Performance evaluation of canine-associated Bacteroidales assays in a multi-laboratory comparison study

Alexander Schriewer[1], Kelly D. Goodwin[2,], Christopher D. Sinigalliano[2], Annie M. Cox[3,*,**], David Wanless[4], Jakob Bartkowiak[4], Darcy L. Ebentier[5], Kaitlyn T. Hanley[1,5], Jared Ervin[6], Louise A. Deering[7], Orin C. Shanks[8], Lindsay A. Peed[8], Wim G. Meijer[7], John F. Griffith[9], Jorge SantoDomingo[10], Jennifer A. Jay[5], Patricia A. Holden[6] and Stefan Wuertz[1,11]*

Abstract

The contribution of fecal pollution from dogs in urbanized areas can be significant and is an often underestimated problem. Microbial source tracking methods (MST) utilizing quantitative PCR of dog-associated gene sequences encoding 16S rRNA of Bacteroidales are a useful tool to estimate these contributions. However, data about the performance of available assays are scarce. The results of a multi-laboratory study testing two assays for the determination of dog-associated Bacteroidales (DogBact and BacCan-UCD) on 64 single and mixed fecal source samples created from pooled fecal samples collected in California are presented here. Standardization of qPCR data treatment lowered

[1]*University of California, Department of Civil and Environmental Engineering, Davis, CA*

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

[3]*National Oceanic and Atmospheric Administration, Northwest Fisheries Science Center Laboratory, Seattle, WA*

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

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

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

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

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

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

[10]*US Environmental Protection Agency, Microbial Contaminants Control Branch, National Risk Management Research Laboratory, Cincinnati, OH*

[11]*Nanyang Technological University, Singapore Centre on Environmental Life Sciences Engineering (SCELSE), School of Biological Sciences, and School of Civil and Environmental Engineering, Singapore*

[*]*Stationed at NOAA/SWFSC, La Jolla, CA*

[**]*Currently at: University of British Columbia, Life Sciences Center, Department of Microbiology and Immunology, Vancouver, Canada*

inter-laboratory variability of sensitivity and specificity results. Both assays exhibited 100% sensitivity. Normalization methods are presented that eliminated random and confirmed non-target responses. The combination of standardized qPCR data treatment, use of normalization via a non-target specific Bacteroidales assay (GenBac3), and application of threshold criteria improved the calculated specificity significantly for both assays. Such measures would reasonably improve MST data interpretation not only for canine-associated assays, but for all qPCR assays used in identifying and monitoring fecal pollution in the environment.

Introduction

Fecal material contributes to microbial pollution at many coastal and freshwater sites. Dog feces, in particular, can account for poor water quality near rivers and beaches, because it contains a significant amount of fecal indicator bacteria (FIB). Loading estimates for canine fecal matter range from 1.5×10^8 to 5.6×10^9 CFU of enterococci/g dry weight, and the loading from one dog fecal event is comparable to 6,940 bird fecal events (Wright *et al.* 2009, Zhu *et al.* 2011). Dog exercise areas have been shown to impact water quality (Garfield and Walker 2008), and dog feces is a potential source of pathogens (Bagcigil *et al.* 2007, Gookin *et al.* 2007, Houf *et al.* 2008, Papini *et al.* 2009, Damborg *et al.* 2009). In addition, dogs can carry antibiotic resistant bacteria (Baptiste *et al.* 2005, Moodley *et al.* 2006, Nuttall *et al.* 2008), and infections between humans and co-habiting dogs appear to be emerging (Weese *et al.* 2006).

Accurate source identification can be used to address beach and shellfish contamination problems and for Total Maximum Daily Load (TMDL) determinations of allowable FIB levels, as a means to estimate supportable inputs of pathogens into a specific watershed (Kern 2002, Hagedorn *et al.* 2011). Once sources of contamination have been identified, scientifically and economically sound remediation efforts can be devised. To best aid watershed management, microbial source tracking (MST) markers should provide a quantitative means for fecal source identification as demonstrated previously (Reischer *et al.* 2006, Lee *et al.* 2010, Schriewer *et al.* 2010, Tambalo *et al.* 2012a). To address this need, MST assays utilizing quantitative PCR (qPCR) to identify fecal contamination from dogs have been developed and utilized in field studies (Kildare *et al.* 2007, Sinigalliano *et al.* 2010, Schriewer *et al.* 2010). The assays discussed here, DogBact (Sinigalliano *et al.* 2010) and BacCan-UCD (Kildare *et al.* 2007), both target Bacteroidales (Dick *et al.* 2005a,b), but have not been compared previously. In addition, it is uncommon for multiple laboratories to devote resources towards a round-robin test and there are no examples of qPCR-based methods having been validated in an interlaboratory exercise. For qPCR-based MST methods, only within-laboratory (intra-laboratory) studies have been performed (Shanks *et al.* 2010a,b).

The objectives of this study were: 1) to compare two promising quantitative PCR (qPCR)-based canine-associated assays, BacCan-UCD (Kildare *et al.* 2007) and DogBact (Sinigalliano *et al.* 2010), that target genes encoding Bacteroidales 16S rRNA in a multi-laboratory comparison study using a single-blinded approach with purified DNA from fecal slurries originating from either one animal or human source (singleton) or two combined sources (doubleton; Boehm *et al.* 2013); and 2) to develop a general procedure for the normalization of source tracking data obtained by qPCR. Eight different laboratories participated in the study, all of which received filters of the same challenge samples. This effort was part of a broader multi-laboratory assessment of MST method performance, the State of California-funded Source Identification Protocol Project (SIPP). Other host targets and aspects of assay performance are reported elsewhere (Boehm *et al.* 2013, Layton *et al.* 2013, Ebentier *et al.* 2013, Raith *et al.* 2013, Sinigalliano *et al.* 2013).

