Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes

Valerie J. Harwood1, Christopher Staley1,2, Brian D. Badgley3, Kim Borges4 & Asja Korajkic5

1Department of Integrative Biology, University of South Florida, Tampa, FL, USA; 2University of Minnesota BioTechnology Institute, St. Paul, MN, USA; 3Department of Crop and Soil Environmental Science, Virginia Tech, Blacksburg, VA, USA; 4University of Maine Fort Kent, Fort Kent, ME, USA; and 5National Exposure Research Laboratory, U.S. Environmental Protection Agency, Cincinnati, OH, USA

Correspondence: Valerie J. Harwood, Department of Integrative Biology, University of South Florida, 4202 E Fowler Ave SCA 110, Tampa, FL 33620 USA. Tel.: (813) 974 1524; fax: (813) 974 3263; e-mail:vharwood@usf.edu

Received 30 September 2012; revised 8 May 2013; accepted 25 June 2013. Final version published online 7 August 2013.

DOI: 10.1111/1574-6976.12031

Editor: Colin Berry

Keywords
water quality; recreational water; MST; fecal indicator bacteria; qPCR; epidemiology.

Abstract
Microbial source tracking (MST) describes a suite of methods and an investigative strategy for determination of fecal pollution sources in environmental waters that rely on the association of certain fecal microorganisms with a particular host. MST is used to assess recreational water quality and associated human health risk, and total maximum daily load allocations. Many methods rely on signature molecules (markers) such as DNA sequences of host-associated microorganisms. Human sewage pollution is among the greatest concerns for human health due to (1) the known risk of exposure to human waste and (2) the public and regulatory will to reduce sewage pollution; however, methods to identify animal sources are receiving increasing attention as our understanding of zoonotic disease potential improves. Here, we review the performance of MST methods in initial reports and field studies, with particular emphasis on quantitative PCR (qPCR). Relationships among human-associated MST markers, fecal indicator bacteria, pathogens, and human health outcomes are presented along with recommendations for future research. An integrated understanding of the advantages and drawbacks of the many MST methods targeting human sources advanced over the past several decades will benefit managers, regulators, researchers, and other users of this rapidly growing area of environmental microbiology.

Introduction
Establishing the need for microbial source tracking (MST)

Limitations of fecal indicator bacteria (FIB)

Sewage contamination of water bodies poses a definite risk to human health via waterborne pathogens (Cabelli et al., 1982; Cheung et al., 1990; Wade et al., 2003). Monitoring for all waterborne pathogens in environmental waters is currently unrealistic due to the great diversity of pathogens that are known to be present in sewage (including viruses, bacteria, and protozoa), and the disparate methods required for concentrating and analyzing them, yet monitoring for only one or a handful of pathogens may give a false impression of safety if pathogens other than those tested are present. Furthermore, many pathogens are difficult and costly to culture, are difficult to identify, and have patchy distributions or low concentrations in environmental waters (Field & Samadpour, 2007; Stoeckel & Harwood, 2007). For over a century, the approach to this problem has been to monitor for FIB, which have been selected due to low pathogenic potential, high levels in sewage and feces, and relationship to pathogen presence. The major FIB used worldwide include fecal coliforms, *Escherichia coli*, and enterococci (see (Leclerc et al., 2001; Tallon et al., 2005; Griffith et al., 2009) for reviews).

The use of FIB as surrogates for human health risk is replete with assumptions, one of the most important of
which is that FIB consistently covary with pathogen presence; however, FIB concentrations have not been well correlated with pathogens in many studies (Lund, 1996; Bonadonna et al., 2002; Lemarchand & Lebaron, 2003; Anderson et al., 2005; Harwood et al., 2005). This lack of correlation is generally attributed to the widely differing physiology and phylogeny of FIB and pathogens (comprised of bacteria, enteric viruses, and protozoans such as Cryptosporidium spp. and Giardia spp.). Furthermore, many epidemiology studies have failed to find a correlation between human health outcomes and FIB levels, particularly when the pollution is not from a known point source such as a wastewater treatment plant (WWTP) (Prieto et al., 2001; Dwight et al., 2004; Colford et al., 2007). The disconnect can be attributed, at least in part, to the fact that E. coli, enterococci, and other FIB are shed in the feces of many different animals (Harwood et al., 1999; Souza et al., 1999; Leclerc et al., 2001). Due to sometimes limited cohort sizes and FIB levels, some epidemiological studies simply lack the sensitivity to determine significant relationships between health risk and FIB concentration (McBride et al., 1993; Boehm & Soller, 2012). Many waterborne pathogens, particularly viruses, infect only humans. Much of the waterborne disease burden in developed countries is attributed to viral infections (Hopkins et al., 1984; Bosch et al., 1991; Reynolds et al., 2008); therefore, human sewage contamination is considered very risky (Bosch, 1998; Lodder et al., 1999; Field & Samadpour, 2007). Elevated FIB concentrations resulting from input from comingled sources where human sewage is a minimal factor, such as stormwater, generally have relatively lower human health risks than sewage-impacted waters (Ferguson et al., 1996; Wade et al., 2008). The FIB paradigm does not allow one to determine or account for differential human health risk from various fecal sources, as most animals shed these bacteria in their feces.

Added to the issue of differential risk from various contamination sources is the capability of naturalized or environmentally adapted strains of FIB to persist in many habitats, including terrestrial soils, aquatic sediments, and aquatic vegetation (Byappanahalli & Fujioka, 1998; Solo-Gabriele et al., 2000; Byappanahalli et al., 2003; Topp et al., 2003; Whitman et al., 2003; Jeng et al., 2005; Ishii et al., 2006; Ksoll et al., 2007; Badgley et al., 2011). FIB persistence, and possibly growth, in these environmental habitats widens the disconnect between FIB and pathogens. Furthermore, the near-ubiquitous distribution of FIB among host species prohibits the identification of sources of contamination, which in turn interferes with remediation of polluted waters.

**MST and applications**

MST emerged at the end of the 20th century (Wiggins, 1996; Parveen et al., 1997; Hagedorn et al., 1999; Bernard & Field, 2000a; Harwood et al., 2000) as an attempt to determine the dominant sources of fecal contamination in environmental waters (Scott et al., 2005; Field & Samadpour, 2007; Stoeckel & Harwood, 2007; Harwood et al., 2009). The impetus for emergence of this research area derives from (1) the effort to determine the extent to which fecal source (e.g. human, dog, cattle) influences human health risk from contact with water and (2) the desire to attribute FIB loading in water bodies to the correct fecal sources. The basic premise of MST is that certain fecal microorganisms are strongly associated with particular hosts and that identified attributes of these host-associated microorganisms can be used as markers for fecal contamination from the host. The utility of two basic strategies have been tested (1) library-dependent analyses that required the collection and typing of many FIB isolates for some identifying attribute, including antibiotic resistance (Parveen et al., 1997; Harwood et al., 2000), carbon source utilization (Hagedorn et al., 1999), or genetic type (Parveen et al., 1999; Moore et al., 2005) and (2) library-independent analyses that target a particular feature of a specific bacterial species or type, for example, a variable region of the 16S rRNA gene of Bacteroidales (Bernhard & Field, 2000b).

The number and range of potential host sources included in MST studies must be deliberately chosen to suit the water body and particular questions associated with it (Harwood & Stoeckel, 2011). In the case of library-dependent methods, known fecal sources sampled determine the host range of library-dependent methods, that is, one would isolate and type FIB from the feces of the animals judged most likely to contaminate the area. Library-independent methods must rely upon a separate analytical method for each host source to be investigated. The most broadly useful methods employ host-specific markers that can be applied in diverse geographic settings. These methods usually omit a culture step, and so they yield same-day results that can allow timely management decisions for applications such as beach closures (Wade et al., 2006; Field & Samadpour, 2007).

