 (1.1) | -0.3 (1.6) | 1.4 (0.6) | 0.5 (0.3) | -1.6 (0.5) | – | – | – | – |
| | non-target | 104 | 2.9 (2) | 0.8 (2) | 0.7 (1.4) | -1.8 (2.1) | -1.3 (2) | -2.2 (2.1) | -3.6 (1.8) | – | – | – | – |
| BtH | target | 12 | 5.3 (0.1) | 1 (1.4) | 1.4 (0.9) | **-0.2 (1.9)** | 0.2 (0.9) | -0.6 (0.5) | – | – | – | – | – |
| | non-target | 26 | 0.9 (0.7) | -0.9 (0.4) | -1.1 (1.2) | -2.8 (1.2) | -3.7 (1.6) | -4.3 (1.2) | – | – | – | – | – |
| gyrB | target | 12 | 5.2 (0.3) | 1.5 (1.4) | 0.8 (1.1) | – | 0.2 (0.6) | -1.1 (0.5) | – | – | – | -2 (0.7) | – |
| | non-target | 26 | 1.2 (0.8) | -0.3 (0.6) | -0.9 (1.5) | – | -2.8 (1.1) | -3.5 (0.8) | – | – | – | -3.4 (1.2) | – |
| HF183 SYBR | target | 48 | 5.9 (1.1) | 1.7 (1.5) | 1.9 (1.4) | -0.9 (1.5) | 0.5 (1.3) | -0.2 (1.2) | -2.2 (0.8) | -2.1 (0.3) | – | – | -1.5 (0.5) |
| | non-target | 104 | 2.3 (1) | -0.2 (0.8) | -0.9 (1.4) | -3.7 (1.3) | -2.1 (1.4) | -3.2 (1.3) | -5.4 (1.9) | -5.1 (NA)[n] | – | – | NA (NA)[o] |
| HF183 Taqman | target | 60 | 6.9 (0.1) | 2.2 (1.5) | 2.4 (1.1) | -0.3 (1.7) | 1.3 (0.6) | 0.5 (0.3) | -1.7 (0.6) | – | – | – | – |
| | non-target | 130 | 1.2 (0.9) | -0.5 (0.8) | -1.1 (1.5) | -4 (1.3) | -2.8 (1.4) | -3.2 (1.1) | -5.1 (2.1) | – | – | – | – |
| HumM2 | target | 72 | 5.3 (0.3) | 0.9 (1.4) | 1.1 (1) | -1.6 (1.7) | 0.2 (0.7) | -0.8 (0.5) | -2.9 (0.7) | – | – | – | – |
| | non-target | 156 | 0.8 (0.9) | -1.1 (0.7) | -0.9 (0.8) | -3.7 (1.3) | -2.6 (1.2) | -3.2 (0.7) | -6.2 (1.9) | – | – | – | – |
| Mnif | target | 60 | 5.7 (0.5) | 1.3 (1.6) | 2 (1.2) | -2.3 (2.1) | 0.2 (1.1) | -0.5 (0.9) | -3.1 (0.8) | – | – | – | – |
| | non-target | 130 | 3.4 (1.2) | 0.8 (1.3) | 1.6 (2.2) | -1.9 (2.2) | -0.7 (1.9) | -1.3 (1.7) | -4.5 (1.3) | – | – | – | – |

[a]human stools, sewage and septage; [b]non-human animals; [c]mg wet mass, sewage and septage samples excluded; [d]ng total DNA by NanoDrop; [e]EPA method 1600; [f]Entero1(A) (Haugland et al., 2005); [g]*E. coli* membrane filtration; [h]*E. coli* 23S qPCR assay EC23S857 (Chern et al., 2011); [i](Siefring et al., 2008); [j](Layton et al., 2006); [k](Kildare et al., 2007); [l]*B. fragilis* group specific (Matsuki et al., 2002); [m]Fecal *Bacteroides* (Converse et al., 2009); [n]*n* of samples with amplification was too small to calculate standard deviation; [o]No amplification was observed

