2013

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

Blythe A. Layton
Southern California Coastal Water Research Project

Yiping Cao
Southern California Coastal Water Research Project

Darcy L. Ebentier
UCLA Civil and Environmental Engineering

Kaitlyn Hanley
UCLA Civil and Environmental Engineering

Elisenda Balleste
University College Dublin

See next page for additional authors

Authors
Blythe A. Layton, Yiping Cao, Darcy L. Ebentier, Kaitlyn Hanley, Elisenda Balleste, Joao Brandao, Muruleedhara Byappanahalli, Reagan Converse, Andreas H. Farnleitner, Jennifer Gentry-Shields, Maribeth L. Gidley, Michele Gourmelon, Chang Soo Lee, Jiyoung Lee, Solen Lozach, Tania Madi, Wim G. Meijer, Rachel Noble, Lindsay Peed, Georg H. Reischer, Raquel Rodrigues, Joan B. Rose, Alexander Schriewer, Chris Sinigalliano, Sangeetha Srinivasan, Jill Stewart, Laurie C. Van De Werfhorst, Dan Wang, Richard Whitman, Stefan Wuertz, Jenny Jay, Patricia A. Holden, Alexandria B. Boehm, Orin Shanks, and John F. Griffith

WATER RESEARCH 47 (2013) 6897–6908

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/watres

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

Blythe A. Layton [a], Yiping Cao [a], Darcy L. Ebentier [b], Kaitlyn Hanley [b], Elisenda Ballesté [c], João Brandão [d], Muruleedhara Byappanahalli [e], Reagan Converse [f,g], Andreas H. Farnleitner [h], Jennifer Gentry-Shields [i], Maribeth L. Gidley [j], Michèle Gourmelon [k], Chang Soo Lee [l], Jiyoung Lee [l], Solen Lozach [k], Tania Madi [m], Wim G. Meijer [c], Rachel Noble [f], Lindsay Peed [n], Georg H. Reischer [h], Raquel Rodrigues [d], Joan B. Rose [o], Alexander Schriewer [p], Chris Sinigalliano [q], Sangeetha Srinivasan [o], Jill Stewart [i], Laurie C. Van De Werfhorst [r], Dan Wang [s], Richard Whitman [e], Stefan Wuertz [p,t], Jenny Jay [b], Patricia A. Holden [r], Alexandria B. Boehm [s], Orin Shanks [n], John F. Griffith [a,*]

[a] Southern California Coastal Water Research Project, 3535 Harbor Blvd Ste 110, Costa Mesa, CA 92626, United States
[b] UCLA Civil and Environmental Engineering, 5732H Boelter Hall, Los Angeles, CA 90095, United States
[c] UCD School of Biomolecular and Biomedical Science, University College Dublin, Dublin 4, Ireland
[d] Instituto Nacional de Saúde, Dr. Ricardo Jorge, Avenida Padre Cruz, 1649-016 Lisboa, Portugal
[e] U.S. Geological Survey, Great Lakes Science Center, Lake Michigan Ecological Research Station, 1100 N. Mineral Springs Road, Porter, IN 46304, United States
[f] UNC Chapel Hill Institute of Marine Sciences, 3431 Arendell Street, Morehead City, NC 28557, United States
[g] Oak Ridge Institute for Science and Education, PO Box 117, Oak Ridge, TN 37831, United States
[h] Environmental Microbiology and Molecular Ecology Group, Interuniversity Cooperation Centre for Water and Health, Institute of Chemical Engineering, Vienna University of Technology, Gumpendorfer Strasse 1a/166-5-2, A-1060 Vienna, Austria
[i] Department of Environmental Sciences and Engineering, University of North Carolina, 135 Dauer Drive, Chapel Hill, NC 27599-7431, United States
[j] The Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, FL 33149, United States
[k] Laboratoire de Microbiologie, MIC/LNR, Département Ressources Biologiques et Environnement, Unité Environnement, Microbiologie et Phycotoxines, Ifremer, ZI Pointe du diable, Plouzané, France
[l] College of Public Health, Division of Environmental Health Science, The Ohio State University, 406 Cunz Hall, 1841 Neil Avenue, Columbus, OH 43210, United States
[m] Source Molecular Corporation, 4985 SW 74th Court, Miami, FL 33155, United States
[n] U.S. Environmental Protection Agency, Office of Research and Development, National Risk Management Research Laboratory, 26 West Martin Luther King Drive, Cincinnati, OH 45268, United States
[o] Department of Fisheries and Wildlife, 13, Natural Resources, Michigan State University, East Lansing, MI 48824, United States
[p] Department of Civil & Environmental Engineering, University of California, Davis, One Shields Avenue, Davis, CA 95616, United States

* Corresponding author. Tel.: +1 714 755 3228.
E-mail address: johng@sccwrp.org, jfgriffith@icloud.com
0043-1354/$ – see front matter
http://dx.doi.org/10.1016/j.watres.2013.05.060

[q] National Oceanic and Atmospheric Administration, Atlantic Oceanographic and Meteorological Laboratory, 4301 Rickenbacker Causeway, Miami, FL 33149, United States
[r] Earth Research Institute and Bren School of Environmental Science & Management, University of California, Santa Barbara, CA 93106, United States
[s] Stanford University, Dept. of Civil & Environmental Engineering, 473 Via Ortega, Stanford, CA 94305, United States
[t] Singapore Centre on Environmental Life Sciences Engineering (SCELSE), School of Biological Sciences SBS-B1N-27 and School of Civil and Environmental Engineering, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Singapore

ARTICLE INFO

Article history:
Received 2 October 2012
Received in revised form
7 May 2013
Accepted 10 May 2013
Available online 5 July 2013

Keywords:
Microbial source tracking
qPCR
Water quality
Bacteroides
Bacteroidales

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. Here, we evaluated ten of these methods (BacH, BacHum-UCD, *Bacteroides thetaiotaomicron* (BtH), BsteriF1, gyrB, HF183 endpoint, HF183 SYBR, HF183 Taqman®, HumM2, and *Methanobrevibacter smithii nifH* (Mnif)) 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.

1. 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, 2011), 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 United States Environmental Protection Agency's new 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 – as positive versus non-detect – 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 1 mg of wet feces or 1 ng 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.

2. Methods

2.1. 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 (Section 1 and Tables S1–S2). 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.

2.2. Classification of DNQ

The presence/absence sensitivity and specificity metrics were calculated in two ways: once with DNQ (detected, 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 S3. 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).

2.3. 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, *Escherichia 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., 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

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 (per reaction)	LLOQ (quantity per 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* α-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 and Lee (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	2–4 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)	*Methanobrevibacter smithii nifH*	Mnif202F, MnifP, Mnif353R	Genomic *M. smithii* DNA	Sewage	Gull, ambient seawater	15 mg wet wt	5 genome equivalents	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.

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. We chose to focus on total DNA mass on each challenge filter as measured by NanoDrop spectrophotometry. We defined quantitative benchmarks for sensitivity and specificity 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, we chose a balanced subset of the data and performed an *in silico* dilution experiment. This subset consisted of assays run 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 lab) 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.

2.4. *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 contributions 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.

3. Results

3.1. *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.

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 Section 2.1.

3.2. *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' lab versus 46% and 37% across all labs, respectively, DNQ positive). In general, the assays performed well with their developers' execution and test sample quantities (Table 3). However, in our 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

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' own metrics are shown in bold.

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

together, HF183 Taqman was the only assay categorized as both quantitatively sensitive and specific (Fig. 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

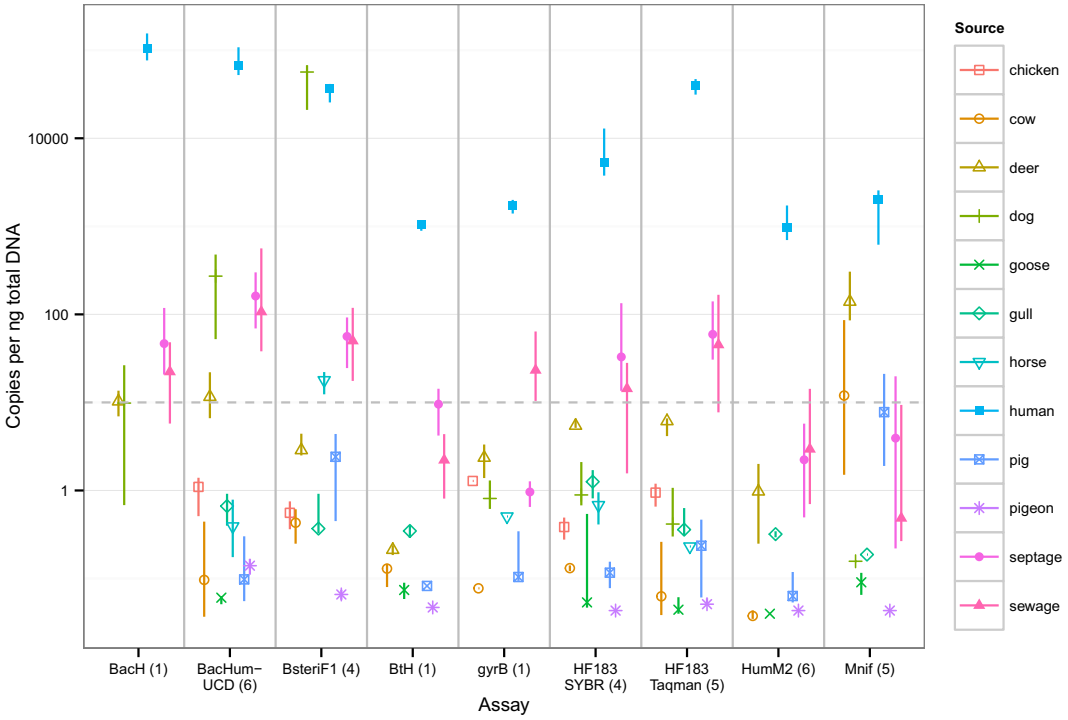


Fig. 1 – Copies per nanogram total DNA in each fecal source for quantitative assays. Each point is the median value for a given source, and the bars represent the interquartile ranges (25th to 75th percentiles). Fecal sources are indicated by a unique combination of color and shape. The solid markers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

Table 4 – Human-associated marker abundance under all available fecal source characterizations. Values are median (standard deviation) of log10-transformed copy numbers across all labs. Underline indicates the most sensitive assay (largest copy abundance in target samples) and bold 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).