MST applications are diverse, ranging from assessment of beach water quality (Brownell et al., 2007; Abdelzaher et al., 2010; Korajkic et al., 2011) to source allocations for total maximum daily load plans (Simpson et al., 2002; Arnone & Walling, 2007), the legal arena (Weidhaas et al., 2010; Teaf et al., 2011), and food safety (Santo Domingo & Sadowsky, 2007; Graves, 2011). Likewise, the host species implicated as potential pollution sources in
various water bodies range from human (Bernhard & Field, 2000b; McQuaig et al., 2006) to agricultural animals (Bernhard & Field, 2000b; Shanks et al., 2008; Weidhaas et al., 2010) to pets (Kildare et al., 2007) and wild animals such as gulls (Sinigalliano et al., 2010; Lu et al., 2011). Many MST publications have focused solely or in part on human source contamination, as this issue tends to be the greatest concern for managers and regulators, resulting in a confusing proliferation of MST methods for markers designed to detect human waste.

This review focuses on library-independent methods for human source contamination and includes a brief overview of methods for identifying fecal contamination from other animals. The review emphasizes qPCR methods with the rationale that the movement in the research area is toward these quantitative methods; the consequences of which are that (1) the more promising end-point PCR methods have been adapted as qPCR methods and (2) much of the new method development in the last several years has been for qPCR methods; however, important studies using end-point PCR are included. This review substantially expands the ground covered by others (Bernhard & Field, 2000b; Scott et al., 2005; McQuaig et al., 2006, 2009; Ufnar et al., 2006; Field & Samadpour, 2007; Shanks et al., 2010a) by focusing on a critical review of the performance of the methods and the correlation between MST markers and pathogen detection or human health risk.

Performance criteria for method validation

Throughout this review, the terminology used will be consistent with that in Stoeckel and Harwood (Stoeckel & Harwood, 2007). The validation of MST markers (see Fig. 1) is accomplished by assessing certain performance criteria, including sensitivity and specificity (Bernhard & Field, 2000b; Scott et al., 2005; McQuaig et al., 2006; Ufnar et al., 2006; Stoeckel & Harwood, 2007). Sensitivity refers to the proportion of known positive samples (e.g. positive controls) that are correctly identified as positive (true-positive rate). Sensitivity of a method may be influenced by physical or chemical properties of the matrix or sample type (e.g. particulate matter, inhibitors such as humic acids) (Haugland et al., 2005; Siefring et al., 2008); therefore, the intrinsic sensitivity of a method (distribution of the marker in the population) is most accurately assessed in a neutral matrix, that is, by adding feces or sewage to a buffer solution (Harwood et al., 2009). Sensitivity should also be measured in a variety of water types found in the study region in order to determine whether the method is subject to inhibition from constituents of the water (Harwood et al., 2011). Because sensitivity may vary due to geographic variability of the distribution of the marker in the human population, sensitivity is often determined across a geographic range, and should always be verified in a new geographic region (Bernhard & Field, 2000b; Scott et al., 2005; McQuaig et al., 2006; Ufnar et al., 2006; Stoeckel & Harwood, 2007; Harwood et al., 2009).

The limit of detection (LOD) is a quantitative or semi-quantitative expression of the lowest amount of target that can be detected. A useful MST method must not only be sensitive to low target numbers in the analytical stage, but it must also have a target that is at high enough levels in feces or waste so that it can be diluted and still be detected (Stoeckel & Harwood, 2007; Harwood & Stoeckel, 2011). LOD of the analytical stage is frequently calculated as the minimum number of gene copies as per PCR that can be reliably detected (Bernhard & Field, 2000b; McQuaig et al., 2006, 2009; Kildare et al., 2007; Kirs & Smith, 2007; Okabe et al., 2007; Reischer et al., 2007; Wolf et al., 2008); however, this method has little relevance to the performance of the method in the field. It can also be calculated as the amount of source material (typically sewage or feces) required to reliably obtain a positive result (Seurinck et al., 2005; Betancourt & Fujioka, 2006; Layton et al., 2006; McQuaig et al., 2006; Harwood et al., 2009), or as the maximum dilution factor at which the target can be reliably detected (Harwood et al., 2009), both of which provide better insight into the methods’ potential for performance in the field. Similarly, the limit of quantification (LOQ) is fewest gene copies or the least amount of fecal material that can be accurately quantified, which is frequently an order of magnitude greater (less sensitive) than the LOD.

The literature is confusing with respect to the methodological limits, as some are expressed as per assay (how little can be in the tube and still be detected or quantified), while others are expressed as protocol limits of detection, which include loss through processing steps. Another issue is the fact that some markers, like the 16S rRNA gene, are present in multiple copies as per genome, while others are single-copy genes. These differences must be taken into account when converting from gene copies to genome equivalents; the latter is the most direct comparison to the number of microorganisms present. Yet another potential point of confusion is how quantitative results are normalized, that is, one may express results as gene copies (or genome equivalents) per ng DNA, or one may normalize to a more familiar unit in environmental microbiology such as gene copies in 100 mL water. Normalizing the results from fecal matter is also an issue; in some instances, LOD or LOQ is normalized to ng DNA, to dry weight of feces, to wet weight of feces, or to dilution factor. Clearly, some agreement among MST researchers is needed on this issue to be able to compare method performance.
Fig. 1. Diagram of the process of developing and validating a new MST marker. Suggested sample size for Level 2 sensitivity testing is a minimum, based on the authors’ experience and values in the literature, and may be considerably larger depending upon the desired geographic coverage and confidence in method performance. *The threshold for rejecting a candidate human-associated marker (HAM) HAM based on false-negative (Level 1 sensitivity) or false-positive (Level 2 specificity) results also depends upon the desired confidence in the method; in practice, many methods with c. 90% specificity and even lower sensitivity have proven useful for MST.
Method specificity is defined as the proportion of fecal/waste samples from nontarget hosts (species other than those whose waste the method is designed to detect) that produce negative test results, or 1 minus the false-positive rate (Stoeckel & Harwood, 2007). Fecal and/or sewage samples are the reference material generally used for negative controls for specificity testing. No standards regarding the number of reference samples or host types used for specificity testing have been established, and a wide range of these parameters has been employed in various studies, (Bernhard & Field, 2000b; Scott et al., 2005; Ufnar et al., 2006; Stoeckel & Harwood, 2007; Griffith et al., 2009; Harwood et al., 2009; McQuaig et al., 2009; Mieszkin et al., 2009; Shanks et al., 2010a), thus the appropriate procedure for measuring this important criterion remains undefined. Certainly, analyzing as many nontarget fecal samples as possible for specificity is desirable, as it provides more confidence that limitations to method specificity are understood (Harwood & Stoeckel, 2011). Some of the MST methods discussed below showed enviable specificity in initial testing; however, when more fecal samples were tested, the presence of the marker in other animal feces was determined.

A recent trend has been to adapt end-point (presence/absence) PCR methods to the quantitative PCR (qPCR, also called real-time PCR) format, or to develop new qPCR methods (Dick & Field, 2004; Seurinck et al., 2005; Kildare et al., 2007; McQuaig et al., 2009). qPCR is rapid, does not require gel electrophoresis of PCR amplicons, and measures fluorescence with each PCR cycle to quantify the specific DNA sequence targeted by the method (Walker, 2002). If a probe is used in addition to primers, for example, TaqMan® chemistry, additional specificity can be gained. An important performance characteristic of qPCR is amplification efficiency, which is measured from the standard curve and estimates the multiplication factor of the DNA sequence with each qPCR cycle. Ideally, amplification efficiency is 1.0 (gene copies doubled with each round of qPCR, also expressed as 100%); however, this parameter may be affected by factors such as reagent concentrations, PCR cycling conditions, primer design, and amplicon size (Heid et al., 1996). Amplification efficiency ($E_a$) can be estimated from the slope of the standard curve by the equation $E_a = (10^{-1/slope}) - 1$ (Deprez et al., 2002). We are unaware of any guidance on acceptable ranges for amplification efficiency; however, values closer to 1.0 indicate greater efficiency.