**Table 5. Sensitivity of human-associated assays in singleton human, sewage and septage samples calculated with detected, not quantifiable (DNQ) values as positive or negative on a per-filter basis.**

| Assay | n[a] | Human (%) | | Sewage (%) | | Septage (%) | |
|---|---|---|---|---|---|---|---|
| | | DNQ+ | DNQ– | DNQ+ | DNQ– | DNQ+ | DNQ– |
| BacH | 4 | 100 | 100 | 100 | 50 | 100 | 75 |
| BacHum-UCD | 24 | 100 | 100 | 92 | 92 | 100 | 100 |
| BsteriF1 | 16 | 100 | 100 | 100 | 88 | 100 | 100 |
| BtH | 4 | 100 | 100 | 100 | 75 | 100 | 100 |
| gyrB | 4 | 100 | 100 | 100 | 50 | 75 | 0 |
| HF183 endpoint | 28 | 96 | NA | 57 | NA | 71 | NA |
| HF183 SYBR | 16 | 100 | 100 | 100 | 81 | 100 | 94 |
| HF183 Taqman | 20 | 100 | 100 | 100 | 85 | 100 | 100 |
| HumM2 | 24 | 100 | 100 | 83 | 46 | 96 | 54 |
| Mnif | 20 | 95 | 95 | 55 | 20 | 85 | 65 |

[a]Number of singleton samples in each target source (varies by number of laboratories running each assay).

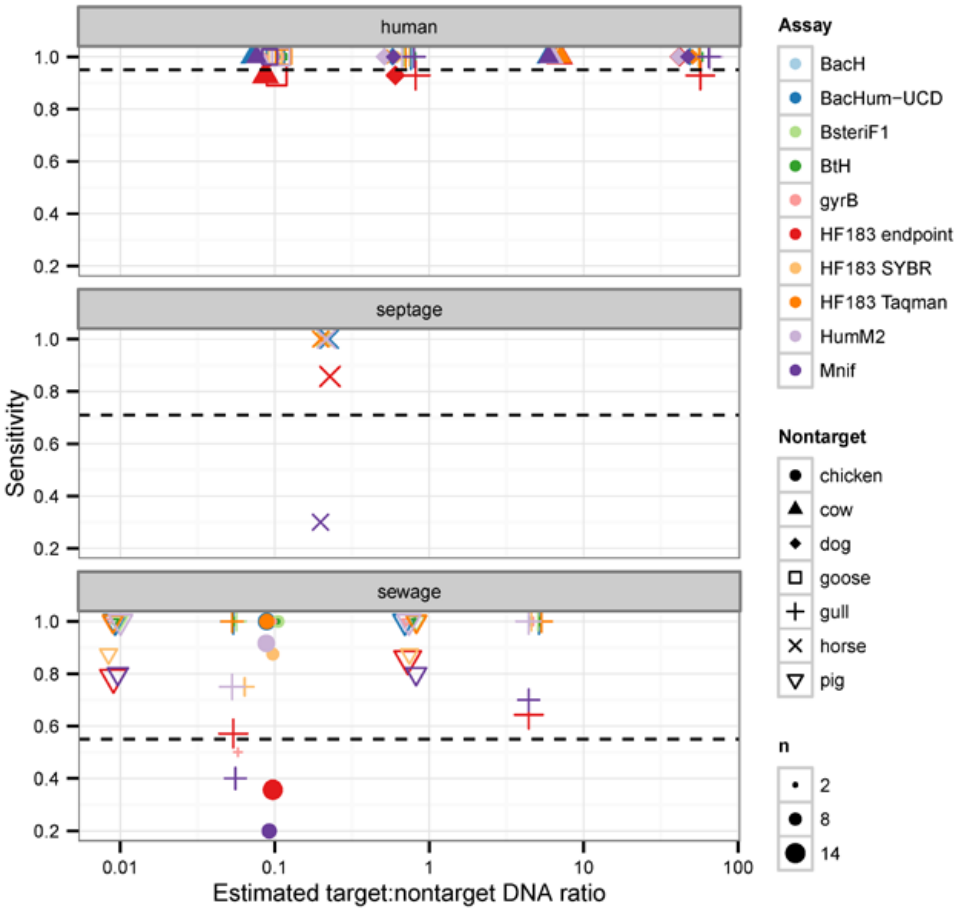


**Figure 2. Presence/absence per-filter sensitivity (DNQ positive) in doubletons versus the estimated target:non-target DNA ratio. The three target sources present in the doubletons are organized into the horizontal panels. The shape of each point indicates the non-target source in the doubleton, and the assays are differentiated by colors. The size of each point indicates the number of measurements that were used to calculate the sensitivity value for that point, which ranged from 2 to 14. The horizontal positions of the points were "jittered" to make more of the data visible. The dashed lines represent the lowest sensitivity of any assay in the target singletons (DNQ positive).**

regarding how to classify DNQ results obtained in MST field studies (Stewart *et al.* 2013). In practice, it may be beneficial to perform both a human bacteria-associated assay (highly sensitive, less specific) and a human viral assay (highly specific, less sensitive); however, the optimal method for concentrating human viruses from environmental water samples is yet to be determined (see Harwood *et al.* 2013).