Methods

Preparation of Challenge Samples and Processing Briefly, 64 challenge samples were created by mixing fresh feces, sewage, or septage, in artificial freshwater. Feces used were from chicken, cow, dog, deer, goose, gull, horse, human, pig, and pigeon. This set of 64 blind samples was comprised of 19 single-source ("singleton") and 13 mixed-source ("doubleton") samples in duplicate. Each doubleton sample contained human feces, septage, or sewage in combination with one non-human source. Except for sewage (untreated influent from nine treatment facilities) and septage (6 septage collection trucks or community systems) a minimum of 12 individual samples were collected across California for each animal type. The challenge samples were filtered (50 ml or 5 ml; see Supplemental Information (SI) ftp://ftp.sccwrp.org/pub/download/DOCUMENTS/AnnualReports/2013AnnualReport/ar13_540_553SI.

pdf) over polycarbonate membrane filters with a diameter of 47 mm diameter and a pore size of 0.4 µm (Isopore, Millipore) and distributed frozen to the participating laboratories. The methods for creation and laboratory analysis of the challenge samples are described in detail elsewhere (Boehm *et al.* 2013 and SI).

Canine-Associated qPCR

The samples were analyzed for two different dog-associated Bacteroidales qPCR assays by eight laboratories from the United States and European Union. Some of the laboratories tested both assays, with the result that the DogBact assay (Sinigalliano *et al.* 2010) was run by six and the BacCan-UCD assay (Kildare *et al.* 2007) by four laboratories. Two different nucleic acid extraction methods were used: Gene Rite DNA EZ Extraction kit (GeneRite, North Brunswick, NJ) and DNeasy Blood & Tissue Kit (Qiagen, Valencia, CA; for further details see SI).

Standardized Calculation of qPCR Standard Curves and Results

After obtaining Cycle Threshold (C_t) values, laboratories differed in how final concentrations were calculated and how non-detects (ND), or detected but not quantifiable (DNQ), were defined. Some laboratories decided not to report DNQ data. To eliminate differences in results based on individual laboratory treatment, the raw C_t values for challenge samples and standard curves were collected from all laboratories and processed in the same way.

Regression analysis was performed on each individual standard curve to remove outliers; C_t values with a residual value larger than +3 or smaller than -3 were ignored when determining the final standard curve for each lab. The initial calculation of LLOQs using the precision of replicate standard curves resulted in values below those that could be confidently detected as lowest standard; for this reason we chose as a more conservative approach to only consider concentrations that yielded amplification in at least 50% of the replicates. Based on these LLOQs, copy numbers per reaction (cp) were calculated for each laboratory from the average C_t value of the non-outlier standard replicates at the lowest concentration included in the standard curve (Table 1). Each sample replicate was considered within the range of quantification (ROQ) if the C_t <LLOQ, DNQ if the C_t >LLOQ, and ND if there was no amplification detected by cycle 40.

Because laboratories also differed in numbers of dilutions and replicates of samples tested, a uniform strategy was applied to standardize the interpretation of results into ROQ, DNQ, and ND categories (see SI for details). Basically, the majority category among replicates defined whether the sample was categorized as ROQ, DNQ, or ND. For 50:50 situations in replicates between ROQ and DNQ, or DNQ and ND, both cases were categorized as DNQ. Specifically, for samples run in triplicate, C_t values were averaged if two of three, or three of three replicates were ROQ. If two of three, or three of three replicates were DNQ, the sample was

Table 1. Standard curve statistics for canine associated assays with standardized post processing.

Assay	Lab	Slope	y-Intercept	R^2	Efficiency* (%)	LLOQ (C_t)	LLOQ (cp/rxn)
DogBact	1	-3.46	38.4	0.992	94.7	34.8	11
	2	-3.51	44.8	0.951	92.8	38.2	77
	3	-3.35	41.5	0.990	98.8	37.5	16
	4	-3.23	39.0	0.982	104.0	35.6	11
	5	-3.54	41.5	0.986	91.5	37.6	13
	6	-3.41	46.5	0.997	96.5	38.6	217
BacCan-UCD	1	-3.43	44.7	1.000	95.8	37.9	100
	3	-3.27	45.7	0.997	102.3	38.9	122
	7	-3.60	46.0	0.992	89.5	38.8	100
	8	-3.36	43.0	0.971	98.5	32.7	1205

* Efficiency = $-1+10^{(-1/\text{slope})}$

considered DNQ. If two of three, or three of three, replicates were ND, the sample was considered ND. In the rare case that the replicates contained one each ROQ, DNQ and ND, the samples were considered DNQ. For samples run in duplicate, C_t values were averaged if two of two replicates were ROQ, and a single C_t values was used if one of two replicates was ROQ. If two of two replicates were DNQ, or one DNQ and one ND resulted, the sample was considered DNQ. If two of two replicates were ND, the sample was considered ND. For samples run in dilution series, all C_t values were considered and, accounting for dilution factors, converted into cp numbers for undiluted samples template using each laboratory's standardized standard curve. These concentrations were averaged and, according to individual LLOQ values, categorized as ROQ, DNQ, and ND. Thus, sample concentrations below LLOQ were considered DNQ. When C_t values were reported at >40 cycles, they were considered ND.

Sensitivity and Specificity

Sensitivity (%) was calculated as the number of challenge samples correctly identified as positive for the host feces divided by the total number of samples that contain the host feces:

$$\text{Sensitivity} = TP/(TP + FN) \qquad \textbf{Eq. 1}$$

where TP and FN are true positives and false negatives, respectively.

Specificity (%) was calculated as the number of challenge samples correctly identified as negative for the host feces divided by the total number of samples that did not contain the host feces:

$$\text{Specificity} = TN/(TN + FP) \qquad \textbf{Eq. 2}$$

where TN and FP are true negatives and false positives, respectively.