The section entitled ‘Initial evaluation of methods’ describes the performance characteristics observed in initial publications of methods for human-associated markers (HAM). Table 1 provides specifics on sensitivity, specificity, LOD, and LOQ, as well as a notation on whether a field study was performed, while Fig. 2 provides a visual overview of the phytype/gene target of various methods for detection of human waste. The section ‘Further evaluation of methods’ describes studies that further test method performance, including comparative studies. Some of the studies included in this section incorporated substantial field validation, which is reported under ‘Selected Field Studies’ and captured for human-associated methods in Fig. 2. Note that in some cases, initial studies showed highly promising results for new methods, only to reveal important flaws with further testing (Further evaluation of methods). Throughout the Evaluation sections, correlations with FIB are noted when applicable. The penultimate section, ‘Correlation with pathogens and risks of waterborne illness’ focuses on correlation of MST marker detection with pathogens and/or human health risk, and the final section provides a brief synopsis of MST methods for animal fecal sources.

### Initial evaluation of methods

**Human-associated Bacteroidales**

One of the first library-independent methods developed for the detection of human fecal contamination was based on end-point PCR detection of the 16S rRNA gene of human-associated *Bacteroidales* (Bernhard & Field, 2000b). Some detail is devoted to this method because it served as a precursor to many of the human-associated methods discussed below and in Table 1. The obligate anaerobic *Bacteroides-Prevotella* taxon was targeted due to high concentrations in feces and tendency to coevolve with the host (Bernhard & Field, 2000b). A forward primer targeting apparently human-specific members of the *Bacteroidales* was paired with a general reverse primer (Bac708R) targeting a much broader phylogenetic group. Two ‘human-specific’ forward primers HF134F and HF183F were designed, the most useful of which was HF183F due to its sensitivity and specificity. The HF183F PCR product was amplified by a second round of PCR to increase sensitivity, which was over 84% ($n = 13$) in human fecal samples and 100% ($n = 3$) in sewage. Specificity, determined by testing negative control fecal samples from domestic and wild animals ($n = 46$), was 100%. The LOD was $10^{-12}$ g DNA or $10^5$ gene copies. Sewage could be detected at a concentration as low as $1.4 \times 10^{-6}$ g L$^{-1}$ (dry weight/volume), which was one order of magnitude higher (less sensitive) than the LOD for fecal coliforms. This marker, which is frequently termed HF183 in the literature and whose sequence has been found in the 16S rRNA gene of the cultured *Bacteroides dorei* (Haugland et al., 2010), has since been used in many field tests. It has also been extensively compared with other human markers described below.
<table>
<thead>
<tr>
<th>Citation</th>
<th>Gene target</th>
<th>Primers and probes</th>
<th>Sensitivity* (n)</th>
<th>Specificity† (n)</th>
<th>Reference types</th>
<th>Limit of detection‡</th>
<th>Limit or range of quantification§</th>
<th>Field testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seurinck et al. (2005)</td>
<td>16S rRNA gene</td>
<td>SYBR Green HF183 SShBac-R</td>
<td>86% (7); 100% SEW (4)</td>
<td>95% (19)</td>
<td>SEW; Feces: HUM, cow, Hr, dog, Ch</td>
<td>≥ 1 ng wet feces L⁻¹ in freshwater</td>
<td>10⁻²–10⁻⁷ copies µL⁻¹ (feces); 4.7 x 10⁻⁵ copies L⁻¹ (freshwater)</td>
<td>No</td>
</tr>
<tr>
<td>Layton et al. (2006)</td>
<td>16S rRNA gene</td>
<td>TaqMan HuBac566f HuBac692r HuBac594Bhq1 probe</td>
<td>100% (3)</td>
<td>68% (18)</td>
<td>Feces: HUM, cow, pig, Hr, dog</td>
<td>0.3 mg feces L⁻¹</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td>Kildare et al. (2007)</td>
<td>16S rRNA gene</td>
<td>TaqMan BacHum-160f BacHum-241r BacHum-193p probe</td>
<td>67% feces (18); 100% SEW (14)</td>
<td>97% (41)</td>
<td>SEW; Feces: HUM, cow, Hr, dog, cat, gull</td>
<td>3.5 copies**</td>
<td>30 copies</td>
<td>Yes</td>
</tr>
<tr>
<td>Reischer et al. (2007)</td>
<td>16S rRNA gene</td>
<td>TaqMan Bac_H_f Bac_H_r Bac_HPC probe Bac_H_PT probe</td>
<td>95% feces (21); 100% wastewater (20)</td>
<td>&gt; 99% (302)</td>
<td>SEW; Feces: HUM, cow, pig, Hr, Sh, goat, Ch, Tu, dog, cat, 3 bird species, fox, 3 ruminants</td>
<td>6 copies</td>
<td>30–10⁷ copies</td>
<td>Yes</td>
</tr>
<tr>
<td>Okabe et al. (2007)</td>
<td>16S rRNA gene</td>
<td>TaqMan q5601f q5625r q56124MGB probe</td>
<td>100% (3)</td>
<td>10% (10)</td>
<td>Feces: HUM, cow, pig</td>
<td>4.3 copies</td>
<td>4.3–4.3 x 10⁵ copies</td>
<td>Yes</td>
</tr>
<tr>
<td>Silkie &amp; Nelson (2009)</td>
<td>16S rRNA gene</td>
<td>BacHum-160f BacHum-241r BacHum-193p probe</td>
<td>100% (12)</td>
<td>70% (41)</td>
<td>SEW, dog, cow, Hr, Canada geese (pooled samples)</td>
<td>ND</td>
<td>ND</td>
<td>No</td>
</tr>
<tr>
<td>Ahmed et al. (2009a)</td>
<td>16S rRNA gene</td>
<td>SYBR Green TaqMan</td>
<td>All assays: 100% (50)</td>
<td></td>
<td>SEW; Feces: HUM, cow, pig, Hr, Sh, goat, Ch, dog, 2 bird, kangaroo</td>
<td>Copies: HF183: 1 BacHum: 1 HuBac: 7 BacH: 5 Human-Bac: 3</td>
<td>ND</td>
<td>No</td>
</tr>
</tbody>
</table>
Table 1. Continued