## Challenge Filter Sample Units of Measure

The amount of fecal matter on a filter can be described using several units of measure; this study found that changing the challenge filter sample units of measure can change which assays performed best. The relative quantities of fecal material in the study's challenge filter samples changed considerably among fecal sources when different units of measure were used to describe fecal concentrations. For example, one fecal source may have low *Enterococcus* levels, but a high wet mass compared to another source (Ervin *et al.* 2013). Accordingly, the assay that performed best on a "per unit" basis depended on which unit of measure was used. Data from all laboratories were used in this analysis, even though there were differing sizes of data sets and clear laboratory-to-laboratory variation, and study results should be interpreted with those factors in mind. To date, the present study and associated manuscripts (Boehm *et al.* 2013; Raith *et al.* 2013; Sinigalliano *et al.* 2013) are the first to examine the effects of changing fecal units of measure on the performance outcomes of molecular MST methods.

The present study focused on total DNA mass as the primary challenge filter sample unit of measure. The amount of fecal matter varied substantially from filter to filter, both within a given fecal source and across sources (Ervin *et al.* 2013), yet it was not possible to directly quantify the fecal material on each filter for every unit of measure. Total DNA mass was the only unit of measure with measurements on individual filters using the same quantification method from a majority of laboratories (832 total DNA measurements). Further, because total DNA yield varied extensively among laboratories (Figure SI-2), normalizing to total DNA minimized bias and put the assays on the most level playing field possible (see Kelty *et al.* 2012).

When the data were normalized to ng total DNA, HF183 Taqman was the only assay categorized as both sensitive and specific. Numerous studies from around the world have demonstrated the success of the original HF183 endpoint PCR assay (e.g., Griffith *et al.* 2003, Gawler *et al.* 2007, Ahmed *et al.* 2012), and many qPCR assays have been developed to target the same region of the *Bacteroides* 16S rRNA gene (see Supplemental Information). In the present work, the HF183 endpoint assay was much less sensitive to sewage than the HF183 qPCR assays (Table 5), suggesting that a qPCR version of this method is preferable where sewage contamination is a concern.

## Influence of Target Source

Assay sensitivity varied among the three "target" sources (human stools, sewage or septage). While every assay was highly or perfectly sensitive to human stools, success was more varied with sewage and septage sources. There are several possible explanations. Firstly, the sewage and septage challenge filter samples had very low quantities of fecal material compared to the human stool samples (Ervin *et al.* 2013), which affects sensitivity on a "per filter" basis. Secondly, sewage and septage are mixed sources with fecal inputs from humans as well as other animal species. The mixed nature of these sources could affect sensitivity both in terms of decreased amount of target per unit of fecal material as well as possible cross-reactivity to the non-human inputs. Lastly, the assays in this study were initially designed to be human fecal-associated, not necessarily sewage or septage-associated (though many were validated with sewage samples). Given the differences between stool samples and sewage/septage, there is a need for methods that can discriminate sewage and septage, such as community analysis (Cao *et al.* 2013). In consideration of the differences among target sources, Table 4 was recreated with the stool and sewage/septage target sources analyzed separately (Table SI-4).

Almost every assay had lower sensitivity to sewage than septage (Table 5, DNQ negative). Septic tanks may be more hospitable environments for fecal anaerobes than sewerage systems, and thus the microbes targeted by the assays in this study may be more numerous in septage samples than sewage. This hypothesis is supported by the greater copy numbers of general Bacteroidales per total DNA mass observed in septage versus sewage samples (Figure SI-3). Further, it has been previously shown that septage has higher concentrations than sewage of *E.coli* uidA, *Enterococcus* 16S rRNA and BtH gene copies (Srinivasan *et al.* 2011), and that only a small

percentage of microorganisms in sewage are fecal-derived (McLellan *et al.* 2010).

The present study's findings contrast with those of some of the original assay publications. For example, Kildare *et al.* (2007) found the BacHum-UCD marker to be less prevalent in human stools than wastewater samples, though gene copy abundance in these sources was not reported. Similarly, researchers in France found HF183 SYBR to be less prevalent in stools than wastewater (Mauffret *et al.* 2012). In the present study, HumM2 and HF183 Taqman were orders of magnitude less sensitive to sewage but substantially more abundant in human stools than reported by Shanks *et al.* (2009, 2010).