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 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.

3.3. *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

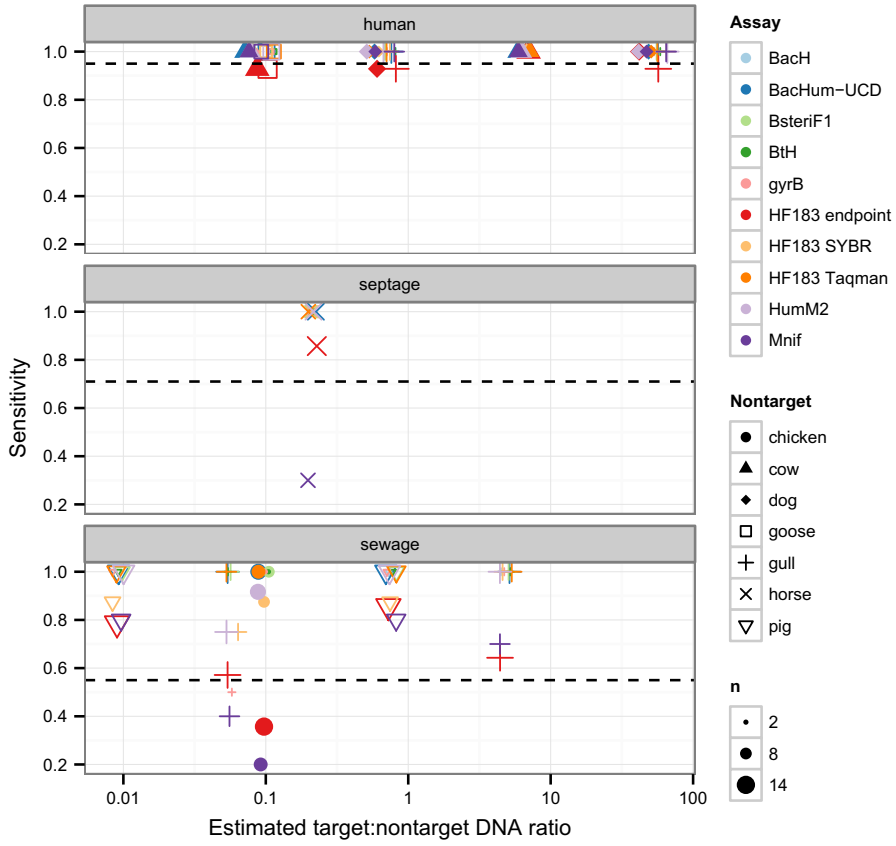


Fig. 2 – Presence/absence per-filter sensitivity (DNQ positive) in doubletons versus the estimated target:non-target DNA ratio. The three target sources present in the doubletons are organized into the horizontal panels. The shape of each point indicates the non-target source in the doubleton, and the 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).

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 (Fig. 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.

3.4. *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, Fig. 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.

4. Discussion

4.1. *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 here we considered only singletons and pooled all labs' data together. However, variable performance was observed among laboratories and this variability may skew the overall DNQ classification results (Fig. S1). 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 our knowledge, this work and the other manuscripts from the present study (Raith et al., 2013; Sinigalliano et al., 2013) are the first to comprehensively examine the effects of varying DNQ classification on MST assay performance.

In our analysis, we 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 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).

4.2. *Challenge filter sample units of measure*

The amount of fecal matter on a filter can be described using several units of measure, and we found that changing the challenge filter sample units of measure can change which assays performed best. The relative quantities of fecal material in our 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 labs were used in this analysis, even though there were differing sizes of datasets and clear lab-to-lab variation (see Section 4.6), and our results should be interpreted with those factors in mind. To our knowledge, the present work and other manuscripts from this study (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.

We 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 S2), 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.

4.3. *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 S4).

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 (Fig. S3). 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).

Our 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 our results and those of previous studies may due to differences in the wastewater samples. In the present work, we sampled a relatively small number of treatment plants ($n = 9$), 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.

4.4. *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. In our analysis, we uncovered an interesting effect of gull feces on sensitivity to sewage (Fig. 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 our knowledge, 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.

4.5. *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 (Fig. S1). These differences are often driven by DNQ classification, which reflects the varying LLOQ values among laboratories (Table S3). 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 datasets between assays (ranging from 1 to 7 laboratories), which creates a statistical bias in the performance metrics. This bias is exacerbated by the clear lab-to-lab 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 dataset where lab-to-lab variability is not a factor (e.g. Table 3).

5. Conclusions

- HF183 Taqman consistently excelled across numerous performance benchmarks
- In practice, it may be beneficial to use two assays targeting different genes and/or bacterial species, such as HF183 Taqman with BtH or HumM2
- Further work is needed to determine whether the additional uncertainty associated with using multiple human-associated assays adds value to source tracking efforts
- While these assays performed well with fresh fecal pollution sources from California in an artificial water source, several issues still need to be thoroughly addressed prior to implementation in local management settings: persistence and decay, standardization of protocols, performance with reference feces from different geographical areas and animal species, and potential influence of the environmental sample matrix on amplification

Acknowledgments

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. We also greatly appreciate the students and staff of all labs that participated in the SIPP method evaluation study. Work in the Farnleitner Lab was funded by the Austrian Science Fund (FWF) projects # P22309-B20 and DK-plus #W1219-N22. Work in the Rose Lab 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. This article is Contribution 1758 of the U.S. Geological Survey Great Lakes Science Center. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.watres.2013.05.060.

REFERENCES

Ahmed, W., Masters, N., Toze, S., 2012. Consistency in the host specificity and host sensitivity of the *Bacteroides* HF183 marker for sewage pollution tracking. Letters in Applied Microbiology 55 (4), 283–289.

Bernhard, A.E., Field, K., 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 (10), 4571–4574.

Boehm, A.B., Soller, J.A., 2011. Risks associated with recreational waters: pathogens and fecal indicators. In: Meyers, R.A. (Ed.), Encyclopedia of Sustainability Science and Technology. Springer.

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

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clinical Chemistry 55 (4), 611–622.

Cao, Y., Van De Werfhorst, L.C., Dubinsky, E.A., Badgley, B.D., Sadowsky, M.J., Andersen, G.L., Griffith, J.F., Holden, P.A., 2013. Evaluation of molecular community analysis methods for discerning fecal sources and human waste. Water Research 47 (18), 6849–6861.

Chern, E.C., Siefring, S., Paar, J., Doolittle, M., Haugland, R.A., 2011. Comparison of quantitative PCR assays for *Escherichia coli* targeting ribosomal RNA and single copy genes. Letters in Applied Microbiology 52 (3), 298–306.

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

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

Ervin, J.S., Russell, T.L., Layton, B.A., Yamahara, K.M., Wang, D., Sassoubre, L.M., Cao, Y., Kelty, C.A., Sivaganesan, M., Boehm, A.B., Holden, P.A., Weisberg, S.B., Shanks, O.C., 2013. Characterization of fecal concentrations in human and other animal sources by physical, culture, and quantitative real-time PCR methods. Water Research 47 (18), 6873–6882.

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

Gawler, A.H., Beecher, J.E., Brandão, J., Carroll, N.M., Falcão, L., Gourmelon, M., Masterson, B., Nunes, B., Porter, J., Rincé, A., Rodrigues, R., Thorp, M., Walters, J.M., Meijer, W.G., 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 (16), 3780–3784.

Griffith, J.F., Weisberg, S., McGee, C., 2003. Evaluation of microbial source tracking methods using mixed fecal sources in aqueous test samples. Journal of Water and Health 1 (7), 141–151.

Harwood, V.J., Boehm, A.B., Sassoubre, L.M., Vijayavel, K., Stewart, J.R., Fong, T.-T., Caprais, M.P., Converse, R.R., Diston, D., Ebdon, J., Fuhrman, J.A., Gourmelon, M., Gentry-Shields, J., Griffith, J.F., Kashian, D.R., Noble, R.T., Taylor, H., Wicki, M., 2013. Performance of viruses and bacteriophages for fecal source determination in a multi-laboratory, comparative study. Water Research 47 (18), 6829–6943.

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

Haugland, R.A., Varma, M., Sivaganesan, M., Kelty, C., Peed, L., 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.

Johnston, C., Ufnar, J.A., Griffith, J.F., Gooch, J.A., Stewart, J.R., 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 (6), 1946–1956.