<table>
<thead>
<tr>
<th>Citation</th>
<th>Gene target</th>
<th>Primers and probes</th>
<th>Sensitivity* (n)</th>
<th>Specificity† (n)</th>
<th>Reference sample types</th>
<th>Limit of detection‡</th>
<th>Limit or range of quantification§</th>
<th>Field testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jenkins et al. (2009)</td>
<td>16S rRNA gene</td>
<td>TaqMan BacHum-UCD HF183</td>
<td>BacHum-UCD: 25% human feces (12), 0% SEW (5)</td>
<td>Both assays: 100% (25)</td>
<td>SEW; feces: HUM cow, donkey</td>
<td>ND</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td>Lee et al. (2010)</td>
<td>16S rRNA gene</td>
<td>BacHuman</td>
<td>TaqMan BacHuman-F BacHuman-R BacHum-TP probe</td>
<td>100% (16)</td>
<td>81.5% (54); 1.9% of true positives if quantified relative to Bact spp. marker</td>
<td>SEW; Feces: cow, pig, Hr, dog, cat, deer, goose, gull, racoon</td>
<td>6.5 copies to 10⁷ copies</td>
<td>Yes</td>
</tr>
<tr>
<td>Haugland et al. (2010)</td>
<td>16S rRNA gene</td>
<td>HF183 BunF2 BfragF1 BvulgF1 PcopriF1 BsteriF1 BthetaF2</td>
<td>All assays: 100% (1 feces composite, 14 SEW)</td>
<td>60%</td>
<td>SEW; Feces composites: HUM, cow, pig, Ch, dog, cat</td>
<td>ND</td>
<td>10–10⁴ copies</td>
<td>No</td>
</tr>
<tr>
<td>Ahmed et al. (2010)</td>
<td>16S rRNA gene</td>
<td>HuBac HumM2 HumM3 HF183 BunF2 BfragF1 BvulgF1 PcopriF1 BsteriF1 BthetaF2</td>
<td>TaqMan HuBac HumM2F HumM3F HF183-1 BunF2B fragF1 BvulgF1 PcopriF1 BsteriF1 BthetaF2</td>
<td>87% (15)</td>
<td>SEW (54); (22 composite)++</td>
<td>HUM, cow, dog, cat, Ch</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Shanks et al. (2010a)</td>
<td>16S rRNA gene</td>
<td>HuBac</td>
<td>TaqMan HuBac HUM (16)</td>
<td>93% (30)</td>
<td>SEW; Feces: HUM cow, pig, Sh, goat, Ch, Tu dog, cat, 4 bird, 5 ruminant, 3 marine mammals</td>
<td>ND</td>
<td>10–10⁴ copies</td>
<td>No</td>
</tr>
<tr>
<td>Citation</td>
<td>Gene target</td>
<td>Primers and probes</td>
<td>Sensitivity* (n)</td>
<td>Specificity† (n)</td>
<td>Reference sample types</td>
<td>Limit of detection‡</td>
<td>Limit or range of quantification§</td>
<td>Field testing</td>
</tr>
<tr>
<td>--------------------</td>
<td>--------------------------------------</td>
<td>--------------------</td>
<td>---------------------------</td>
<td>---------------------------</td>
<td>--------------------------------------------</td>
<td>--------------------</td>
<td>-------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Yampara-Iquise et al. (2008)</td>
<td>B. thetaiotomicron α-mannanase B. theta x</td>
<td>TaqMan Bth-F Bth-R Bth-P probe</td>
<td>100% feces (10); 100% SEW (20)</td>
<td>100% (160)</td>
<td>SEW; Feces: HUM, cow, pig, HR, Ch, Tu, dog, goose</td>
<td>9.3 copies</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td>Shanks et al. (2009)</td>
<td>Hypothetical protein HumM2 Putative RNA polymerase sigma factor HumM3</td>
<td>TaqMan Hum2F Hum2R HumM2P probe Hum3F Hum3R HumM3P probe</td>
<td>Both assays: 100%, (16 feces, 20 SEW)</td>
<td>HumM2: 99.2% HumM3: 97.2% (265)</td>
<td>SEW; Feces: HUM, cow, pig, HR, Sh, Ch, Tu, 7 ruminant, dog, cat, 4 bird, seal, lion, dolphin</td>
<td>ND</td>
<td>10–1 × 10⁶ copies</td>
<td>Yes</td>
</tr>
<tr>
<td>Lee &amp; Lee (2010)</td>
<td>B. fragilis gyrB</td>
<td>TaqMan Bf904F Bf958R Bf923MGB probe Human-Bac1</td>
<td>Both assays 100% (10)</td>
<td>gyrB: 97% (30) Human-Bac1: 57%</td>
<td>Feces: HUM cow, pig, dog</td>
<td>110 CFU per reaction</td>
<td>ND</td>
<td>No</td>
</tr>
<tr>
<td>Bifidobacterium</td>
<td>Gourmelon et al. (2010b)</td>
<td>16S rRNA gene</td>
<td>TaqMan W257F W255R W256P probe</td>
<td>90% feces (10); 100% WWTP effluent (8)</td>
<td>WWTP effluent; Feces: cow, pig, HR, Sh, Ch, wild bird</td>
<td>ND</td>
<td>10³ copies per 100 mL</td>
<td>Yes</td>
</tr>
<tr>
<td>Enterococcus</td>
<td>Ahmed et al. (2008)</td>
<td>Enterococcal surface protein (esp-f)</td>
<td>SYBR Green EspF Esp-p2</td>
<td>67% septic (12); 100% SEW (30)</td>
<td>SEW; Feces: cow, pig, HR, Sh, goat, Ch, dog, deer, duck, pelican, wild bird, kangaroo</td>
<td>10 copies</td>
<td>to 10⁷ copies</td>
<td>Yes</td>
</tr>
<tr>
<td>Methanobrevibacter smithii</td>
<td>Johnston et al. (2010)</td>
<td>nifH</td>
<td>TaqMan Mnf 202F Mnf 253R Mnf probe</td>
<td>100% (16)</td>
<td>100% (2.3) methanogens; 50% (4) bird guano</td>
<td>2.5 genome equivalents</td>
<td>5 genome equivalents</td>
<td>Yes</td>
</tr>
<tr>
<td>F* RNA Coliphages</td>
<td>Kirs &amp; Smith (2007)</td>
<td>Replicase and coat protein</td>
<td>Molecular Beacon Subgroup I-H F, R, and beacons</td>
<td>ND</td>
<td>SEW, Ch</td>
<td>10 copies; 50 copies multiplex</td>
<td>ND</td>
<td>No</td>
</tr>
<tr>
<td>Wolf et al. (2010)</td>
<td>Viral genome</td>
<td>TaqMan VTB4-Fph G-J-GV primer/ probe series</td>
<td>ND</td>
<td>ND</td>
<td>SEW; Feces: HUM cow, pig, Sh, deer, 3 bird</td>
<td>&lt; 10 copies</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td>Table 1. Continued</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td><strong>Gene target</strong></td>
<td><strong>Primers and probes</strong></td>
<td><em><em>Sensitivity</em> (n)</em>*</td>
<td><strong>Specificity† (n)</strong></td>
<td><strong>Sample types</strong></td>
<td><strong>Limit of detection‡</strong></td>
<td><strong>Limit or range of quantification§</strong></td>
<td><strong>Field testing</strong></td>
</tr>
<tr>
<td>Pepper Mild Mottle Virus</td>
<td>Rosario et al. (2009)</td>
<td>Viral genome</td>
<td>TaqMan PMMMV-FP PMMMV-RP PMMMV-Probe</td>
<td>100% SEW (12 raw, 12 treated)</td>
<td>90% pooled (20) 92% individual (24)</td>
<td>SEW; Feces: Ch, cow, dog, Hr, gull, sheep, pig, feral pig, coyote, raccoon; Intestine: Ch, Tu</td>
<td>10^3 copies</td>
<td>ND</td>
</tr>
<tr>
<td>Polyomaviruses JC and BK</td>
<td>McQuaig et al. (2009)</td>
<td>T antigen</td>
<td>TaqMan SM2 P6 KG3 probe</td>
<td>100% septic (14); 100% SEW (41); 23% urine (26)</td>
<td>100% (117) feces; 100% (10) urine</td>
<td>SEW; Feces: cow, pig, Hr, Ch, dog, cat, deer, 4 bird, fox, raccoon; Urine: dog, cat</td>
<td>10 copies</td>
<td>ND</td>
</tr>
<tr>
<td>Polyomavirus JC</td>
<td>Albinana-Gimenez et al. (2009)</td>
<td>T antigen</td>
<td>TaqMan JE3F JE3R JE3P</td>
<td>ND</td>
<td>ND</td>
<td>viral particles, river water</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Pathogenic Viruses</td>
<td>Wolf et al. (2010)</td>
<td>Viral genomes</td>
<td>VTB-1‡‡: PAdV-3 OAdV-3</td>
<td>60% (5)</td>
<td>100% (56)</td>
<td>Feces: HUM, Pig, Sh, deer, cow, Canada goose, black swan, duck</td>
<td>10 copies</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HAdV-F</td>
<td>50% (4)</td>
<td>100% (67)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VTB-2: HAdV-C AtAdV PAdV-S</td>
<td>0% (15)</td>
<td>100% (56)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VTB-3: Norovirus GI</td>
<td>40% (15)</td>
<td>100% (56)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Norovirus GII</td>
<td>50% (20)</td>
<td>100% (51)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Norovirus GIII</td>
<td>75% (4)</td>
<td>100% (67)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GI 9% (56)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VTB-4: F+ RNA bacteriophage genogroups</td>
<td>GII 0% (15)</td>
<td>GII 93% (15)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GII 0% (15)</td>
<td>GII 87.5% (56)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GIII 100% (56)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GIV 100% (15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Citation</td>
<td>Gene target</td>
<td>Primers and probes</td>
<td>Sensitivity* (n)</td>
<td>Specificity† (n)</td>
<td>Limit of detection²</td>
<td>Limit or range of quantification§</td>
<td>Field testing</td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------------------------</td>
<td>--------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>---------------------</td>
<td>-----------------------------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td><strong>Mitochondria</strong></td>
<td><strong>Caldwell et al. (2007)</strong></td>
<td>NADH dehydrogenase subunit 5 (ND5)</td>
<td>TaqMan Human forward Human reverse Human probe</td>
<td>100% feces (16); 60% wastewater (5)</td>
<td>100% (≥ 16)</td>
<td>SEW; Feces: HUM Hr, Sh, goat, Ch, Tu, goose, dog, cat, tilapia</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td><strong>Schill &amp; Mathes (2008)</strong></td>
<td>Cytochrome b (mtCytb)</td>
<td>SYBR Green Human forward Human reverse Human probe</td>
<td>100% (2 feces, 10 SEW)</td>
<td>100% (18)</td>
<td>SEW; Feces: HUM cow, pig, Hr, Sh, Ch, goose, dog, deer</td>
<td>10⁻¹⁻¹⁰⁸ copies µL⁻¹ (multiplex)</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

SEW, human sewage; HUM, human; Hr, horse; Ch, chicken; Tu, turkey; Sh, sheep.