Normalization of Results

Concentrations of DNA, General Bacteroidales (GenBac3), and viable enterococci provided by the research group that created the blinded sample set were used to normalize the results obtained from individual laboratory analyses for DogBact and BacCan-UCD assays. A new approach based on dividing measured gene copies by the relative standard deviation (RSD) of all analyses to account for inter-laboratory variation was developed. Cultivable enterococci were measured by membrane filtration according to EPA method 1600 (USEPA 2006). Between two and three dilutions were filtered for enumeration. If more than one dilution was countable (between 1 and 250 CFU per filter), the counts were averaged to estimate CFU per 100 ml. The concentration of GenBac3 was estimated by three laboratories using EPA Method B (USEPA 2010), with a geometric mean of values among laboratories calculated to estimate cp per 100 ml. The concentration of DNA was estimated by two laboratories using Quant-iT kits (dsDNA High-Sensitivity or dsDNA Broad-Range; Invitrogen-Molecular Probes, Grand Island, NY) that were applied to 50 ml of slurry filtered through a membrane filter and using GeneRite DNA extraction kits. This concentration was multiplied by the exact volume and normalized to 100 ml of slurry. For detailed information see Boehm *et al.* (2013) and SI.

Results

The DogBact (6 laboratories) and BacCan-UCD (4 laboratories) assays for the detection of canine associated Bacteroidales were tested on 64 challenge samples comprised of duplicates of 19 singleton and 13 doubleton samples, prepared from pools of individual fecal samples collected across California. Several laboratories tested both qPCR assays. In general, laboratories were not restricted in how samples were processed or calculated and results reported. This non-restriction led to two obvious differences among the laboratories: 1) the number of replicates and dilutions for samples and standard curves run by each laboratory, and 2) the classification of results into ROQ, DNQ and ND. While the first difference influenced the quality of quantitative data, the second directly affected qualitative outcomes, i.e., the assessment of positive or negative samples (Table 2). To unify these outcomes, it was necessary to treat all laboratories' original data in a standardized fashion. More specifically, the same criteria were used to calculate standard curves, which affected how outliers within the standard-curve datasets were identified and LLOQs calculated. After obtaining the standardized standard curves, the same criteria were applied to individual laboratory-specific datasets for challenge samples in order to classify ROQ, DNQ, and ND results consistently. Because the standardized data treatment resulted in close to zero specificity or sensitivity for data from Laboratory 1 (DogBact)

Table 2. Sensitivity and specificity of assays based on originally submitted and standardized data from either singletons only or from singletons and doubletons together.

	Laboratory Number											
	DogBact							BacCan-UCD				
	1	2	3	4	5	6	AVG*	1	3	7	8	AVG*
Original Submitted Data												
	Singletons, DNQ = ND											
Sensitivity	1	1	1	1	1	1	**1**	1	1	1	1	**1**
Specificity	0.03	0.38	0.68	0.85	0.65	0.35	**0.58**	0.68	0.59	0.82	0.71	**0.7**
	Singletons, DNQ = Positive											
Sensitivity	1	1	1	1	1	1	**1**	1	1	1	1	**1**
Specificity	0.03	0.38	0.68	0.85	0.65	0.35	**0.58**	0.32	0.47	0.82	0.71	**0.54**
	with Doubletons, DNQ = ND											
Sensitivity	1	1	1	1	1	1	**1**	1	1	1	1	**1**
Specificity	0.02	0.27	0.68	0.82	0.63	0.34	**0.55**	0.66	0.61	0.84	0.63	**0.7**
Standardized Data												
	Singletons, DNQ = ND											
Sensitivity	1	1	1	1	1	1	**1**	1	1	1	0	**1**
Specificity	0	0.59	0.91	0.94	0.91	0.79	**0.83**	0.68	0.59	0.79	1	**0.69**
	with Doubletons, DNQ = ND											
Sensitivity	1	1	1	1	1	1	**1**	1	1	1	0	**1**
Specificity	0	0.5	0.91	0.95	0.91	0.77	**0.81**	0.68	0.61	0.82	1	**0.7**
Log (1/RSD) Method												
	Singletons, Values Larger RSD (150%) = ND											
Sensitivity	1	1	1	1	1	1	**1**	1	1	1	0	**1**
Specificity	0	0.76	0.94	0.94	0.94	0.88	**0.89**	0.82	0.76	0.88	1	**0.82**
Normalized via GenBac3												
	Singletons, Values Smaller Log (0.01) = ND											
Sensitivity	1	1	1	1	1	1	**1**	1	1	1	0	**1**
Specificity	0.71	0.88	1	1	1	0.91	**0.96**	0.91	0.91	0.88	1	**0.9**

*AVG: average, results from lab 1 (DogBact) and lab 8 (BacCan-UCD) are excluded

and Laboratory 8 (BacCan-UCD), respectively, this data were excluded from the calculation of average sensitivities and specificities (Table 2). Both datasets represent special cases in which most ROQs had to be classified as DNQ or vice versa after application of standardized criteria.

Effect of Standardized Data Treatment on Assay Sensitivity and Specificity

Samples with an assay response below concentrations of the calculated laboratory's LLOQ were considered DNQ. Consideration of DNQ samples as positive samples resulted without exception in lower average specificity of assays without affecting sensitivities (Table 1 and Table SI-5). The calculated average specificity of the DogBact assay increased significantly after standardized data treatment from 0.46 to 0.81 when all samples were considered and from 0.49 to 0.83 when only singletons were considered. In contrast, the differences in calculated average specificity for the BacCan-UCD assay before and after standardization of results were only minimal with changes of 0.68 to 0.70 for all samples and 0.70 to 0.69 for singletons only. Besides concentrated and diluted dog singleton challenge samples, goose and septage samples were also reported by the majority of laboratories as above DNQ for both assays, and cow and chicken for only the BacCan-UCD assay. Fecal samples from dogs were collected in urban areas, minimizing the probability of direct contact between chicken and dogs. The Bacteroidales population in chicken appears to be rather diverse compared to other animal hosts and similar results were observed for other non-dog markers (unpublished information).