Kelty, C.A., Varma, M., Sivaganesan, M., Haugland, R.A., Shanks, O.C., 2012. Distribution of genetic marker concentrations for fecal indicator bacteria in sewage and animal feces. Applied and Environmental Microbiology 78 (12), 4225–4232.

Kildare, B.J., Leutenegger, C.M., McSwain, B.S., Bambic, D.G., Rajal, V.B., 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.

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R., Sayler, G., 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 (6), 4214–4224.

Lee, C.S., Lee, J., 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 (3), 311–318.

Matsuki, T., Watanabe, K., Fujimoto, J., Miyamoto, Y., Takada, T., Matsumoto, K., Oyaizu, H., Tanaka, R., 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 (11), 5445–5451.

Mauffret, A., Caprais, M.P., Gourmelon, M., 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 (15), 5143–5152.

McLellan, S.L., Huse, S.M., Mueller-Spitz, S.R., Andreishcheva, E.N., Sogin, M.L., 2010. Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent. Environmental Microbiology 12 (2), 378–392.

Pan, Y., Bodrossy, L., Frenzel, P., Hestnes, A.G., Krause, S., Luke, C., Meima-Franke, M., Siljanen, H., Svenning, M.M., Bodelier, P.L.E., 2010. Impacts of inter- and intralaboratory variations on the reproducibility of microbial community analyses. Applied and Environmental Microbiology 76 (22), 7451–7458.

Raith, M.R., Kelty, C.A., Griffith, J.F., Schriewer, A., Wuertz, S., Mieszkin, S., Gourmelon, M., Reischer, G.H., Farnleitner, A.H., Ervin, J., Holden, P.A., Jay, J.A., Boehm, A.B., Rose, J.B., Meijer, W.G., Sivaganesan, M., Shanks, O.C., 2013. Comparison of PCR and quantitative real-time PCR methods for the characterization of ruminant and cattle fecal pollution sources. Water Research 47 (18), 6921–6928.

Reischer, G., Kasper, D., Steinborn, R., Farnleitner, A., Mach, R., 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 (4), 351–356.

Roslev, P., Bukh, A.S., 2011. State of the art molecular markers for fecal pollution source tracking in water. Applied Microbiology and Biotechnology 89 (5), 1341–1355.

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

Seurinck, S., Defoirdt, T., Verstraete, W., Siciliano, S., 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.

Shanks, O.C., Kelty, C.A., Sivaganesan, M., Varma, M., Haugland, R.A., 2009. Quantitative PCR for genetic markers of human fecal pollution. Applied and Environmental Microbiology 75 (17), 5507–5513.

Shanks, O.C., White, K., Kelty, C.A., Sivaganesan, M., Blannon, J., Meckes, M., Varma, M., Haugland, R.A., 2010. Performance of PCR-based assays targeting Bacteroidales genetic markers of human fecal pollution in sewage and fecal samples. Environmental Science and Technology 44 (16), 6281–6288.

Siefring, S., Varma, M., Atikovic, E., Wymer, L., Haugland, R.A., 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 (2), 225–237.

Sinigalliano, C.D., Ervin, J.S., Van De Werfhorst, L.C., Badgley, B.D., Ballesté, E., Bartkowiak, J., Boehm, A.B., Byappanahalli, M.N., Goodwin, K.D., Gourmelon, M., Griffith, J.F., Holden, P.A., Jay, J.A., Layton, B.A., Lee, C., Lee, J., Meijer, W.G., Noble, R.T., Raith, M.R., Ryu, H., Sadowsky, M.J., Schriewer, A., Wang, D., Wanless, D., Whitman, R.L., Wuertz, S., Santo Domingo, J.W., 2013. Multi-laboratory evaluations of the performance of *Catellicoccus marimammalium* PCR assays developed to target gull fecal sources. Water Research 47 (18), 6883–6896.

Srinivasan, S., Aslan, A., Xagoraraki, I., Alocilja, E., Rose, J.B., 2011. *Escherichia coli*, enterococci, and *Bacteroides thetaiotaomicron* qPCR signals through wastewater and septage treatment. Water Research 45 (8), 2561–2572.

Stapleton, C.M., Kay, D., Wyer, M.D., Davies, C., Watkins, J., Kay, C., McDonald, A.T., Porter, J., Gawler, A., 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 (19), 4888–4899.

Stewart, J.R., Boehm, A.B., Dubinsky, E.A., Fong, T.-T., Goodwin, K.D., Griffith, J.F., Noble, R.T., Shanks, O.C., Vijayavel, K., Weisberg, S.B., 2013. Recommendations following a multi-laboratory comparison of microbial source tracking methods. Water Research 47 (18), 6829–6938.

USEPA, 2005. Microbial Source Tracking Guide Document (No. EPA/600-R-05-064). U.S. Environmental Protection Agency, Office of Research and Development, National Risk Management Research Laboratory, Cincinnati, OH.

Yampara-Iquise, H., Zheng, G., Jones, J.E., Carson, C.A., 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 (5), 1686–1693.

Performance of Human Fecal-Associated PCR-Based Assays in a Multi-Laboratory Method Evaluation Study (Layton et al.): Supplemental Information

1. Standard Operating Procedures (SOPs) used by "core" SIPP laboratories

1.1. Five laboratories involved in planning the SIPP study (SCCWRP, Stanford University, UC Santa Barbara, UCLA, and US EPA/Shanks, collectively known as the "core" labs) agreed to standardize their protocols for DNA extraction, DNA quantification, q/PCR, and data analysis. These SOPs were shared with all participating laboratories, but adherence to these SOPs was optional. All core labs and four "outside" laboratories used the GeneRite DNA-EZ ST kit for DNA extraction. All core labs and eight outside laboratories used a NanoDrop spectrophotometer for DNA quantification. Four core labs also quantified DNA using a fluorescence-based kit (Quant-iT), but only NanoDrop data were used in the analysis for this manuscript. Detailed protocols for q/PCR are found in Table S1. Manufacturer and vendor information for q/PCR reagents are found in Table S2. The SIPP SOP for using the FlashGel system to visualize conventional PCR products is given in Section 1.2 below. Details of data analysis are described in Ebentier et al. (this issue). Briefly, all core laboratory qPCR data (from both human and general assays) were quantified using lab-specific standard curves. These lab-specific curves were generated by pooling data from standard curves run on each qPCR plate, as well as a "master" standards plate containing triplicates of four separate dilution series. The lower limit of quantification (LLOQ, copies per reaction) was defined as the lowest point on the standard curve for which >50% of the replicates amplified. When samples had amplification in <3 PCR triplicates, a Cq of 40 was substituted for the non-detect replicates and the sample mean of 3 Cq values was used to quantify the copies present in the sample. Outside lab qPCR data were used as-is (as reported by each lab) in this manuscript, with the exception of BLOQ samples. For outside labs that reported BLOQ as quantified copy numbers or non-detects, a LLOQ was determined from their reported standard curves and BLOQ samples were treated as DNQ (see also Section 2 below).

Table S1. SOPs for quantitative and conventional PCR used by core SIPP laboratories. dNTP = deoxynucleoside triphosphate, rxn = reaction, BSA = bovine serum albumin. (see attached Excel file)

Table S2. Ordering information for (q)PCR reagents used by core SIPP labs

Item	Manufacturer	Part number	Specifications	Vendor
Primers	Operon	N/A	resuspend in DNase-free water	Operon
Probes	Operon	N/A	resuspend in TE pH 8.0 (from MoBio), put in as many 1000 pmol aliquots as possible	Operon
qPCR MM	Applied Biosystems	4318157	AB Universal MM	Life Technologies
Conventional	Takara	RR001AM	$MgCl_2$ separate	
BSA	GIBCO	15260-037		Invitrogen
DNAse free water	MoBio	100371-020		VWR
qPCR core kit for SYBR® Assay ROX	Eurogentec	RT-SN10-05		Anaspec.com
AE Buffer	Qiagen	19077		Qiagen

1.2. FlashGel System SOP

All conventional PCR products were visualized using Lonza FlashGel 16+1 double tier cassettes (2.2% agarose gel, VWR cat# 95015-624).

1. Remove the white seals on the gel cassette (leave the vent seals intact).
2. Rinse the gel cassette with dI water (TAE buffer is acceptable as well) and remove any excess liquid from the gel cassette by tilting the cassette and blotting. Do not touch the wells; they should have a bubble of liquid over them.
3. Insert the cassette into the gel dock.
4. Load samples by mixing 2uL sample with 2uL Lonza loading dye (5X loading dye VWR catalog number: 95015-630) and inject into well.
5. Load 4uL of ladder (DNA marker (100bp-4kb) VWR catalog number: 95015-632) into one well (other size ladders can be used (i.e. 1kb); they all worked well in preliminary tests.
6. Set the power source to 275V and allow the gel to run for ~3.5 minutes (Note: Single tier

cassettes can run up to 7 min and double tier cassettes up to 5 min).

7. Pictures can be taken anytime during the gel electrophoresis and it is recommended to take one picture every minute since strong UV supply will potentially degrade PCR products.

2. Variable assay performance across laboratories

Table S3. Lower limit of quantification (LLOQ, in units of copies per filter) for each lab and assay. The lab "names" are random. BLOQ = below limit of quantification; DNQ = detected, not quantifiable; ND = not detected. Much of the information shown is a summary of data presented in Boehm et al. (this issue); the LLOQ values calculated in this work and not previously reported are shown in bold.