* Sensitivity is reported as the percentage of human waste samples returning positive results (true positives). Sample types are separated to reflect resulting variation in sensitivities.
† Specificity is reported as the percentage of negative results for animal waste samples (true negatives).
‡ Limit of detection is reported as the lowest concentration or target copy number at which a positive result was obtained.
§ Where applicable, the range at which the assay was tested and returned results corresponding to the estimated known quantities of target. Single values indicate the lower limit of quantification.
¶ Validation statistic was not calculated.
** Copies designate gene copies/reaction.
†† qPCR methods were tested for specificity against 22 composite fecal samples.
‡‡ All targets specified were amplified using unique forward and reverse primers and probes.
§§ Sensitivity and specificity were tested against individual target and nontarget fecal samples, respectively.
Methods targeting the 16S rRNA gene of Bacteroidales

Many qPCR methods targeting the 16S rRNA gene of human-associated Bacteroidales have been developed (Seurinck et al., 2005; Layton et al., 2006; Kildare et al., 2007; Okabe et al., 2007; Reischer et al., 2007; Lee et al., 2010; Shanks et al., 2010a) (Figs 2 and 3). Several of these primers overlap in the short 600-bp region targeted by most assays, and most of the primers and probes described show 100% identity to portions of the 16S rRNA gene of B. dorei (Fig. 3). In this review, we will refer to methods that target this sequence and closely related ones as 'human-associated Bacteroides.' Less closely related or well-defined groups will be referred to by the designation of the Order Bacteroidales. Differences in sensitivity or specificity among assays may result from the location of primers and probes at positions that are heterogeneous among other members of this genus. The characteristics and performance of these methods are summarized in Table 1, including performance measurements of sensitivity, specificity, LOD, and LOQ.

Seurinck et al. developed a qPCR method based on SYBR Green chemistry using the HF183 primer and a new reverse primer that specifically targeted the human-associated sequence (Seurinck et al., 2005). The authors did not name the new reverse primer; here, we designate it SSHBac-R. The method was highly sensitive and specific (Table 1); however, one false-positive result was generated by a chicken fecal sample. As is frequently the case with HAM, the distribution in individual fecal samples was not ubiquitous (6 of 7 samples positive). Table 1 shows the LOD and LOQ of the qPCR method. End-point PCR using the same primer set was unexpectedly two orders of magnitude less sensitive than the qPCR method, that is, c. $10^2$ vs. $10^3$ gene copies L$^{-1}$, respectively.

Another method targeting human-associated Bacteroides proved to be less specific. A TaqMan© qPCR method (designated HuBac) showed 100% sensitivity for human fecal samples (Layton et al., 2006), but specificity was only 68% when tested against domestic animal fecal samples, showing cross-reactivity with all species (Table 1). Further studies of the performance of this marker (see ‘Further evaluation of methods’) have confirmed the imperfect specificity of this method.

The TaqMan qPCR method designated BacHum-UCD (Kildare et al., 2007) was much more specific than the HuBac method (Table 1). Although its sensitivity toward human fecal samples was imperfect at 67%, it was 100% sensitive to raw sewage samples. This work also compared the new BacHum-UCD method with existing qPCR methods, finding that the LOQ for the SYBR Green HF183 qPCR method described above (Seurinck et al., 2005) was about 10-fold more sensitive than BacHum-UCD (3 copies compared to 30). HF183 specificity was slightly inferior to that of BacHum-UCD, showing cross-reactivity to 25% and 14% of dog and cat fecal samples; however, HF183 specificity among all samples tested was 95%. The low specificity of the HuBac TaqMan method (Layton et al., 2006) was also confirmed in this study (61%).

Yet another method that targets the 16S rRNA gene of human-associated Bacteroides is BacH (Reischer et al., 2007). The sensitivity of the method was nearly 100% (Table 1) with a specificity of over 99% (one false-positive cat sample), and LOD and LOQ were comparable to other methods. Field testing was included in this study (see ‘Selected Field Studies’). The Human-Bac1 qPCR method (Okabe et al., 2007) showed 100% sensitivity toward a limited number of human fecal samples; specificity, however, was only 10%, although the marker concentration was several orders of magnitude less for animal samples compared with human samples.

Fig. 2. Meta-analysis showing the number of different existing assays for human-associated MST markers by target species and gene, as well as whether the assays have been used in field studies (*) or are correlated with pathogen presence (▲). Stacked bars represent different gene targets. 1–11 represents Bacteroidales 16S rRNA gene targets (HF183, HuBac, BacHum-UCD, BacH, HumanBac, BuniF2, BfragF1, BvulF1, PcoprF1, BsterF1, and BthetaF2), 12–15 denotes Bacteroidales non-16S rRNA gene targets (HumM2, HumM3, B. theta α-1,6-mannanase, gyrB), 16– Bifidobacterium 16S rRNA gene, 17– Enterococcus faecium esp, 18– Methanobrevibacter smithii nifH, 19– F-RNA coliphage genome, 20– F-RNA replicase coat protein, 21– Pepper mild mottle virus genome, 22– Human polyomavirus t antigen, 23– mitochondrial NADH dehydrogenase subunit 5, and 24– mitochondrial cytochrome b.
The BacHuman qPCR method targets similar 16S rRNA gene sequences (Lee et al., 2010). It was found to be 100% sensitive to sewage, but it was not tested on individual human fecal samples and was only 81.5% specific (Table 1). As an alternate measurement of specificity, Lee et al. determined the percentage of BacHuman copies relative to general Bacteroides marker (BacGeneral) copies in each fecal sample. In false-positive analyses, the relative quantity of the BacHuman target represented c. 1.8% of the general Bacteroides target detected in a fecal sample, while the relative proportion of BacHuman detections in human samples (true-positive analyses) was nearly 50 times larger. Field testing showed weak correlation of total Bacteroidales with E. coli (Supporting Information, Table S1), and the BacHuman marker comprised a greater percentage of the overall Bacteroidales signal in a stream contaminated with human sewage as opposed to a stream contaminated with cattle feces in which the BacHuman marker represented < 1% of the total Bacteroidales signal.

The alternative approach chosen by Haugland, et al. (Haugland et al., 2010) relied on cultured members of the Bacteroidales rather than gene sequences from uncultured bacteria. This work developed several more TaqMan qPCR methods targeting the hypervariable V2 region of the 16S rRNA gene (Haugland et al., 2010). The authors reasoned that methods targeting cultured species may be valuable in MST because calibration can then be accomplished using genomic DNA rather than a plasmid, and the qPCR results can then be compared against culture-dependent methods. Primer sets were developed to target six different Bacteroides species,
including B. dorei and one Prevotella species. Forward primers were HF183, BsteriF1, BuniF2, BfragF1, BthetaF2, BvulgF1, and PcopriF1 markers. When the methods were tested using sewage and composited fecal samples from humans and domestic animals, all species markers were present in human waste in high abundance (3.17–4.3 log_{10} copy number ng\(^{-1}\) total DNA). None of the markers were completely specific to humans, and the presence and abundance of each marker varied among different animals (Table 1). Performance of the HF183 marker was superior in that it was abundant in human feces, was detected in low quantity in chicken and dog feces (0.35 and 0.36 log_{10} copy number per 1 ng total DNA, respectively), and was absent from cow, pig, and cat waste.