As expected, the highest signals were reported for both assays for the dog singleton samples, but pigeon (DogBact) and chicken (BacCan-UCD) singleton samples exhibited similar concentrations among some of the laboratories (Figures 1 and 2). Interestingly, for the DogBact assay at least one replicate filter from every sample was reported in the quantifiable range (ROQ) by at least one of the laboratories. This was also the case for the BacCan-UCD marker with the exception of gull, pigeon, and diluted sewage and human samples. Although standardized data treatment removed false positive sewage and diluted septage responses among the BacCan-UCD results by re-classifying the ROQ values to DNQ, most other non-target

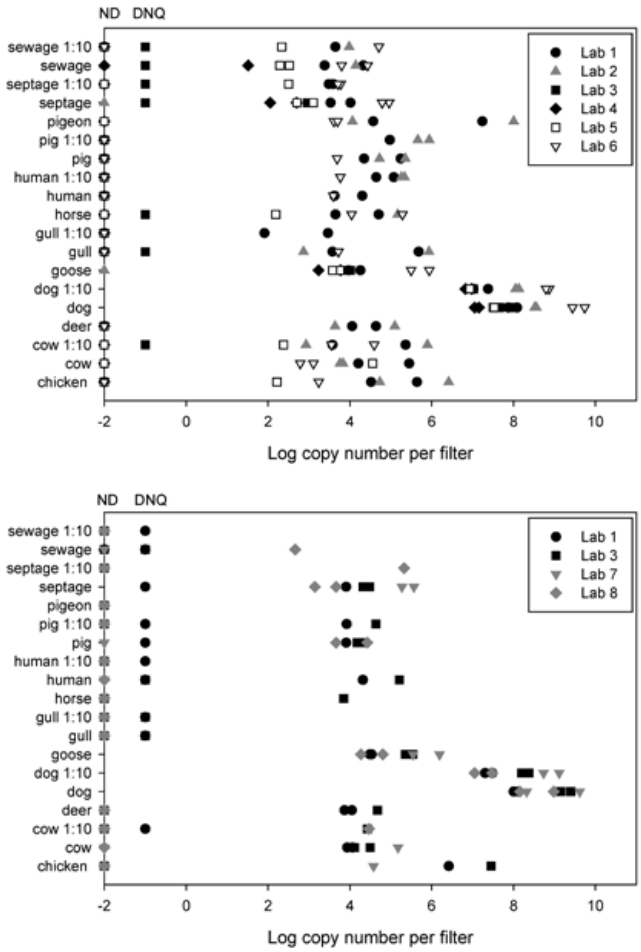

Figure 1. Originally submitted results from 38 single-source (duplicate) samples for DogBact (top) and BacCan-UCD (bottom). ND = not detected. DNQ = detected but not quantifiable.

samples were measured clearly above LLOQ values. These results, particularly the differences observed between replicate filters, motivated additional data analysis.

Effect of Different Normalization Procedures on Assay Sensitivity and Specificity

What distinguishes qPCR from PCR, as the name suggests, is the ability to generate quantitative data, which allows data conversion beyond binary presence/absence observations. The following section describes methods to normalize the dog-associated Bacteroidales concentration data to ideally identify non-target responses (false positives).

Consideration of Relative Standard Deviation

If the observed non-target responses were due to random effects, such as qPCR signal artifacts caused

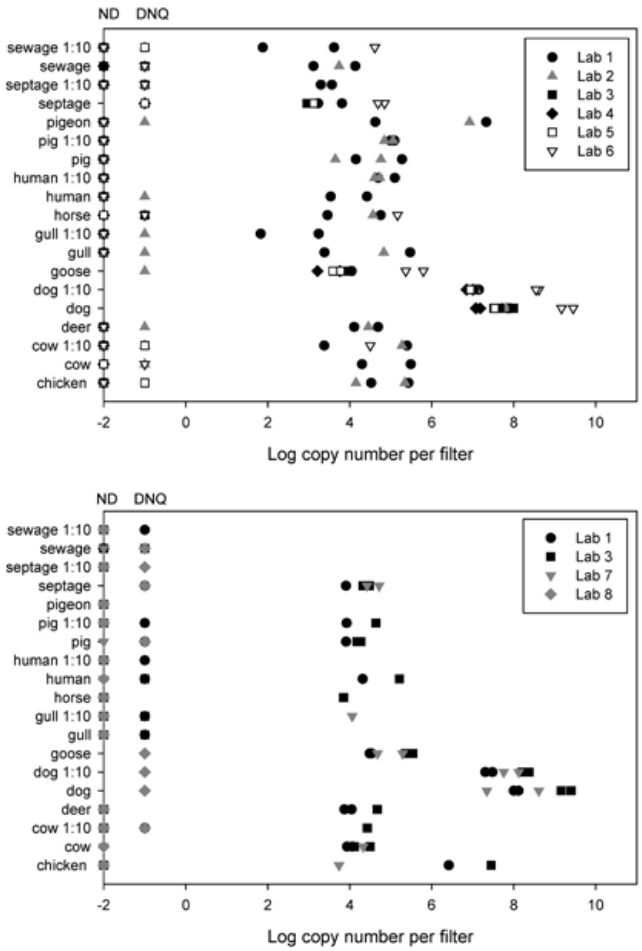

Figure 2. Results from 38 single-source (duplicate) samples for DogBact (top) and BacCan-UCD (bottom) after standardized data treatment. ND = not detected. DNQ = detected but not quantifiable.