Assay	Lab	Standard reference material	LLOQ, copies/filter	BLOQ reported
BacH	B	Circular plasmid	230	DNQ
BacHum-UCD	A	Circular plasmid	50	DNQ
	G	Circular plasmid	644	DNQ
	H	Circular plasmid	**500**	copy numbers
	J	Circular plasmid	100	DNQ
	L	Circular plasmid	3850	DNQ
	N	Circular plasmid	2404	ND
BsteriF1	A	Linear plasmid	50	DNQ
	G	Linear plasmid	1857	DNQ
	J	Linear plasmid	1000	DNQ
	M	Linear plasmid	957	DNQ
BtH	F	genomic DNA	700	DNQ
gyrB	D	genomic DNA	20000	ND
HF183 SYBR	C	Linear plasmid	250	DNQ
	E	Synthesized oligo	100	DNQ
	G	Linear plasmid	387	DNQ
	M	Linear plasmid	991	DNQ
HF183 Taqman	A	Linear plasmid	50	DNQ
	G	Linear plasmid	2222	DNQ
	J	Linear plasmid	1000	DNQ
	L	Linear plasmid	2948	DNQ
	M	Linear plasmid	1057	DNQ
HumM2	A	Linear plasmid	50	DNQ
	G	Linear plasmid	22486	DNQ
	H	Linear plasmid	**500**	copy numbers
	J	Linear plasmid	1000	DNQ
	L	Linear plasmid	3407	DNQ
	M	Linear plasmid	1113	DNQ
Mnif	G	genomic DNA	2450	DNQ
	H	genomic DNA	**313**	copy numbers

	I	genomic DNA	**1000**	copy numbers
	K	genomic DNA	200	ND
	M	genomic DNA	3282	DNQ

Figure S1: Presence/absence sensitivity (top row) and specificity (bottom row) in singleton samples under both DNQ classifications (columns) in each lab. Some data are obscured by multiple points at the same location. The lab "names" are random. (see attached PDF)

3. Variable DNA extraction yield among laboratories

Figure S2(a). DNA yield according to DNA extraction kit and quantification method for all labs involved in the method evaluation study (Boehm et al., this issue). Note that the two laboratories that used a MoBio kit and a Qubit fluorometer reported lower DNA yield than all other labs. (see attached PDF)

Figure S2(b). DNA yield among labs measured by NanoDrop. The lab "names" in this figure are random. (see attached PDF)

4. Table 4 (main text) recreated with human stool and wastewater (sewage/septage) target sources considered separately

Table S4. Human-associated marker abundance under all available fecal source characterizations. Values are median (standard deviation) of log10-transformed copy numbers across all labs. See footnotes to Table 4. The assay with the highest gene copy abundance in a target source (most sensitive) for each unit of measure is shown in bold. The assay with the lowest copy abundance in non-human sources (most specific) is shaded. Note that this definition of "specific" is different than what is used in Table 4.

Assay	Source	*n*	MgWet	DNA	ENT MF	ENT qPCR	*E. coli* MF	*E. coli* qPCR	GenBac	AllBac	BacUni	Bfrag	FecalB
BacH	human stool	4	**7.5 (0.2)**	**5 (0.4)**	**4.8 (0.2)**	–	**1.5 (0.2)**	**1 (0.2)**	–	–	–	–	–
	wastewater	8	–	1.5 (0.9)	1.7 (1)	–	1.3 (1.1)	0.2 (1)	–	–	–	–	–
	non-human	26	2.8 (0.9)	0.9 (0.9)	-1.5 (1.8)	–	-1.5 (1.2)	-2.2 (0.9)	–	–	–	–	–
BacHum-UCD	human stool	24	7.1 (1)	4.8 (1.2)	4.4 (1)	2.3 (0.4)	1.1 (1)	0.6 (1)	**-1 (0.3)**	–	-0.4 (0.1)	–	–
	wastewater	48	–	**2.1 (1)**	**2.5 (0.9)**	-0.7 (0.4)	**2.3 (1)**	**1 (0.9)**	**-1.3 (0.4)**	–	-0.4 (0.2)	–	–
	non-human	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	human stool	16	6.8 (0.2)	4.6 (0.2)	4.2 (0.2)	2.4 (0.4)	0.9 (0.2)	0.4 (0.2)	-1.2 (0.3)	–	–	–	–
	wastewater	32	–	1.7 (0.5)	2.1 (0.4)	-0.9 (0.5)	1.7 (0.6)	0.6 (0.4)	-1.8 (0.4)	–	–	–	–
	non-human	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	human stool	4	5.3 (0.1)	3 (0.2)	2.6 (0.1)	**2.8 (0.5)**	-0.7 (0.1)	-1.2 (0.1)	–	–	–	–	–
	wastewater	8	–	0.6 (0.8)	1.1 (0.5)	**-0.5 (0.8)**	0.8 (0.7)	-0.4 (0.4)	–	–	–	–	–
	non-human	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	human stool	4	5.2 (0.3)	3.2 (0.1)	2.6 (0.3)	–	-0.7 (0.3)	-1.2 (0.3)	–	–	–	-1.7 (0.7)	–
	wastewater	8	–	0.9 (0.9)	0.8 (0.5)	–	0.2 (0.4)	-0.5 (0.6)	–	–	–	-2.2 (0.7)	–
	non-human	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	human stool	16	5.9 (1.1)	3.7 (0.8)	3.3 (1.1)	1.4 (1.5)	0 (1.1)	-0.6 (1.1)	-2.5 (1.1)	-2 (0.2)	–	–	-1 (0.5)
	wastewater	32	–	1.3 (0.8)	1.4 (1.2)	-1.2 (0.7)	1 (1.3)	0 (1.2)	-2.1 (0.6)	-2.1 (0.3)	–	–	-1.6 (0.5)
	non-human	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)	–	–	NA (NA)
HF183 Taqman	human stool	20	6.9 (0.1)	4.6 (0.2)	4.3 (0.1)	2.5 (0.6)	1 (0.1)	0.5 (0.1)	-1.1 (0.4)	–	–	–	–
	wastewater	40	–	1.7 (0.7)	2.1 (0.5)	-0.6 (0.9)	1.6 (0.6)	0.7 (0.4)	-1.8 (0.6)	–	–	–	–
	non-human	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	human stool	24	5.3 (0.3)	3 (0.5)	2.6 (0.3)	0.9 (0.4)	-0.7 (0.3)	-1.2 (0.3)	-2.7 (0.3)	–	–	–	–
	wastewater	48	–	0.5 (0.7)	0.8 (0.4)	-2.3 (1)	0.4 (0.5)	-0.5 (0.5)	-3 (0.8)	–	–	–	–
	non-human	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	human stool	20	5.7 (0.5)	3.3 (0.7)	3.1 (0.5)	0.8 (0.5)	-0.3 (0.5)	-0.8 (0.5)	-2.5 (0.5)	–	–	–	–
	wastewater	40	–	0.1 (0.9)	1.5 (1)	-3.3 (0.6)	1.2 (1.1)	-0.1 (0.9)	-3.5 (0.6)	–	–	–	–
	non-human	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)	–	–	–	–

5. Abundance of general qPCR gene copies per ng total DNA

Figure S3: General assay copy abundance normalized by total DNA mass. Each point is the median value for a given fecal source, and the bars represent the interquartile ranges (25^{th} to 75^{th} percentiles). Fecal sources are indicated by a unique combination of color and shape. The number in parenthesis after the assay name is the number of labs that ran the assay. The fecal *Bacteroides* qPCR assay data are excluded because no filter-matched NanoDrop total DNA measurements were available.

6. *Bacteroides* 16S rRNA gene and oligonucelotide sequence analysis

6.1. Goal: to examine each *Bacteroides* 16S rRNA gene assay's *in silico* specificity by comparing the primer and probe sequences with the human and non-human *Bacteroides* 16S rRNA reference gene sequences available in the National Center for Biotechnology Information (NCBI) database.

6.2. *Bacteroides* primer/probe sequence analysis methods

All data analyses were conducted in Geneious Pro 5.6.3 (Drummond et al., 2012). We used sequence data available from NCBI (the Gene and Nucleotide databases, accessed via Geneious on June 12, 2012) to examine the sequence specificity of the *Bacteroides* 16S rRNA gene assays. The primers and probes were aligned with 1) the reference sequences used to develop the assays, 2) the sequences corresponding to accession numbers reported in the assay developers' papers and 3) 75 other human-associated *Bacteroides* spp. 16S rRNA sequences, including those from fully sequenced genomes and ATCC reference type strains. The database was also searched for non-human fecal *Bacteroides* 16S rRNA sequences corresponding to the fecal sources used in the challenge filter samples. The primers and probes were tested for specificity against the non-human sequences, with a maximum of 2 base mismatches allowed.

6.3. Primer/probe alignment and mismatch results

The alignment of human fecal-associated *Bacteroides* 16S rRNA gene sequences submitted by the assay developers' labs with the related primers and probes reveal a fair amount of overlap in

the amplicon regions of these assays (Figure S4). For example, the positions of the BacHum-UCD and forward and reverse primers are nearly identical to the HF183 SYBR primers. Indeed, the following primer/probes overlap one another within the same 71-bp region: BacH-pC, BacH-pT, HF183F, BacHum160f, BsteriF1DE, BacHr, and BacHum193p (see Table 1 for oligo naming conventions, taken from the original publication of the assays whenever possible). The most unique primers and probes – those with no overlap with any other oligo in the alignment – are BacHf, BthetP1, BthetR1, and Bac708R. While previously the *Bacteroides* species associated with the HF183 marker was unknown (Bernhard and Field, 2000a), it now appears that all of the above assays except BsteriF1 target *B. dorei* (Figure S4).