The discrepancies between the present study's results and those of previous studies may due to differences in the wastewater samples. In the present work, a relatively small number ($n = 9$) of treatment plants were sampled, some of which receive industrial wastewater (up to 20% of total input volume and as much as 50% during certain times of day; C. McGee, pers. comm.). Stapleton *et al.* (2009) found several orders of magnitude fewer gene copies of human *Bacteroides* in industrial wastewater compared to sewage. In addition, the microbial community present in the sewerage infrastructure (biofilms) may be quite different among locations due to a number of factors. Thus, it is possible that the microbial profile of the sewage influent used in the present study may vary considerably from those found elsewhere, which could explain some of the contrasting results. Before these methods are employed in local MST studies, management agencies may benefit from performing small studies to establish the assays' sensitivity to the wastewater sources present in their watersheds.

### Doubletons

In environmental water samples, there will be numerous sources of bacterial DNA, including multiple fecal hosts and indigenous microbes. The doubleton challenge samples represent an idealized model of very a simple two-host system. The present study's analysis uncovered an interesting effect of gull feces on sensitivity to sewage (Figure 2). It appears that gull feces decreased sensitivity to sewage in several assays, yet this effect was not observed in the sewage/pig samples or in the human/gull samples. To date, no other method evaluation studies have tested these assays against a sewage/gull matrix. This finding has implications for application of these assays at beaches with large native seagull populations; however, not every assay was affected and the number of samples in this category was relatively small. Further study is needed before definite recommendations can be made on this issue.

### Effect of Individual Laboratory Performance

An important source of variability in assay performance is the effect of individual laboratories, whether due to differing protocols or varying levels of experience with the technology. The sources of inter-laboratory variability in assay performance include: DNA purification approach and efficiency, DNA yield measurements, qPCR chemistry, type of qPCR standard reference material, qPCR instrument, laboratory infrastructure (i.e., spatial separation of tasks) and technician skill level. Even with standardized protocols, laboratories may produce different results (Pan *et al.* 2010). In the present study, there are several instances of assay performance differing across laboratories (Figure SI-1). These differences are often driven by DNQ classification, which reflects the varying LLOQ values among laboratories (Table SI-3). Issues associated with repeatability among laboratories in this study are explored in depth in Ebentier *et al.* (2013). It is clear that SOPs, which should include everything from laboratory setup to data handling and stringent quality assurance guidelines, must be established for accurate performance assessment and successful implementation of these methods.

Another important limitation of this work is the imbalance in size of the data sets between assays (ranging from one to seven laboratories), which creates a statistical bias in the performance metrics. This bias is exacerbated by the clear laboratory-to-laboratory variability in performance. This bias and variability make it difficult to compare performance metrics across assays, and could be why the overall performance in the present study often does not match the original reports. Method performance needs to be determined with an unbiased data set where laboratory-to-laboratory variability is not a factor (e.g., Table 3).

## Literature Cited

Ahmed, W., N. Masters and S. Toze. 2012. Consistency in the host specificity and host sensitivity of the *Bacteroides* HF183 marker for sewage pollution tracking. *Letters in Applied Microbiology* 55:283-289.

Bernhard, A.E. and K. Field. 2000. A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. *Applied and Environmental Microbiology* 66:4571-4574.

Boehm, A.B. and J.A. Soller. 2012. Recreational water risk: Pathogens and fecal indicators. pp. 8758-8770 *in*: R.A. Meyers (ed.), Encyclopedia of Sustainability Science and Technology. Springer. New York, NY.

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* 47:6812-6828.

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* 55:611-622.

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

Chern, E.C., S. Siefring, J. Paar, M. Doolittle and R.A. Haugland. 2011. Comparison of quantitative PCR assays for *Escherichia coli* targeting ribosomal RNA and single copy genes. *Letters in Applied Microbiology* 52:298-306.

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

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

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

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

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

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

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* 47:6929-6943.

Haugland, R., S. Siefring, L. Wymer, K. Brenner and A. 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.

Haugland, R.A., M. Varma, M. Sivaganesan, C. Kelty, L. Peed and O.C. Shanks. 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* 33:348-357.

Johnston, C., J.A. Ufnar, J.F. Griffith, J.A. Gooch and J.R. Stewart. 2010. A real-time qPCR assay for the detection of the *nifH* gene of

*Methanobrevibacter smithii*, a potential indicator of sewage pollution. *Journal of Applied Microbiology* 109:1946-1956.

Kelty, C.A., M. Varma, M. Sivaganesan, R.A. Haugland and O.C. Shanks. 2012. Distribution of genetic marker concentrations for fecal indicator bacteria in sewage and animal feces. *Applied and Environmental Microbiology* 78:4225-4232.

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

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

Lee, C.S. and J. Lee. 2010. Evaluation of new gyrB-based real-time PCR system for the detection of *B. fragilis* as an indicator of human-specific fecal contamination. *Journal of Microbiological Methods* 82:311-318.