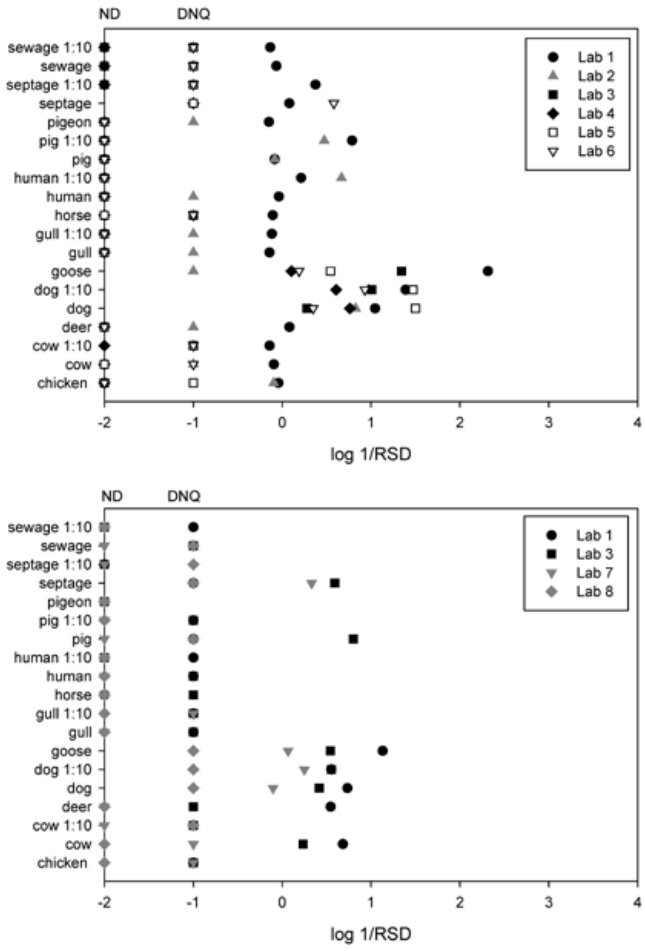

Figure 3. Presentation of DogBact(top) and BacCan-UCD (bottom) results as 1/RSD. Results of post standardized data treatment from 38 single-source (duplicates) samples are used. ND = not detected. DNQ = detected but not quantifiable (assigned when either of the two replicate samples was DNQ or one ND and the other detected).

by irregularities in reaction mixture or material, possible cross contamination, or sample mix-up, it would be unlikely for responses to appear in both duplicates of the same sample to the same extent. On the other hand, true positives should deliver an intense (high concentration) and reproducible response. To eliminate positives due to such errors ('random positives'), we used the relative standard deviation between the two replicates was used. By displaying the reciprocal 1/RSD values, true positives should be found towards the right side of data plots. This treatment led to a significant reduction of non-target data points; yet, chicken, human and pig samples remained above DNQ criteria reported by at least two laboratories for the DogBact assay, and cow and septage samples remained above DNQ criteria for BacCan-UCD (Figure 3). For both assays, goose and dog samples were the only samples for which data points from the majority of laboratories remained positive. Taking into account that these standard deviations comprised all possible sources of variability, from sample filtration to sample extraction and quantification, an RSD of 150% was considered reasonable; thus, a threshold value of log ($1/RSD_{150}$) = -0.176 was used for sensitivity and specificity calculations (Table 2). The resulting average specificities were 0.89 and 0.82 for DogBact and BacCan-UCD assays, respectively. As expected, the sensitivity was not affected by this data conversion and remained at 1.00.

Normalization of Quantitative qPCR Data by Total DNA Concentration

The total DNA concentration of a sample is a good indication of its total biomass. This particular

study used a defined sample set of known feces; therefore, a correlation between target concentration and total DNA was expected (barring random variation and possible cross-contamination, as discussed above). Thus, normalizing qPCR concentrations by measured total DNA concentrations represented another way to reduce random non-target signals. This normalization (log (copies assay/ng source DNA)) lead to a clear separation of dog singleton samples among all laboratories. For DogBact, the target samples were separated from the other samples at a value greater than four log gene copies per nanogram DNA on the x-axis (Figure 4). For the BacCan-UCD assay, this same clustering above four on the x-axis (log scale) was seen for target samples, with the exception that cross-reactivity with chicken samples was observed by two laboratories that reported results in the same range as dog singleton samples (Figure 4). It is unclear as to why a strong separation occurred at this particular value, but a set of samples clustered at about four magnitudes higher than other samples was observed. This threshold appeared to separate samples that gave true positive signals because when this observational threshold was used to classify data in a binary analysis, the average DogBact and BacCan-UCD specificities became 1.00 and 0.98, respectively. Again, the sensitivities of 1.00 for both assays were not affected by the normalization.

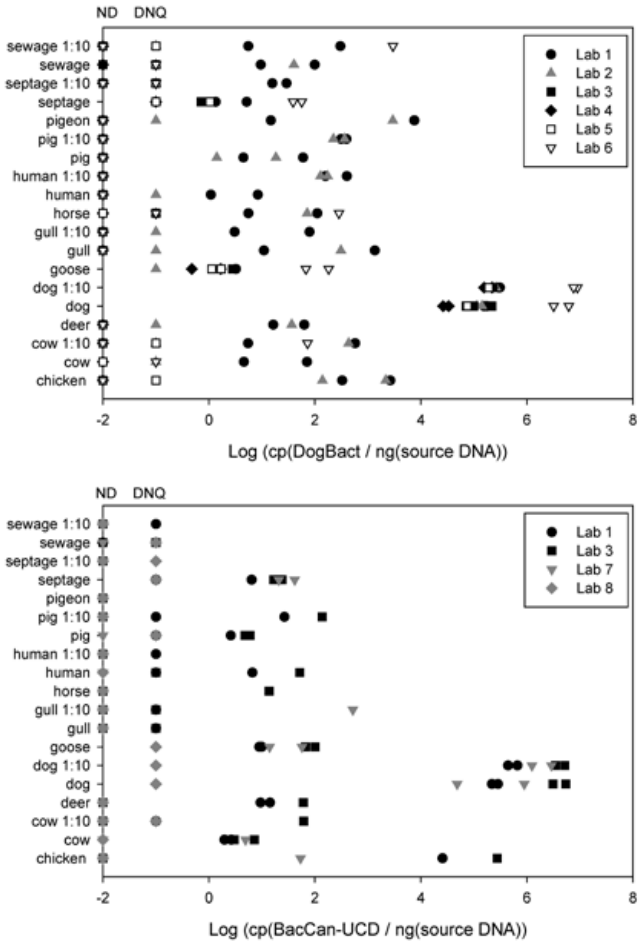

Figure 4. Normalization of DogBact (top) and BacCan-UCD (bottom) results by ng DNA measured in the sources. Results of post standardized data treatment from 38 single-source (duplicate) samples are used. ND = not detected. DNQ = detected but not quantifiable.