The sequence specificity of each primer and probe was tested against all available non-human fecal-associated *Bacteroides* 16S rRNA gene sequences (n = 645) that corresponded to the fecal hosts in this study. These data included sequences from cow (n=40), chicken (65), dog (74), goose (110), gull (324), horse (19) and pig (13) hosts. Note that no sequences were found for pigeon or deer. A maximum of two base pair mismatches was allowed in the analysis. The BthetP1, BthetR1, HFsybR, BacHum241r and Bac708R sequences clearly target conserved regions of the 16S gene; these oligos had *in silico* binding – often without any mismatches – in the vast majority of non-human animal sequences (summarized in Table S3). Thus it is the forward primers that confer human specificity for these assays (BacHum-UCD, BsteriF1, HF183 Taqman, HF183 SYBR and HF183 endpoint). The HF183F and BacHum160f primers matched only one cow sequence, with two mismatches on the 3' end. The BsteriF1 forward primer matched one cow and one dog sequence, each with one mismatch on the 3' end. The BacH assay had the most specific primer/probe set according to this analysis: the forward primer and both probes matched a few cow (n=2), dog (7) and gull (2) sequences, while the reverse primer had no nonspecific matches at all. Accordingly, BacH also had the least cross-reactivity to the non-human challenge samples (Figure 1). Unfortunately there were not enough fecal-associated sequences in the database to perform similar analyses for the functional gene assays (BtH, gyrB, HumM2, Mnif).

6.4. Primer/probe mismatch discussion

The most distinguishing feature of every PCR-based assay is the DNA sequences of the primers and probes, which play an important role in the host-specificity and sensitivity of the assay. All of the *Bacteroides* 16S rRNA gene-based assays in this study targeted the exact same region of the gene – originally identified by Bernhard & Field (2000b; 2000a) – and several assays shared nearly identical primer and probe regions (Figure S4). Given these similarities, one might expect the assays to perform in a similar, if not identical fashion, but this was not quite the case. Performance differences among nearly-identical assays can be due to annealing temperature, salt concentration, mastermix chemistry, thermalcycler platform, and so forth. The finding that *in silico* the probe and reverse primer of HF183 Taqman assay showed binding with hundreds of non-human sequences, yet was one of the most specific assays in the study, indicates that the HF183 forward primer is robustly human-associated. The conserved nature of the reverse primer and probe regions may also contribute to the increased sensitivity that HF183 enjoys.

While there were some parallels between the *in silico* sequence matching and the qPCR results (e.g. BsteriF1 and dog), we expected to see more matches between the oligonucleotides and non-human sequences given the number of hosts that showed cross-reactivity. However, this analysis was limited by the number of non-human sequences available in the NCBI database. For example, we were not able to compare any sequences from deer. This is unfortunate, as deer was a source of false-positive results for every assay, sometimes at levels equivalent to sewage or septage (Figure 1). Some contamination was detected in a small percentage of filter blanks and sample processing controls (see Boehm et al., SI), so it may be that the deer samples were contaminated with a human source during the sample collection or creation process. As such, it would have been ideal to sequence the amplicons from reactions that generated false positive results had we had unlimited resources.

Table S5: Summary of nonspecific *in silico* primer binding (≤2 mismatches) in *Bacteroides* 16S rRNA gene-based assays

Assay	Oligonucleotide	Match
BacH	BacHf	Gull, Cow, Dog
	BacH-pC/BacH-pT	Gull, Dog, Cow
	BacHr	N/A
BacHum-UCD	BacHum160f	Cow
	BacHum193p	Pig, Gull, Goose, Dog, Cow, Chicken
	BacHum241r	Pig, Horse, Gull, Goose, Dog, Cow, Chicken
BsteriF1	BsteriF1DE	Dog, cow
	BthetP1	Pig, Horse, gull goose, dog, cow, chicken
	BthetR1	Pig, Horse, gull goose, dog, cow, chicken
HF183 Taqman	HF183F	Cow
	BthetP1	Pig, Horse, gull goose, dog, cow, chicken
	BthetR1	Pig, Horse, gull goose, dog, cow, chicken
HF183 SYBR	HF183F	Cow
	HFsybR	Pig, Horse, gull goose, dog, cow, chicken
HF183 endpoint	HF183F	Cow
	Bac708R	Pig, horse, gull, goose, dog, cow, chicken

Figure S4. Unique human-associated Bacteroides 16S rRNA gene sequences from NCBI (including assay developers' sequences and reference genomes) aligned with the primers and probes used in this study over the region of the assay amplicons. The sequence that the original HF183 assay was based on (Accession # AF233408) was used as the reference sequence in the alignment. (see attached PDF)

7. Positive and negative predictive values

In addition to sensitivity and specificity, positive and negative predictive values (PPV and NPV) were calculated under both DNQ classifications each assay (singletons only). PPV and NPV were defined as: PPV = true positives/(true positives + false positives) and NPV = true negatives/(true negatives + false negatives). All three target sources (human stools, sewage and septage) were counted as true positives. These metrics responded similarly to the different classifications of DNQ samples as sensitivity and specificity (Table S4). Applying the same 80% criteria as used for sensitivity and specificity, only HF183 endpoint met the benchmark for PPV with DNQ positive. However, with DNQ negative, five assays met the 80% mark for PPV: BtH, gyrB, HF183 SYBR, HF183 Taqman, and HumM2. All assays met or exceeded 80% for NPV under both DNQ classifications. It is important to recognize that the PPV and NPV metrics are valid only for this dataset and cannot be extrapolated to environmental samples, because our challenge filter set does not reflect the prevalence of human fecal contamination found in the environment (Altman and Bland, 1994).

Table S6. Positive and negative predictive values of the singletons calculated with DNQ samples treated as positive or negative.

		Positive predictive value		Negative predictive value	
Assay	n	DNQ+	DNQ-	DNQ+	DNQ-
BacH	38	67%	69%	100%	88%
BacHum-UCD	228	41%	58%	97%	98%
BsteriF1	152	45%	53%	100%	97%
BtH	38	50%	92%	100%	96%
gyrB	38	50%	86%	94%	81%
HF183 endpoint	266	90%	NA	89%	NA
HF183 SYBR	152	68%	80%	100%	96%
HF183 Taqman	190	46%	84%	100%	98%
HumM2	228	63%	84%	96%	86%
Mnif	190	53%	54%	87%	80%

8. References

Altman, D.G., Bland, J.M., 1994. Diagnostic tests 2: Predictive values. BMJ : British Medical Journal 309 (6947), 102.

Bernhard, A.E., Field, K., 2000a. 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 (10), 4571–4574.

Bernhard, A.E., Field, K., 2000b. Identification of Nonpoint Sources of Fecal Pollution in Coastal Waters by Using Host-Specific 16S Ribosomal DNA Genetic Markers from Fecal Anaerobes. Applied and Environmental Microbiology 66 (4), 1587–1594.

Boehm, A.B., Van De Werfhorst, L.C., Griffith, J.F., Holden, P.A., Jay, J.A., Shanks, O.C., Wang, D., Weisberg, S.B., n.d. Performance of forty-one microbial source tracking methods: a twenty-seven lab evaluation study. Water Research (in press).

Drummond, A.J., Ashton, B., Buxton, S., Cheung, M., Cooper, A., Duran, C., Field, M., Heled, J., Kearse, M., Markowitz, S., Moir, R., Stones-Havas, S., Sturrock, S., Thierer, T., Wilson, A., 2012. Geneious v5.6.

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

Table S1. SOPs for quantitative and conventional PCR conditions used by "core" SIPP laboratories. dNTP = deoxynucleoside triphosphate, rxn = reaction, BSA = bovine serum albumin. See Table S2 for vendor information