### Other Bacteroidales gene targets

**Bacteroides** genes other than 16S rRNA gene have been incorporated into MST methods, for example, a TaqMan qPCR assay targeting the single-copy α-1,6-mannanase gene of *B. thetaiotaomicron* (BT) (Yampara-Iquise et al., 2008). The sensitivity of the method was 100% among human fecal samples and sewage, and it was 100% specific against composite fecal samples from domestic animals and wildlife collected from across Missouri. Srinivasan et al. recently evaluated this marker relative to *E. coli* and enterococci levels by qPCR through wastewater and septage treatment (Srinivasan et al., 2011), finding that its removal through wastewater treatment was similar to that of *E. coli* and enterococci.

Shanks et al. developed two TaqMan qPCR methods for the detection of human-associated *Bacteroidales* that targeted a hypothetical protein (HumM2) and a putative RNA polymerase sigma factor (HumM3) (Shanks et al., 2009). Both assays showed 100% sensitivity when tested against feces and wastewater sampled throughout the U.S. HumM2 had slightly higher specificity (99.2%, cross-reacting with sheep and elk) compared to HumM3 (97.2%, cross-reacting with chickens). Reactions included an internal amplification control to test for inhibition. The authors found that primary-treated wastewater effluent contained higher concentrations of the human markers than *Enterococcus* 16S rRNA genes.

A different gene, *gyrB* of *B. fragilis*, was the target of a TaqMan qPCR (Lee & Lee, 2010). End-point PCR showed 100% sensitivity among human fecal samples with a specificity of 97% (false positive for one pig sample). The qPCR results showed levels of the marker in negative control fecal samples (cow, dog, and pig) that were 100-fold lower than levels in human feces. Comparison with the Human-Bac1 method (Okabe et al., 2007) again confirmed the low specificity of Human-Bac1, as cross-reaction occurred in samples from 70% of pigs, 40% of cows, and 30% of dogs for overall specificity of only 57%.

### Bifidobacterium

Members of the genus *Bifidobacterium* are anaerobic, gram-positive rods that are abundant in intestinal microbrial communities of humans and some animals (Bonjoch et al., 2004; King et al., 2007). Several end-point PCR methods targeting the 16S rRNA gene have been developed for use in human MST studies (Bonjoch et al., 2004; King et al., 2007; Lamendella et al., 2008), but subsequent research has demonstrated that *B. adolescentis* is not confined to the human gastrointestinal tract. An evaluation of end-point PCR methods targeting *Bifidobacterium* showed that a set of *B. adolescentis* primers amplified fecal DNA from cows, as well as from pigs, sheep, a chicken, a coyote, a deer, and an alpaca (Lamendella et al., 2008). Another study found that the same primers amplified DNA from cows, pigs, and a sheep (Dorai-Raj et al., 2009).

A TaqMan qPCR assay targeting *Bifidobacterium adolescentis* (ADO) (Gourmelon et al., 2010b) showed high sensitivity in human fecal samples and WWTP effluent from France (Table 1). Specificity was nearly 95%, although some false positives were obtained from some cow and bird samples. A comparison of the *B. adolescentis* qPCR method to the *Bacteroides* HF183 qPCR assay (Seurinck et al., 2005) showed that HF183 was more specific (100%) than the *B. adolescentis* method and was not detected in the river samples with known fecal contamination from cattle. F\(^{−}\) RNA coliphage subgroups were also evaluated in this study, but the human-associated groups II and III were found in runoff contaminated with cow feces as well as in sewage, and animal-associated group I coliphages were found in effluent from both humans and nonhumans.

### Enterococcus faecium esp gene

An end-point PCR method targeting the enterococcal surface protein (esp) gene of *E. faecium* was developed to identify human fecal contamination (Scott et al., 2005; Ahmed et al., 2008). This method utilized a culture step to increase target cell numbers. Although the method was 97% sensitive and 100% specific in this study, other studies have detected some level of *Ent. faecium* esp in the feces of animals (Harada et al., 2004; Whitman et al., 2007; Layton et al., 2009).

A SYBR green qPCR method targeting the esp gene that does not require a culture step was developed and tested in Australia (Ahmed et al., 2008). The method employed the previously utilized forward primer specific
to esp of *Ent. faecium* (Scott et al., 2005) and a reverse primer containing a sequence shared by *Ent. faecium* and *Ent. faecalis* (Hammerum & Jensen, 2002). Sensitivity was high, particularly in domestic sewage, while septic system samples, which represent a smaller population, were not as frequently positive (Table 1). Specificity was 100% when tested against twelve types of fecal samples from domestic animals and wildlife in Australia. Sewage dilutions were detectable at dilutions as great as 10⁻⁷, which corresponded to 48 ± 7 CFU cultured enterococci, and esp gene concentrations were generally one to two orders of magnitude lower than cultured enterococci. However, a qPCR method targeting the *Bacteroides* HF183 marker (Bernhard et al., 2003; Seurinck et al., 2005) allowed detection of sewage at much greater dilutions (five orders of magnitude) than the *Ent. faecium* esp method.

### Methanobrevibacter smithii

*Methanobrevibacter smithii*, a methanogenic archaean specific to the human large intestine and vaginal tract, was found in 96% of 700 individual fecal specimens (Ufnar et al., 2006; Dridi et al., 2009). An end-point PCR method targeting the *nifH* gene was only 29% sensitive among individual human fecal samples but showed 93% sensitivity among sewage samples and 100% specificity (Ufnar et al., 2006). The prevalence of *M. smithii* by this method has been further compared against other HAM including human polyomaviruses (HPyVs) and human *Bacteroides* as well as the presence of adenovirus (see ‘Further evaluation of methods’).

A TaqMan qPCR method was subsequently developed that targets a shorter fragment of the *nifH* gene (Johnston et al., 2010) (Table 1). The method utilized a competitive internal control DNA fragment to detect and allow adjustment for PCR inhibition. The method was tested using 23 nontarget methanogenic species, all of which were negative. The qPCR method showed 100% sensitivity in blind-tested marine and freshwater samples containing known sewage or a sewage spike, and it was able to detect the marker at 100-fold lower concentrations than the end-point method; however, cross-reactivity with bird feces was observed (Table 1).

A subsequent study conducted in Australia evaluated the host specificity and sensitivity of the *nifH* gene marker by testing it against fecal and wastewater samples from 11 animal species, including humans (*n* = 272) (Ahmed et al., 2012). The host specificity reported in this particular study was 96%, while sensitivity of the marker to human sewage was 81%. This study also tested prevalence of the marker in environmental water samples and found it to be relatively low compared with other markers tested (esp, HF183, HPyVs, and Adenoviruses). The authors concluded that the *nifH* marker alone may not be sensitive enough to detect fecal pollution in environmental waters but that relatively high host-specificity merits its use in conjunction with other human markers.

### Lachnospiraceae

Newton et al. developed a TaqMan qPCR assay targeting the V6 region of the 16S rRNA gene of members of the family *Lachnospiraceae* that are closely related to *Blautia* spp. (Lachno2) (Newton et al., 2011). This taxon is the second most abundant bacterial group in human fecal samples, making up 0.3–0.9% of the population in wastewater samples. Although no sensitivity or specificity testing was performed against sewage or feces, the concentration of the Lachno2 target was compared with HF183 by qPCR (Bernhard & Field, 2000b; Kildare et al., 2007), culturable *E. coli* and enterococci concentrations, enterococci concentrations enumerated by qPCR, and presence of adenovirus at sites in Milwaukee’s harbor. A strong correlation was observed between concentrations of Lachno2 and HF183 (*r* = 0.86) and Lachno2 and enterococci enumerated by qPCR (*r* = 0.91). Furthermore, logistic regression showed a strong correlation between human marker concentration (Lachno2 plus HF183) and the detection frequency of adenovirus, indicating a link between these markers and human pathogen presence.