Normalization of Quantitative qPCR Data by GenBac3 Concentration

Many of the MST markers used in the SIPP study targeted source-associated Bacteroidales, in part because this order is so abundant in a variety of feces. Thus, normalizing host-associated Bacteroidales qPCR data against non-host-associated (universal) Bacteroidales qPCR data (GenBac3 results) may provide improved normalization compared to other methods (e.g., versus normalization to enterococci). Normalization to GenBac3 did indeed lead to enhanced cluster formation of data points for the BacCan-UCD assay (Figure 5). In contrast, DogBact results were mostly unaffected.

Using the assumption that a host-associated marker would need to represent at least 1% of the universal Bacteroidales marker concentration to be of utility in real-world samples, gives a ratio of dog-associated Bacteroidales to general Bacteroidales of 0.01 (i.e., Log (0.01) = -2) as a conservative binary threshold for calculation of sensitivities and specificities. This analysis resulted in average specificities of 0.96 and 0.90 for DogBact and BacCan-UCD assays, respectively. Again, the average sensitivities remained at 1.00 for both assays.

Discussion

Determining Most Probable Cross-Reactivity

Limited information about cross reactivity with non-target fecal samples for the two dog-associated Bacteroidales markers used in this study was available. In designing the probe, Sinigalliano *et al.* (2010) converted a published dog-associated Bacteroidales conventional PCR assay (Dick *et al.* 2005a) to the DogBact qPCR assay used in this study (Sinigalliano *et al.* 2010). The SIPP study is the first to report cross-reactivity data for this qPCR assay (additional specificity testing data are given

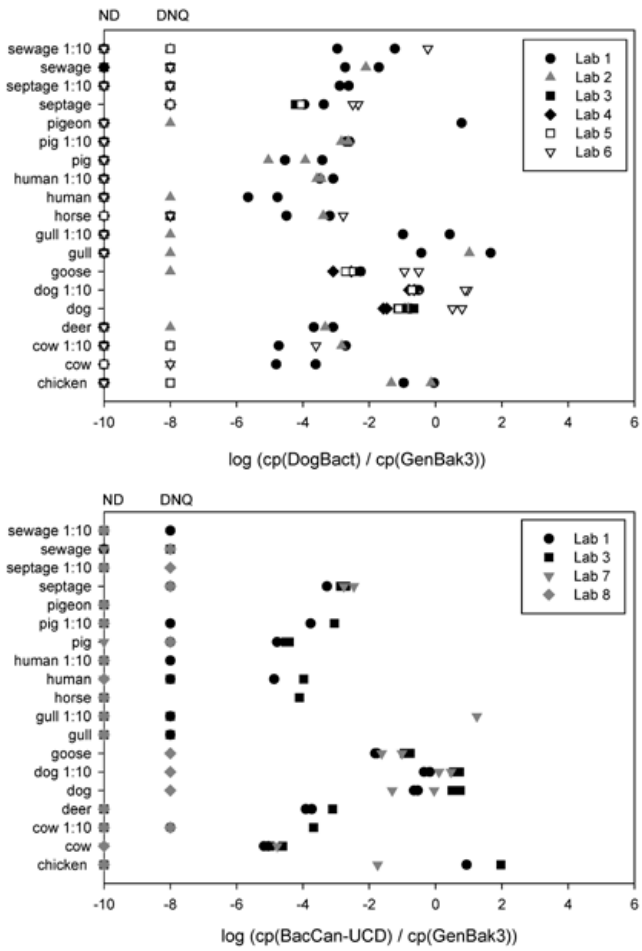

Figure 5. Normalization of DogBact (top) and BacCan-UCD (bottom) results by GenBac3 copy numbers. Results of post standardized data treatment from 38 single-source (duplicates) samples are used. ND = not detected. DNQ = detected but not quantifiable.

in SI and Boehm *et al.* 2013). Kildare *et al.* (2007) previously reported cross-reactivity of BacCan-UCD with individual human (4/18), pooled cat (1/7), and wastewater influent (4/14) samples. Cross-reactivity also occurred with pooled cow (11/11), horse (2/10), and goose (5/10) fecal samples as well as sewage (10/12; Silkie and Nelson, 2009). Other studies modified the BacCan-UCD assay significantly, either by using only conventional PCR and removing the second reverse primer (Ahmed *et al.* 2008) or by choosing a different annealing temperature (Tambalo *et al.* 2012b), and thus cannot be used for a valid comparison of reported cross-reactivity of the assay.

The standardized data treatment removed variations among laboratory results due to data processing. Comparing the standardized datasets for singleton challenge samples of participating laboratories confirmed that some cross reactivity was observed for both assays for goose and septage, and also for chicken for BacCan-UCD. However, other cross-reactivities were not consistent between sample replicates or across laboratories, indicating random errors perhaps introduced by cross-contamination, qPCR artifacts, or other unidentified reasons. Differences in detection limits of the different laboratories could also have been a factor for low level cross -reactivity as the reported LLOQ values varied significantly among laboratories (Table 1).

Consequently, reports of false positives with non-targets must be viewed with caution when observed in a minority of non-target samples or by only one laboratory, whether from individual or composite samples.