Assay	Developer reference	Target gene	oligo names	Reference material (standards)	rxn volume (µl)	template DNA volume (µl per rxn)	Enzyme and buffer (units per rxn)	Mastermix additives (units per rxn)	F primer µM per rxn	R primer µM per rxn	Probe µM per rxn	thermal cycling conditions	Bio-Rad CFX96 threshold setting	ABI StepOnePlus threshold setting	PCR product visualization
BacH[1]	Reischer et al. (2007)	*Bacteroides* 16S	BacHf, BacH-pT, BacH-pC, BacHr	plasmid	25	2	12.5 µl iQ Supermix, contains hot-start iTaq™ DNA polymerase, dNTPs and buffer	BSA 0.4 µg/µl	0.2	0.2	0.1	95ºC for 3 min, 50 cycles of (95ºC for 15 sec, 61ºC for 15 sec, 72ºC for 45 sec), read at 61ºC anneal	N/A[3]	N/A[3]	N/A
BacHum-UCD	Kildare et al. (2007)	*Bacteroides* 16S	BacHum160f, BacHum193p, BacHum241r	plasmid	25	2	12.5 µl Applied Biosystems Taqman™ Universal Mastermix (cat# 4318157)	BSA 0.05 mg/ml	0.4	0.4	0.08	AB-Uni[2]	100 RFU	0.03	N/A
BsteriF1	Haugland et al. (2010)	*B. stericoris* 16S	BsteriF1DE, BthetP1, BthetR1	plasmid	25	2	12.5 µl Applied Biosystems Taqman™ Universal Mastermix (cat# 4318157)	BSA 0.05 mg/ml	1	1	0.08	AB-Uni	100 RFU	0.08	N/A
BtH[1]	Yampara-Iquise et al. (2008)	*B. thetaiotaomiron* a-1-6 mannanase	BtH-F, BtH-P, BtH-R	genomic *B. thetaiotaomicron* DNA	20	2	10 µl LightCycler Taqman 480 Probe Master Mix (5X)	BSA 0.2 mg/ml	0.2	0.2	0.1	95ºC for 15 min, 45 cycles of (95ºC for 15 sec, 60ºC for 1 min, 72ºC for 5 sec), 40ºC for 30 sec, read during 72ºC step	N/A[4]	N/A[4]	N/A
gyrB[1]	Lee et al. (2010)	*B. fragilis* gyrB	Bf904F, Bf923MGB, Bf958R	genomic *B. fragilis* DNA	25	5	12.5 µl Applied Biosystems Taqman™ Universal Mastermix (cat# 4318157)	N/A	0.5	0.5	0.25	AB-Uni	N/A	automatic threshold determination	N/A
HF183 endpoint	Bernhard & Field (2000)	*Bacteroides* 16S	HF183F, Bac708R	plasmid	25	2	0.625 U TaKarRa ExTaq (cat# RR001AM), 2.5 µl 10X TaKaRa PCR buffer	200 µM each dNTP, 2 mM MgCl2, 0.05 mg/ml BSA	0.2	0.2	N/A	94ºC for 2 min, 35 cycles of (94ºC for 30 sec, 63ºC for 45 sec, 72ºC for 45 sec), 72ºC for 7 min	N/A	N/A	FlashGel, see SOP in Supp Info section 1.2. Correct product is 525 bp
HF183 SYBR	Seurinck et al. (2005)	*Bacteroides* 16S	HF183F, HFsybR	plasmid	25	2	0.625 U Hot GoldStar DNA Polymerase from Eurogentec qPCR core kit for Sybr® Green I (cat #RT-SN10-05), 2.5 µl Eurogentec real-time PCR 10X Buffer (MgCl2-free)	200 µM each dNTP, 2 mM MgCl2, 0.75 µl Sybr® Green I	0.25	0.25	N/A	50ºC for 2 min, 95ºC for 10 min, 40 cycles of (95ºC for 30 sec, 53ºC for 1 min, 60ºC for 1 min)	100 RFU	0.03	melt curve with ramping from 60-94.8ºC at 0.4ºC per 10 sec
HF183 Taqman	Haugland et al. (2010)	*Bacteroides* 16S	HF183F, BthetP1, BthetR1	plasmid	25	2	12.5 µl Applied Biosystems Taqman™ Universal Mastermix (cat# 4318157)	BSA 0.05 mg/ml	1	1	0.08	AB-Uni	100 RFU	0.08	N/A
HumM2	Shanks et al. (2009)	*B. fragilis* hypothetical protein BF3236	HumM2F, HumM2P, HumM2R	plasmid	25	2	12.5 µl Applied Biosystems Taqman™ Universal Mastermix (cat# 4318157)	BSA 0.05 mg/ml	1	1	0.08	AB-Uni	100 RFU	0.08	N/A
Mnif	Johnston et al. (2010)	*Methanobrevi-bacter smithii nifH*	Mnif202F, MnifP, Mnif353R	genomic *M. smithii* DNA	25	2	12.5 µl of 2x Quantitect Mastermix (Qiagen cat# 204343)	N/A	0.5	0.5	0.12	50ºC for 2 min, 95ºC for 15 min, 45 cycles of (95ºC for 1 sec then 50ºC for 1 min)	100 RFU	0.03	N/A

[1]not run by core labs, SOP reported by developer lab

[2]AB-Uni: 50ºC for 2 min, 95ºC for 10 min, 40 cycles of (95ºC for 15 sec then 60ºC for 1 min)

[3]developer lab used Eppendorf Mastercycler® ep realplex (Eppendorf AG, Hamburg, Germany); threshold settings not reported.

[4]developer lab used Roche LightCycler 480 Real-Time PCR System; threshold settings not reported

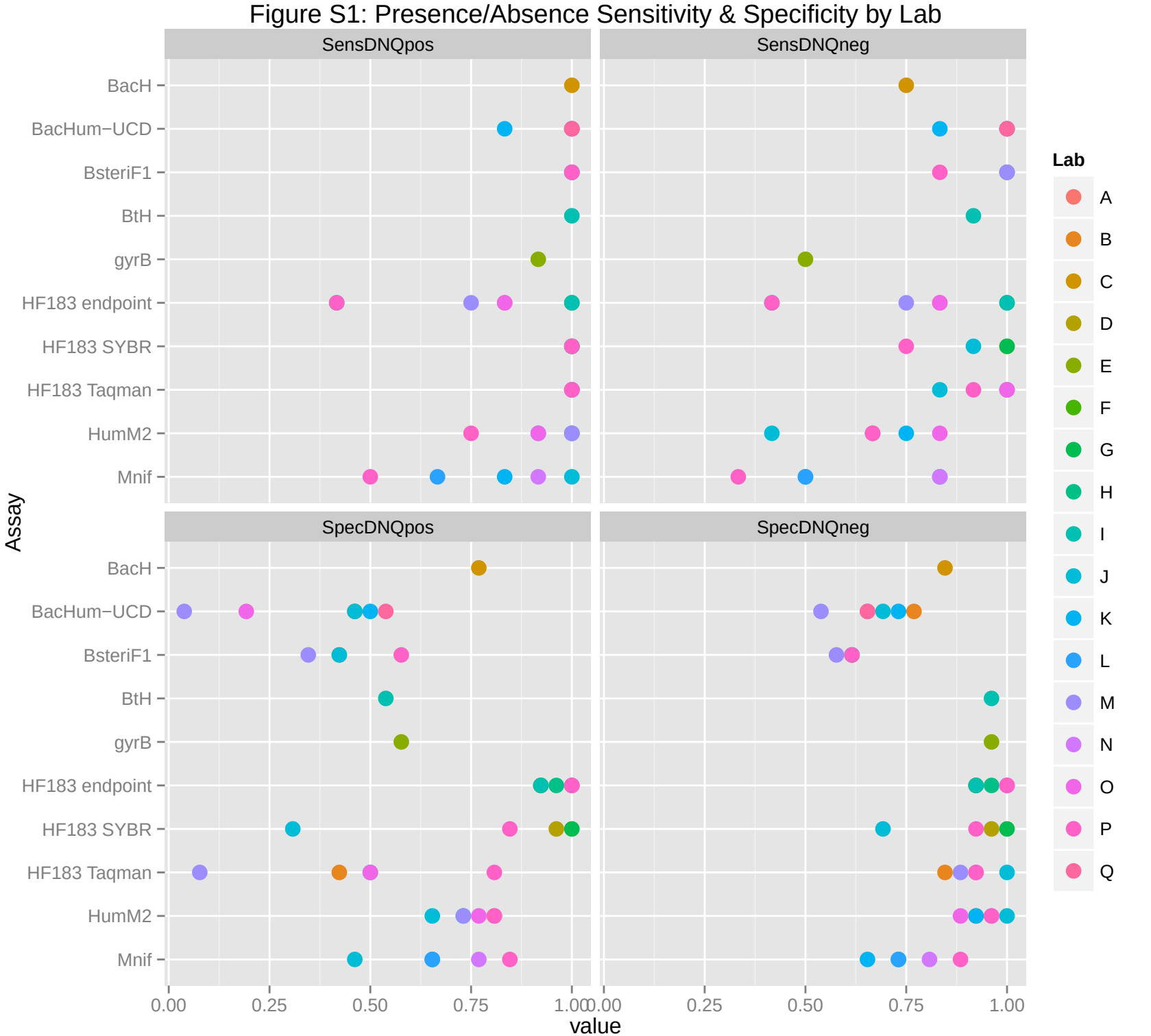
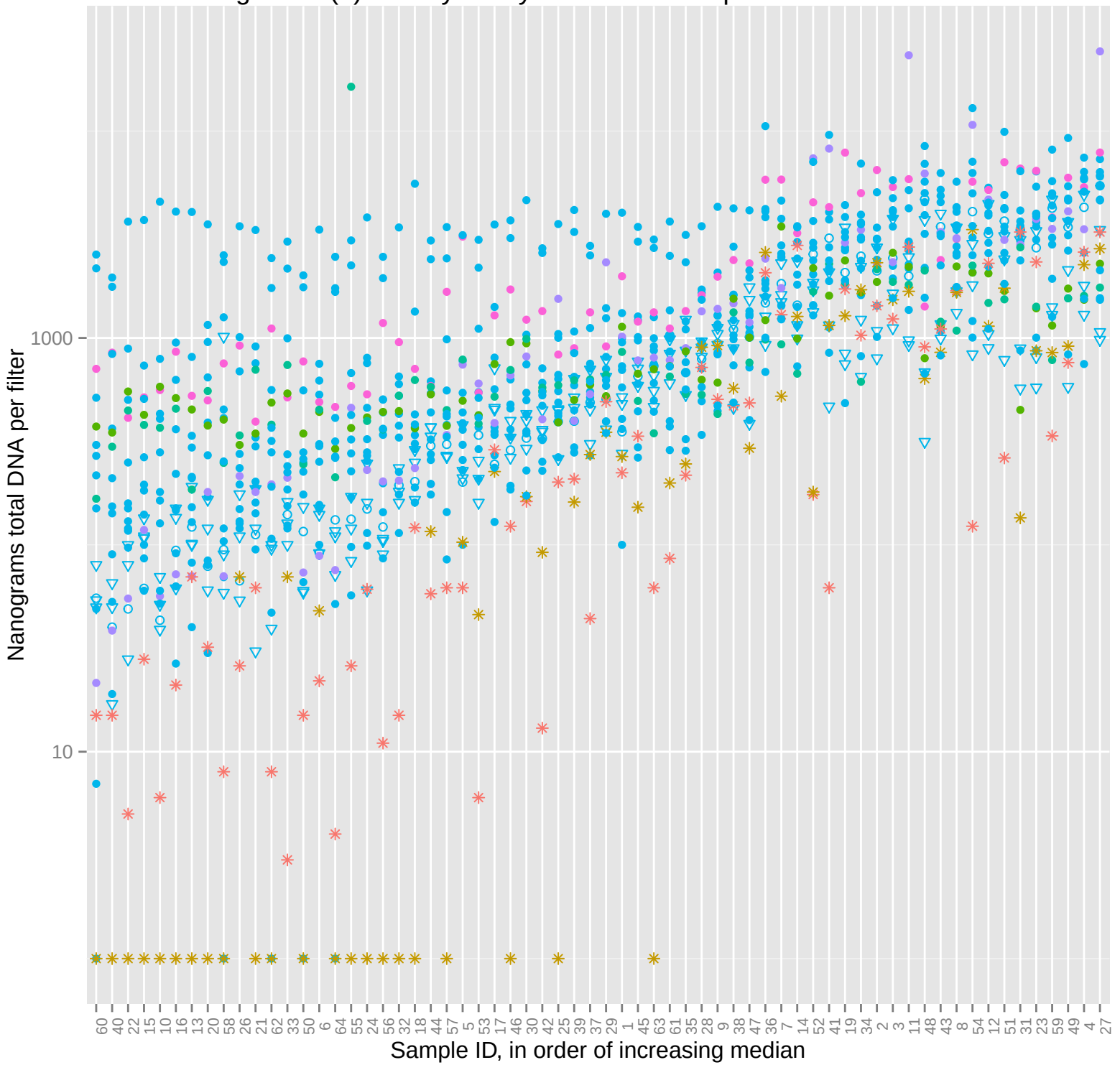