### Escherichia coli

Although *E. coli* has been used and studied extensively as an indicator bacterium, less success has been achieved using this organism for MST applications. Substantial efforts were made with earlier library-based techniques such as antibiotic resistance, ribotyping, or DNA fingerprinting (Dombek et al., 2000; Harwood et al., 2000; Carson et al., 2001; Stoeckel & Harwood, 2007), but we know of no library-independent assays in use that are based on host-specific *E. coli* markers. It is worth mentioning, though, that this may change in the near future, as molecular techniques provide increasing power to examine genetic differences in bacterial groups, and recent studies have shown promise in this regard. One study found evidence of a potentially human-specific strain of *E. coli* that belongs to the B2 clonal subgroup VIII with an O81 serotype (Clermont et al., 2008). This clone represented c. 4% of the *E. coli* isolates recovered from human fecal samples from four continents and was not identified from any animal fecal isolates in the study. More recently, full genome sequencing was used to identify several candidate genes that may be specific to *E. coli* strains of either environmental or enteric origin (Luo
et al., 2011). Further work must be carried out to determine the host specificity of these markers.

**F⁺ RNA coliphage genotyping**

Genotyping of F⁺ (F-specific) RNA coliphages has been suggested as a possible tool for discriminating between human and animal sources of fecal contamination (Havelaar et al., 1990; Hsu et al., 1995; Stewart-Pullaro et al., 2006; Gourmelon et al., 2007; Kirs & Smith, 2007; Wolf et al., 2008). The differential host distribution of these coliphages was first demonstrated by serotyping (Havelaar et al., 1990). Among the four subtypes, groups II and III are more prevalent in human sewage, while groups I and IV are generally associated with animal feces (Hsu et al., 1995; Beekwilder et al., 1996; Stewart-Pullaro et al., 2006). A multiplex qPCR for simultaneously typing and quantifying F⁺ RNA coliphages was developed and compared to methods relying on culture of coliphages (Kirs & Smith, 2007) (Table 1). In seeded samples of sewage and chicken slurry, the percentage of each group of coliphage corresponded to the makeup of the sample.

Another multiplex qPCR method was developed to detect and genotype F⁺ RNA coliphages in New Zealand (Wolf et al., 2008); however, no formal sensitivity or specificity testing was conducted. A subsequent study utilized modified primers and probe, and the method was used in a ‘virtual toolbox’ approach that included multiplex qPCR assays for noroviruses, adenoviruses, and atadenovirus (Wolf et al., 2010). F⁺ RNA coliphages were rarely detected in human feces, but were much more prevalent in pig, deer, and cattle feces. In particular, the human-associated group II F⁺ RNA coliphages were prevalent in pig and cattle feces (80% and 50%, respectively). Environmental samples including wastewater influent, shellfish, abbattoir (slaughterhouse) effluent, biosolids, and river water were analyzed. Although the human-associated groups II and III were the most prevalent and at the highest concentration in sewage influent samples, animal-associated group I coliphages were detected in 55% of sewage samples at c. 10⁵ gene copies. Only groups II and III were detected in biosolids. Animal-associated, group IV F⁺ RNA coliphages were only found in sewage (18% of samples); however, abbattoir effluent contained only group I coliphages. The imperfect specificity of the F⁺ RNA coliphage groups toward human or animal fecal waste merits host-associated, rather than host-specific designation for these markers.

**Pepper mild mottle virus**

The plant pathogen pepper mild mottle virus (PMMoV) is an RNA virus that is a major component of the viral metagenome in human feces (Zhang et al., 2006; Rosario et al., 2009). The TaqMan qPCR method (Zhang et al., 2006) showed 100% sensitivity toward raw and treated wastewater collected across the U.S. (Table 1); however, specificity was around 70% due to cross-reactivity with chicken and gull feces or intestinal homogenates (Rosario et al., 2009). This virus shows great persistence through wastewater treatment, although it does decay in environmental waters, with a half-life of about 1.5 days in seawater. PMMoV may therefore be useful as a conservative tracer of sewage contamination.

**HPyVs**

HPyVs are highly specific to humans with a high prevalence in human populations (Taguchi et al., 1982; McQuaig et al., 2009). The viruses are shed primarily in urine, but also in feces, and high titers have been reported in municipal sewage (Bofill-Mas et al., 2000; McQuaig, 2009). An end-point PCR method was developed for the detection of HPyVs (including BK virus and JC virus) in environmental waters (McQuaig et al., 2006). The method showed 100% sensitivity and specificity. The HPyVs method was adapted to a TaqMan qPCR assay following the modification of one of the original primers (McQuaig et al., 2009). The TaqMan qPCR method for HPyVs was evaluated against end-point PCR assays for human Bacteroides HF183 (Bernhard & Field, 2000b), M. smithii (Ufnar et al., 2006), and adenovirus (Pina et al., 1998). The HPyVs qPCR method was 100% specific (Table 1) when tested against fecal and urine samples. HPyVs were detected in 23% of human urine samples and in 100% of sewage and septic system samples. All other markers were also 100% sensitive in sewage but were less sensitive among septic samples. A qPCR assay that targets only polyomavirus JC has been developed; however, sensitivity and specificity were not tested for this method (Albinana-Gimenez et al., 2009).

**Pathogenic viruses**

qPCR assays targeting a variety of viral pathogens (e.g. adenovirus and norovirus) have also been explored as a means of determining sources of fecal contamination (Wolf et al., 2010); however, these pathogens are more commonly targeted as representatives of human health risk rather than source-specific markers (see ‘Correlation with Pathogens’). Four multiplex qPCRs were developed targeting norovirus, adenovirus, atadenovirus, and F⁺ RNA bacteriophage to distinguish among human and animal sources (Table 1). Each assay was able to detect as few as 10 gene copies, and all assays were reported to show amplification of the specific viral target but not
nontarget viruses. When tested against fecal samples from humans and animals, only host-specific viruses were amplified (e.g. only human-specific viruses were detected in human fecal samples). The assays were also tested against environmental samples as well as sewage influent and abattoir effluent, and while some animal viruses were detected in sewage and biosolid samples, human-specific viral targets were detected at higher concentrations. While the assays were shown to have a high level of sensitivity, variable rates of virus shedding, uncertain incidences of viral infection, and survival of viral particles in the environment impair the ability of these methods to accurately reflect source contributions to environmental waters.

Human mitochondrial DNA (mtDNA)

The first ‘fecal source tracking’ method based on a eukaryotic genetic marker was the end-point PCR assay targeting the human mitochondrial NADH dehydrogenase subunit (Martellini et al., 2005). mtDNA was proposed as a marker based upon the premise that it should be abundant in feces (Caldwell et al., 2007). The PCR method was initially developed in both single and multiplex formats with primers specific to humans, cows, sheep, and pigs. Sensitivity and specificity were not explicitly reported; however, the human mtDNA marker was detected in human feces, treated wastewater, and ultraviolet disinfected effluent, but not in swinery effluent. Human, cattle, and pig markers were detected in sewage.

Several groups have adapted human mtDNA PCR methods to qPCR formats (Table 1) (Caldwell et al., 2007; Schill & Mathes, 2008; Caldwell & Levine, 2009). A multiplex qPCR method targeting the mitochondrial NADH dehydrogenase subunit 5 (ND5) gene from humans, cattle, and pigs was developed (Caldwell et al., 2007). Amplification efficiency for the human marker was 92% for a single qPCR and 107% for the triplex reaction. The multiplex assay detected human mtDNA from individual human fecal samples and sewage samples with 100% and 60% sensitivity, respectively. Specificity of the human marker was 100% when tested against cow (n = 4) and pig (n = 3) farm samples and also 100% against an unreported number of fecal samples from nine other animals. Due to some false-negative results in samples from mixed sources, the authors noted that adequate DNA concentration was clearly a concern for detecting mtDNA markers. Some false-positive results for the cattle marker were also noted when testing human feces, which the authors hypothesized could be caused by beef consumption. A later study by the same group confirmed the sensitivity of the human marker: it was detected in 100% of sewage samples from two WWTP by multiplex qPCR (Caldwell & Levine, 2009).

qPCR methods targeting the cytochrome b gene in mtDNA from humans and eight other vertebrates were developed (Schill & Mathes, 2008). Average amplification efficiency for all markers was 97.6%. The human mtDNA method detected individual human fecal samples and sewage samples with 100% sensitivity. Specificity of the human marker was 100% when tested against fecal samples from domestic animals and wildlife as well as cloned genes from other hosts. The authors also analyzed sewage samples by qPCR for the Bacteroidales microbial markers, that is, AllBac for general fecal contamination (Layton et al., 2006) and both HF183 (Seurinck et al., 2005) and BT (Carson et al., 2005) as HAM, finding all to be 100% sensitive. The LOD for the human mtDNA marker was reported as 1.8 mg human feces per 100 mL, which is more than 100 times lower than the LOD for the multiplex qPCR method (Caldwell & Levine, 2009). The more extensive DNA extraction and concentration method of Schill et al. may account for this variation (Schill & Mathes, 2008). Although no field testing was performed in this study, sixteen freshwater surface samples and groundwater samples were used to test for matrix interference, and no matrix inhibition was detected.