Consequences for MST Studies

Statistical models can help to adjust host distributions relative to each other by translating measured concentrations into probable distributions by accounting for uncertainties and errors (Wang *et al.* 2010). The accuracy of such treatment strongly depends on the strength of the original validation of assay specificity and sensitivity. However, not all assays used in the SIPP study were validated in a similar fashion, and this may have confounded assay comparisons. Many papers in the SIPP study presented differential performance characteristics for assays depending on how qPCR data were binned into binary results, with particular focus on treatment of DNQ results (Layton *et al.* 2013, Raith *et al.* 2013, Sinigalliano *et al.* 2013). Validation requires adequate replication (Armbruster and Pry 2008, CODEX 2010), and variations in LOD and LLOQ calculations (Stewart *et al.* 2013) can directly affect how quantitative results are binned for binary analysis. It is likely that if all assays had been similarly and rigorously validated, creating accurate limit of detection and limit of quantification statistics, much of the need for interpretation would have disappeared. Clearly, the overall SIPP study has highlighted the necessity of uniform protocols to implement MST across laboratories (Ebentier *et al.* 2013). In addition, the results in total suggest a requirement for defined assay validation procedures prior to moving assays into inter-laboratory verification exercises. Validation needs to include definition of matrix effects, and for MST applications, sample aging. Genotyping or metagenomic sequencing of samples (Tringe and Rubin 2005, Rosario *et al.* 2009) would also benefit future validation efforts.

In the absence of additional validation exercises, any type of positive MST results will be subject to uncertainty, whether the observed signal is a result of a small load of real host feces or due to a larger load of cross-reactive non-host feces. One possible solution is normalization by other related parameters, assuming that that the ratio of true associated host signal to, e.g., total DNA would be much bigger than if the reason for the signal were the presence of much more (cross-reactive) DNA material. While measurements like *Enterococcus* and total DNA concentrations are relatively simple to perform, they may not be related to host-associated Bacteroidales data. Normalization by total DNA concentration was effective in separating target from non-target samples in this study. This result was likely due to the fact that the samples contained only mixed fecal material so that total DNA concentration was an adequate proxy for the DNA concentration of the fecal bacteria predominant in the sample; this relationship is unlikely to hold with real environmental samples. In contrast, normalization by *Enterococcus* did not aid assay performance (Table SI-1). Notably, the ratio of enterococci to each MST marker is expected to vary. The *Enterococcus* population can vary significantly in feces of different animals and in relation to Bacteroidales and hence may not be optimal for normalization (Silkie and Nelson 2009, Ervin *et al.* 2013).

Normalization of these Bacteroidales-based MST markers to general Bacteroidales may provide a better method of normalization. Correlations between host associated and general Bacteroidales concentrations have been observed (Silkie and Nelson 2009, Schriewer *et al.* 2010), although variations for individual fecal samples do occur. In this study, normalization of host-associated Bacteroidales (DogBact and BacCan-UCD) assays to general Bacteroidales (GenBac3 assay) improved assay performance. Elimination of samples with less than 1% of host-assay contribution towards the general Bacteroidales concentration as non-target resulted in an improvement of specificity from 0.83 to 0.96 for the DogBact assay and from 0.69 to 0.90 for the BacCan-UCD assay after standardization of data treatment. The normalization of host-associated Bacteroidales with general Bacteroidales marker results for environmental samples seems plausible, but criteria for the identification of main sources need to be carefully considered based on the assays tested.

Literature Cited

Ahmed, W., D. Powell, A. Goonetilleke and T. Gardner. 2008. Detection and source identification of faecal pollution in non-sewered catchment by means of host-specific molecular markers. *Water Science and Technology* 58:579-586.

Armbruster, D.A. and T. Pry. 2008. Limit of blank, limit of detection and limit of quantitation. *The Clinical Biochemist Reviews* 29:49-52.

Bagcigil, A.F., S. Ikiz, B. Dokuzeylul, B. Basaran, E. Or and N.Y. Ozgur. 2007. Fecal shedding of *salmonella* spp. in dogs. *Journal of Veterinary Medical Science* 69:775-777.

Baptiste, K.E., K. Williams, N.J. Willams, A. Wattret, P.D. Clegg, S. Dawson, J.E. Corkill, T. O'Neill and C.A. Hart. 2005. Methicillin-resistant staphylococci in companion animals. *Emerging Infectious Diseases* 11:1942-1944.

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.

CODEX. 2010. Guidelines on Performance Criteria and Validation of Methods for Detection, Identification and Quantification of Specific DNA Sequences and Specific Proteins in Foods. Codex Committee on Methods of Analysis and Sampling CAC/GL 74-2010. Codex Alimentarius Commission - WHO. Rome, Italy.

Damborg, P., J. Top, A.P.A. Hendrickx, S. Dawson, R.J.L. Willems and L. Guardabassi. 2009. Dogs are a reservoir of ampicillin-resistant *Enterococcus faecium* lineages associated with human infections. *Applied and Environmental Microbiology* 75:2360-2365.

Dick, L.K., A.E. Bernhard, T.J. Brodeur, J.W.S. Domingo, J.M. Simpson, S.P. Walters and K.G. Field. 2005a. Host distributions of uncultivated fecal Bacteroidales bacteria reveal genetic markers for fecal source identification. *Applied and Environmental Microbiology* 71:3184-3191.

Dick, L.K., M.T. Simonich and K.G. Field. 2005b. Microplate subtractive hybridization to enrich for Bacteroidales genetic markers for fecal

source identification. *Applied and Environmental Microbiology* 71:3179-3183.

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.

Garfield, L. and M. Walker. 2008. Microbial water quality and influences of fecal accumulation from a dog exercise area. *Journal of Environmental Health* 71:24-29.

Gookin, J.L., S.H. Stauffer, M.R. Coccaro, M.J. Marcotte and M.G. Levy. 2007. Optimization of a species-specific polymerase chain reaction assay for identification of *Pentatrichomonas hominis* in canine fecal specimens. *American Journal of Veterinary Research* 68:783-787.

Hagedorn, C., A.R. Blanch and V.J. Harwood (eds.). 2011. Microbial Source Tracking: Methods, Applications, and Case Studies. Springer. New York, NY.

Houf, K., S. De Smet, J. Bare and S. Daminet. 2008. Dogs as carriers of the emerging pathogen *Arcobacter*. *Veterinary Microbiology* 130:208-213.

Kern, J., B. Petrauskas, P. McClellan, V.O. Shanholtz and C. Hagedorn. 2002. Bacterial source tracking: A tool for total maximum daily load development. *in*: T. Younos (ed.), Advances in Water Monitoring Research. Water Resource Publications. Highlands Ranch, CO.