Matsuki, T., K. Watanabe, J. Fujimoto, Y. Miyamoto, T. Takada, K. Matsumoto, H. Oyaizu and R. Tanaka. 2002. Development of 16S rRNA-gene-targeted group-specific primers for the detection and identification of predominant bacteria in human feces. *Applied and Environmental Microbiology* 68:5445-5451.

Mauffret, A., M.P. Caprais and M. Gourmelon. 2012. Relevance of Bacteroidales and F-specific RNA bacteriophages for efficient fecal contamination tracking at the level of a catchment in France. *Applied and Environmental Microbiology* 78:5143-5152.

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

Pan, Y., L. Bodrossy, P. Frenzel, A.G. Hestnes, S. Krause, C. Luke, 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* 76:7451-7458.

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

Reischer, G., D. Kasper, R. Steinborn, A. Farnleitner and R. Mach. 2007. A quantitative real-time PCR assay for the highly sensitive and specific detection of human faecal influence in spring water from a large alpine catchment area. *Letters in Applied Microbiology* 44:351-356.

Roslev, P. and A.S. Bukh. 2011. State of the art molecular markers for fecal pollution source tracking in water. *Applied Microbiology and Biotechnology* 89:1341-1355.

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

Seurinck, S., T. Defoirdt, W. Verstraete and S. Siciliano. 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* 7:249-259.

Shanks, O.C., C.A. Kelty, M. Sivaganesan, M. Varma and R.A. Haugland. 2009. Quantitative PCR for genetic markers of human fecal pollution. *Applied and Environmental Microbiology* 75:5507-5513.

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

Siefring, S., M. Varma, E. Atikovic, L. Wymer and R.A. Haugland. 2008. Improved real-time PCR assays for the detection of fecal indicator bacteria in surface waters with different instrument and reagent systems. *Journal of Water and Health* 6:225-237.

Sinigalliano, C.D., J.S. Ervin, L.C. Van De Werfhorst, B.D. Badgley, E. Ballesté, J. Bartkowiak, A.B. Boehm, M.N. Byappanahalli, K.D. Goodwin, M. Gourmelon, J.F. Griffith, P.A. Holden, J.A. Jay, B.A. Layton, C. Lee, J. Lee, W.G. Meijer, R.T. Noble, M.R. Raith, H. Ryu, M.J. Sadowsky, A. Schriewer, D. Wang, D. Wanless, R.L. Whitman, S. Wuertz and J.W. Santo Domingo. 2013. Multi-laboratory evaluations of the performance of *Catellicoccus marimammalium* PCR assays developed to target gull fecal sources. *Water Research* 47:6883-6896.

Srinivasan, S., A. Aslan, I. Xagoraraki, E. Alocilja and J.B. Rose. 2011. *Escherichia coli*, enterococci, and *Bacteroides thetaiotaomicron* qPCR signals through wastewater and septage treatment. *Water Research* 45:2561-2572.

Stapleton, C.M., D. Kay, M.D. Wyer, C. Davies, J. Watkins, C. Kay, A.T. McDonald, J. Porter nad A. Gawler. 2009. Evaluating the operational utility of a Bacteroidales quantitative PCR-based MST approach in determining the source of faecal indicator organisms at a UK bathing water. *Water Research* 43:4888-4899.

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* 47:6829-6838.

United States Environmental Protection Agency (USEPA). 2005. Microbial Source Tracking Guide Document. EPA/600-R-05-064. USEPA Office of Research and Development, National Risk Management Research Laboratory. Cincinnati, OH.

Yampara-Iquise, H., G. Zheng, J.E. Jones and C.A. Carson. 2008. Use of a *Bacteroides thetaiotaomicron*-specific alpha-1-6, mannanase quantitative PCR to detect human faecal pollution in water. *Journal of Applied Microbiology* 105:1686-1693.

## Acknowledgements

The authors are deeply grateful to Meredith Raith for her vital contribution to sample creation, data generation and data analysis. Steve Weisberg and Chuck Hagedorn helped develop the manuscript. The authors also greatly appreciate the students and staff of all laboratories that participated in the SIPP method evaluation study. Work in the Farnleitner Laboratory was funded by the Austrian Science Fund (FWF) projects # P22309-B20 and DK-plus #W1219-N22. Work in the Rose Laboratory was funded by the National Oceanic and Atmospheric Administration (NA04OAR4600199). Funding for the method evaluation study was through a Clean Beach Initiative Grant from the California State Water Resources Control Board. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the US Government.

## Supplemental Information

Supplemental Information is available at ftp://ftp.sccwrp.org/pub/download/DOCUMENTS/AnnualReports/2013AnnualReport/ar13_445_459SI.pdf.