Figure S1: Presence/Absence Sensitivity & Specificity by Lab
SensDNQpos
SensDNQneg
SpecDNQpos
SpecDNQneg
BacH
BacHum−UCD
BsteriF1
BtH
gyrB
HF183 endpoint
HF183 SYBR
HF183 Taqman
HumM2
Mnif
Assay
0.00
0.25
0.50
0.75
1.00
value
Lab
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q

Figure S2(a): DNA yield by extraction and quantification methods



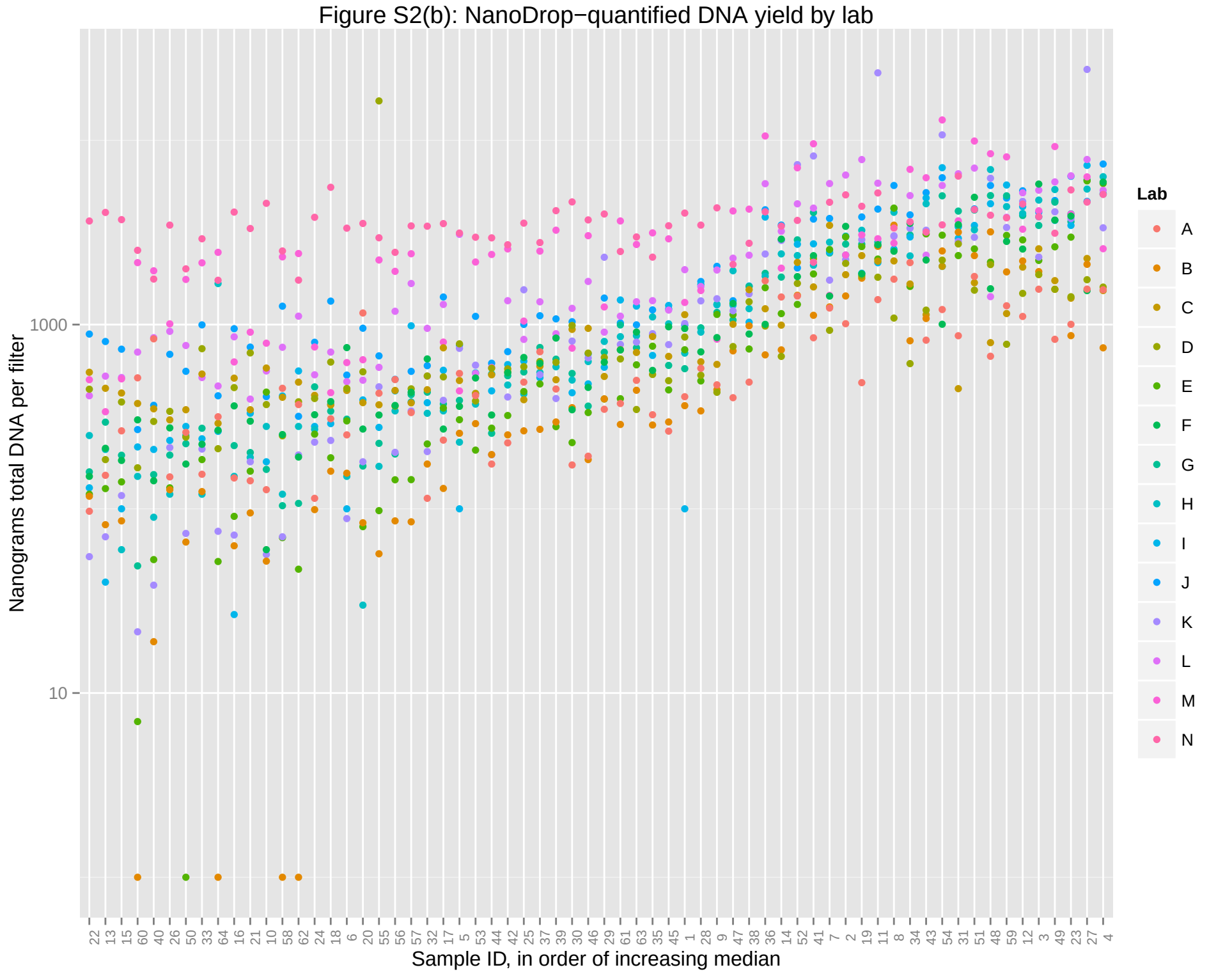
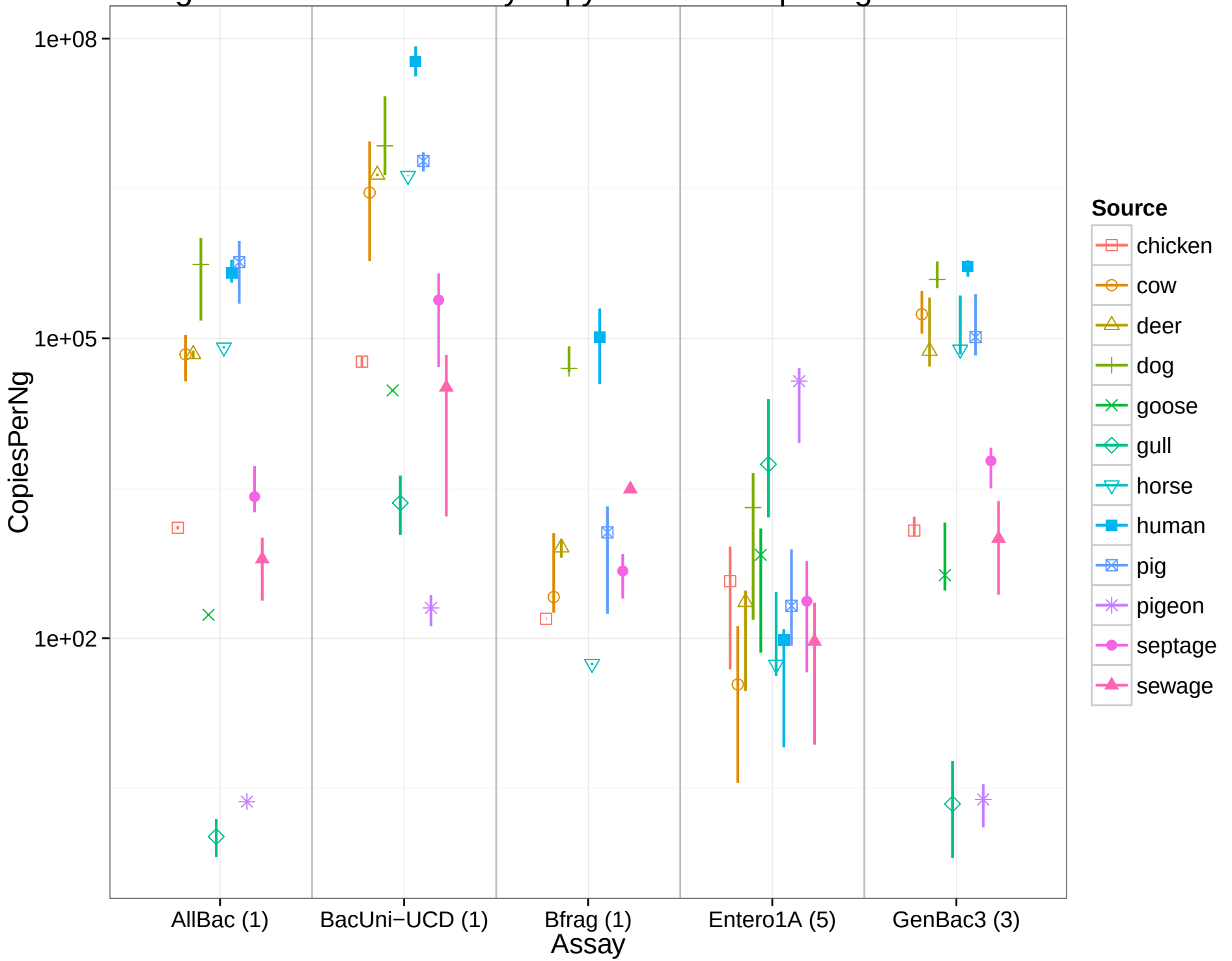