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 dogspecific fecal Bacteroidales: a Bayesian approach. *Water Research* 41:3701-3715.

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

Lee, D.Y., S.C. Weir, H. Lee and J.T. Trevors. 2010. Quantitative identification of fecal water pollution sources by TaqMan real-time PCR assays using Bacteroidales 16S rRNA genetic markers. *Applied Microbiology and Biotechnology* 88:1373-1383.

Moodley, A., M. Stegger, A.F. Bagcigil, K.E. Baptiste, A. Loeffler, D.H. Lloyd, N.J. Williams, N. Leonard, Y. Abbott, R. Skov and L. Guardabassi. 2006. Apa typing of methicillin-resistant *Staphylococcus aureus* isolated from domestic animals and veterinary staff in the UK and Ireland. *Journal of Antimicrobial Chemotherapy* 58:1118-1123.

Nuttall, T., N. Williams, R. Saunders and S. Dawson. 2008. Methicillin-resistant staphylococci in companion animals. *European Journal of Companion Animal Practice* 18:280-282.

Papini, R., M. Marangi, F. Mancianti and A. Giangaspero. 2009. Occurrence and cyst burden of *Giardia duodenalis* in dog faecal deposits from urban green areas: Implications for environmental contamination and related risks. *Preventive Veterinary Medicine* 92:158-162.

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.H., D.C. Kasper, R. Steinborn, R.L. Mach and A.H. Farnleitner. 2006. Quantitative PCR method for sensitive detection of ruminant fecal pollution in freshwater and evaluation of this method in alpine karstic regions. *Applied and Environmental Microbiology* 72:5610-5614.

Rosario, K., C. Nilsson, Y.W. Lim, Y.J. Ruan and M. Breitbart. 2009. Metagenomic analysis of viruses in reclaimed water. *Environmental Microbiology* 11:2806-2820.

Schriewer, A., W.A. Miller, B.A. Byrne, M.A. Miller, S. Oates, P.A. Conrad, D. Hardin, H.H. Yang, N. Chouicha, A. Melli, D. Jessup, C. Dominik and S. Wuertz. 2010. Presence of Bacteroidales as a predictor of pathogens in surface waters of the central California coast. *Applied and Environmental Microbiology* 76:5802-5814.

Shanks, O.C., K. White, C.A. Kelty, S. Hayes, M. Sivaganesan, M. Jenkins, M. Varma and R.A. Haugland. 2010a. Performance assessment PCR-based assays targeting Bacteroidales genetic markers of bovine fecal pollution. *Applied and Environmental Microbiology* 76:1359-1366.

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

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

Sinigalliano, C.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.

Sinigalliano, C.D., J.M. Fleisher, M.L. Gidley, M.H. Solo-Gabriele, T. Shibata, L.R.W. Plano, S.M. Elmir, D. Wanless, J. Bartkowiak, R. Boiteau, K. Withum, A.M. Abdelzaher, G.Q. He, C. Ortega, X.F. Zhu, M.E. Wright, J. Kish, J. Hollenbeck, T. Scott, L.C. Backer and L.E. Fleming. 2010. Traditional and molecular analyses for fecal indicator bacteria in non-point source subtropical recreational marine waters. *Water Research* 44:3763-3772.

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.

Tambalo, D.D., B. Fremaux, T. Boa and C.K. Yost. 2012a. Persistence of host-associated Bacteroidales gene markers and their quantitative detection in an urban and agricultural mixed prairie watershed. *Water Research* 46:2891-2904.

Tambalo, D.D., T. Boa, K. Liljebjelke and C.K. Yost. 2012b. Evaluation of two quantitative PCR assays using Bacteroidales and mitochondrial DNA markers for tracking dog fecal contamination in waterbodies. *Journal of Microbiological Methods* 91:459-467.

Tringe, S.G. and E.M. Rubin. 2005. Metagenomics: DNA sequencing of environmental samples. *Nature Reviews Genetics* 6:805-814.

US Environmental Protection Agency (USEPA). 2006. Method 1600: Enterococci in Water by Membrane Filtration Using membrane-Enterococcus Indoxyl-beta-D-Glucoside Agar (mEI). EPA 821-R-06-009. USEPA Office of Water. Washington, DC.

USEPA. 2010. Method B: Bacteroidales in Water by TaqMan® Quantitative Polymerase Chain Reaction (qPCR) Assay. EPA-822-R-10-003. USEPA Office of Water. Washington, DC.

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

Weese, J.S., H. Dick, B.M. Willey, A. McGeer, B.N. Kreiswirth, B. Innis and D.E. Low. 2006. Suspected transmission of methicillin-resistant *Staphylococcus aureus* between domestic pets and humans in veterinary clinics and in the household. *Veterinary Microbiology* 115:148-155.

Wright, M.E., H.M. Solo-Gabriele, S. Elmir and L.E. Fleming. 2009. Microbial load from animal feces at a recreational beach. *Marine Pollution Bulletin* 58:1649-1656.

Zhu, X.F., J.D. Wang, H.M. Solo-Gabriele and L.E. Fleming. 2011. A water quality modeling study of non-point sources at recreational marine beaches. *Water Research* 45:2985-2995.

Acknowledgements

This study was primarily funded by a Clean Beach Initiative grant from the California Water Resources Control Board. It was also supported in part by Oceans and Human Health Center grants from NSF and NIEHS (NSF 0CE0432368/0911373 and NIEHS P50 ES12736, respectively). Work in the Wuertz laboratory was funded by the Division of Environmental Analysis of the California Department for Transportation contract no. 43A0259. Information has been subjected to the US Environmental Protection Agency (USEPA) peer and administrative review and has been approved for external publication. Any opinions expressed in this paper are those of the author(s) and do not necessarily reflect the official positions and policies of the USEPA. Any mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Supplemental Information

Supplemental Information is available at ftp://ftp.sccwrp.org/pub/download/DOCUMENTS/AnnualReports/2013AnnualReport/ar13_540_553SI.pdf.