Figure S2(b): NanoDrop−quantified DNA yield by lab
Nanograms total DNA per filter
1000
10
Sample ID, in order of increasing median
22 13 15 60 40 26 50 33 64 16 21 10 58 62 24 18 6 20 55 56 57 32 17 5 53 44 42 25 37 39 30 46 29 61 63 35 45 1 28 9 47 38 36 14 52 41 7 2 19 11 8 34 43 54 31 51 48 59 12 3 49 23 27 4
Lab
A
B
C
D
E
F
G
H
I
J
K
L
M
N

Figure S3: General assay copy abundance per ng fecal DNA




1 50 100 150 200 250 300 350 400 450 500 550 590
AF233408 (uncultured human fecal bacterium HF8)
1 46 89 139 188 237 285 334 382 432 482 532 580
FWD DQ497994 (Bacteroides galacturonicus)
FWD DQ497990 (Bacteroides coagulans)
FWD BNRRR16SZZ (Bacteroides cellulosolvens)
FWD DQ497992 (Bacteroides xylanolyticus)
FWD AB200228 (Bacteroides pyogenes)
FWD DQ497991 (Bacteroides pyogenes)
FWD AB200229 (Bacteroides pyogenes)
FWD BNRRR16SAB (Bacteroides eggerthii)
FWD AB050107 (Bacteroides eggerthii DSM 20697)
FWD AB264624 (Bacteroides propionicifaciens)
FWD AB222699 (Bacteroides finegoldii DSM 17565)
FWD DQ794972 (uncultured bacterium)
FWD DQ807538 (uncultured bacterium)
FWD DQ823912 (uncultured bacterium)
FWD EF401444 (uncultured bacterium)
FWD EF399798 (uncultured bacterium)
FWD AB200224 (Bacteroides coprocola DSM 17136)
FWD AB064919 (Bacteroides coprophilus)
FWD Bacsa_R0077 - 10261000 (Bacteroides salanitronis DSM 18170)
FWD Bacsa_R0098 - 10261615 (Bacteroides salanitronis DSM 18170)
FWD Bacsa_R0070 - 10260839 (Bacteroides salanitronis DSM 18170)
REV Bacsa_R0045 - 10260187 (reversed) (Bacteroides salanitronis DSM 18170)
FWD Bacsa_R0062 - 10260656 (Bacteroides salanitronis DSM 18170)
FWD AB200219 (Bacteroides plebeius)
FWD AB200220 (Bacteroides plebeius)
FWD AB200217 (Bacteroides plebeius DSM 17135)
FWD AY126616 (Bacteroides massiliensis DSM 17679)
FWD EU136686 (Bacteroides caccae)
FWD Bache_R0031 - 10139856 (Bacteroides helcogenes P 36-108)
REV Bache_R0081 - 10141976 (reversed) (Bacteroides helcogenes P 36-108)
REV Bache_R0012 - 10138942 (reversed) (Bacteroides helcogenes P 36-108)
FWD Bache_R0052 - 10140784 (Bacteroides helcogenes P 36-108)
FWD AB200227 (Bacteroides helcogenes)
FWD EU136692 (Bacteroides helcogenes P 36-108)
FWD AB214328 (Bacteroides intestinalis DSM 17393)
FWD BNRRR16SAA (Bacteroides ovatus)
FWD AY652736 (Bacteroides ovatus)
FWD EU722734 (Bacteroides ovatus)
FWD AY608696 (Bacteroides salyersiae)
FWD AY608697 (Bacteroides nordii)
FWD AJ583243 (Bacteroides cellulosilyticus DSM 14838)
FWD AB247142 (Bacteroides uniformis)
REV BF16SrRNA06 - 3101800 (reversed) (Bacteroides fragilis YCH46)
REV BF16SrRNA03 - 3101818 (reversed) (Bacteroides fragilis YCH46)
REV BF16SrRNA02 - 3101837 (reversed) (Bacteroides fragilis YCH46)
REV BF16SrRNA05 - 3101808 (reversed) (Bacteroides fragilis YCH46)
REV BF16SrRNA01 - 3101824 (reversed) (Bacteroides fragilis YCH46)
REV BF638R_rRNA0007 - 11701091 (reversed) (Bacteroides fragilis 638R)
REV BF638R_rRNA0010 - 11703589 (reversed) (Bacteroides fragilis 638R)
REV BF638R_rRNA0016 - 11703769 (reversed) (Bacteroides fragilis 638R)
REV 16S_rRNA - 3289067 (reversed) (Bacteroides fragilis NCTC 9343)
REV BF638R_rRNA0004 - 11703200 (reversed) (Bacteroides fragilis 638R)
REV BF638R_rRNA0019 - 11702885 (reversed) (Bacteroides fragilis 638R)
FWD EU136678 (Bacteroides fragilis)
REV BT_r12 - 3284719 (reversed) (Bacteroides thetaiotaomicron VPI-5482)
REV BT_r06 - 3284689 (reversed) (Bacteroides thetaiotaomicron VPI-5482)
FWD BT_r13 - 3284683 (Bacteroides thetaiotaomicron VPI-5482)
FWD AY895205 (Bacteroides thetaiotaomicron)
FWD AB050109 (Bacteroides thetaiotaomicron)
FWD EU722739 (Bacteroides thetaiotaomicron)
FWD AY895201 (Bacteroides thetaiotaomicron)
REV BT_r03 - 3284642 (reversed) (Bacteroides thetaiotaomicron VPI-5482)
FWD AY895182 (Bacteroides thetaiotaomicron)
FWD BVU_1609 - 5302575 (Bacteroides vulgatus ATCC 8482)
REV BVU_3841 - 5304800 (reversed) (Bacteroides vulgatus ATCC 8482)
REV BVU_3572 - 5304532 (reversed) (Bacteroides vulgatus ATCC 8482)
FWD BVU_1999 - 5302965 (Bacteroides vulgatus ATCC 8482)
REV BVU_0458 - 5301427 (reversed) (Bacteroides vulgatus ATCC 8482)
FWD BVU_1695 - 5302661 (Bacteroides vulgatus ATCC 8482)
FWD BNRRR16SA (Bacteroides vulgatus)
FWD AY695676 (uncultured Bacteroidales bacterium)
FWD AY978208 (uncultured bacterium)
FWD AY985828 (uncultured bacterium)
FWD AY979363 (uncultured bacterium)
FWD X83953 (Bacteroides stericoris ATCC 43183)
FWD AY597206 (uncultured Bacteroides sp.)
FWD AY597199 (uncultured Bacteroides sp.)
FWD AY597200 (uncultured Bacteroides sp.)
FWD AY597204 (uncultured Bacteroides sp.)
FWD AF233413 (uncultured human fecal bacterium HF10)
FWD AY597203 (uncultured Bacteroides sp.)
FWD AY597201 (uncultured Bacteroides sp.)
FWD AB050110 (Bacteroides uniformis)
FWD AY597202 (uncultured Bacteroides sp.)
FWD AY597198 (uncultured Bacteroides sp.)
FWD AF233412 (uncultured human fecal bacterium HF74)
FWD AB242142 (Bacteroides dorei DSM 17855)
FWD AF233410 (uncultured human fecal bacterium HF102)
FWD AF233409 (uncultured human fecal bacterium HF117)
FWD AF233411 (uncultured human fecal bacterium HF145)
FWD AF233408 (uncultured human fecal bacterium HF8)
FWD BacHf
FWD DQ384027 (uncultured Bacteroidetes bacterium)
FWD DQ384045 (uncultured Bacteroidetes bacterium)
FWD DQ384031 (uncultured Bacteroidetes bacterium)
FWD DQ384023 (uncultured Bacteroidetes bacterium)
FWD DQ384022 (uncultured Bacteroidetes bacterium)
FWD DQ384049 (uncultured Bacteroidetes bacterium)
FWD DQ384046 (uncultured Bacteroidetes bacterium)
FWD HF183F
FWD AY618283 (uncultured Bacteroides sp.)
FWD AY618282 (uncultured Bacteroides sp.)
FWD AY618281 (uncultured Bacteroides sp.)
FWD BacHum160f
FWD BsteriF1DE
REV BacHr
FWD BacHum193p
REV HFsybR
REV BacHum241r
FWD GU130203 (Bacteroides sp. 1BL)
FWD BthetP1
FWD GU130202 (Bacteroides sp. 1B2)
REV BthetR1
FWD JF423926 (uncultured Bacteroides sp.)
FWD JF699654 (uncultured Bacteroides sp.)
REV Bac708R