Further evaluation of methods

Many studies have confirmed the high sensitivity and limited specificity of HF183 and related Bacteroidales-like markers, such as the previously mentioned study by Kildare et al. (Kildare et al., 2007). In Canada, the HF183 method was 100% specific and 87% sensitive against effluent from a local wastewater treatment facility (Edge et al., 2010). A study using the BacHum-UCD (Kildare et al., 2007) qPCR method confirmed the method’s utility, but also found some limited cross-reactivity with cow, dog, and horse fecal samples (Table 1) (Silkie & Nelson, 2009). Twenty-five percent of dog fecal samples and 20% of cat fecal samples were positive for the HF183 marker by end-point PCR; however, qPCR was not performed in this study (McGuigg et al., 2009). The HF183 qPCR method (Seurinck et al., 2005) was used in freshwater to quantify sewage contamination in urban lake in Dhaka, Bangladesh (Ahmed et al., 2010). Sensitivity of the marker was 87% and specificity was 93% (detected in one dog and one cat) against samples from cows, dogs, cats, and chickens.

Comparative studies of human-associated Bacteroidales markers

Several studies have compared PCR methods for human-associated Bacteroidales markers. A comparison of qPCR assays using the HF183 and BacHum-UCD markers was
performed on DNA extracts from fecal and environmental water samples from Kenya (Jenkins et al., 2009) (Table 1). While both methods were 100% specific to human sources, the BacHum-UCD (Kildare et al., 2007) method was less sensitive toward human fecal samples and sewage than the HF183 method. Five human-associated Bacteroides PCR methods [HF183 (Bernhard & Field, 2000b), BacHum-UCD (Kildare et al., 2007), HuBac (Layton et al., 2006), BacH (Reischer et al., 2007), and Human-Bac (Okabe et al., 2007)], were tested against sewage and fecal samples in Australia (Ahmed et al., 2009a). All methods were 100% sensitive, but specificity ranged from 99% (HF183) to 63% for HuBac (Table 1).

A comparison of assays for human-associated Bacteroidales evaluated the performance of ten qPCR methods and five end-point PCR methods against sewage and fecal samples from 22 different animal species (Shanks et al., 2010a). The end-point methods tested were the HF183 and HF134 markers (Bernhard & Field, 2000b), BT (Carson et al., 2005), HumM19, and HumM22 (Shanks et al., 2007). The qPCR methods targeted HuBac, HF183, BsteriF1, BunIF2, BfragF1, BthetaF2, BvulgF1, HumM2, HumM3, and PcopriF1 (Shanks et al., 2010a) (Table 1). HF183 and HumM19 markers showed the best performance among end-point methods, with 100% sensitivity toward sewage samples, more than 10% sensitivity at the lowest total DNA concentration, and 95% specificity. All methods had amplification efficiencies between 89% and 99.5%, \( R^2 \) values > 0.96, range of quantification (ROQ) of 10 to at least 4 \( \times 10^4 \) gene copies, \( C_T \) precision of < 3.0% mean coefficient of variation across the ROQ, and over 98% sensitivity toward sewage. However, only the HF183, HumM2, and BsteriF1 assays also showed > 99% sensitivity, high abundance in human waste, and high specificity toward nontarget hosts (91%). HuBac showed the greatest abundance in nonhuman waste samples and was detected in 77% of animal fecal samples.

A comparative study by McLain, et al. of five human-associated Bacteroidales PCR assays found cross-amplification with fish feces (McLain et al., 2009). End-point PCR assays for HuBac (Layton et al., 2006), BacHum-UCD (Kildare et al., 2007), BacH (Reischer et al., 2007), HF183 (Seurinck et al., 2005), and HF134 (Bernhard & Field, 2000b) were used to amplify fecal DNA from humans, Nile tilapia, channel catfish, rainbow trout, and Atlantic salmon. The HuBac method produced a PCR product with all four types of fish feces, the BacH method gave a positive result with all fish except Atlantic salmon, and the BacHum-UCD and HF183 methods generated PCR products only with rainbow trout. The HF134 method did not amplify DNA from any of the fish feces. Cross-amplification of fish fecal DNA was further evaluated using the HuBac TaqMan qPCR method (Layton et al., 2006), and fish DNA amplified with a similar efficiency to human fecal DNA; however, it is important to note that the HuBac qPCR method is among the least specific of the human-associated Bacteroides methods.

**Single laboratory studies evaluating multiple HAM**

The performance of human-associated MST markers other than Bacteroidales has been compared in several studies. \( F^+ \) RNA coliphage using oligonucleotide probes (Beekwilder et al., 1996) and species-specific Bacteroidales detection by end-point PCR were evaluated in environmental water samples, sewage, and fecal samples in France (Gourmelon et al., 2007). \( F^+ \) RNA coliphages were detected in relatively few fecal samples (21%) and were therefore not tested for specificity, but were more frequently detected in the pig manure slurry (60%) and sewage (100%; groups II and III). Both the HF183 and HF134 markers for human-associated Bacteroidales (Bernhard & Field, 2000b) were highly sensitive and specific, although both cross-reacted with one chicken fecal sample. Fourteen of 28 environmental water samples contained no \( F^+ \) coliphage, but all of them contained the general Bacteroidales marker (Bac 32F/708R) (Bernhard & Field, 2000a). Furthermore, the probability of detecting at least one human marker was significantly correlated with \( E. coli \) concentrations (\( P < 0.001 \)), and the same was true of human-associated \( F^+ \) coliphage groups II and III (\( P = 0.015 \)).

End-point PCR methods using the HPyVs marker, human-associated Bacteroides HF183 (Bernhard & Field, 2000b) and the esp gene of Ent. faecium (Scott et al., 2005), were compared with concentrations of cultured FIB (\( E. coli \), enterococci, Clostridium perfringens) as well as human pathogens [adenoviruses, enteroviruses, sapoviruses, and torque teno viruses (TTV)] in Australia (Ahmed et al., 2009b). Sensitivity and specificity for each method were determined using sewage samples and individual fecal samples from cattle, pigs, sheep, dogs, and ducks. All three methods showed 100% sensitivity to sewage samples. HF183 showed cross-reactivity with one dog sample and had a specificity of 98% while esp and HPyVs had specificities of 100%. The LOD was ten gene copies for esp and HPyVs, and one gene copy for HF183.

These markers (HF183, esp, and HPyVs) were further validated against sewage that was serially diluted in several water types (Ahmed et al., 2009b). HF183 could be detected in freshwater, seawater, and distilled water in more dilute samples than culturable FIB (\( E. coli \), enterococci, and \( C. perfringens \)), that is, at dilutions of \( 10^{-7} \) and \( 10^{-8} \). The esp marker was detected in all waters at a \( 10^{-4} \) dilution, corresponding to FIB concentrations of...
End-point PCR methods targeting presumably human-associated Bacteroides spp. BT (Walters et al., 2007), B. vulgatus (BV) (Wang et al., 1994), HF183 (Bernhard & Field, 2000b), Bifidobacterium adolescentis (biADO) (Matsuki et al., 1998), and Bifidobacterium catenulatum/pseudocatentulatum (biCATg) (Matsuki et al., 1998) were tested against human, ruminant, and nonruminant animals (horse, donkey, dog, chicken, goose, and pig fecal samples) in Ireland (Dorai-Raj et al., 2009). Sensitivity was evaluated against both individual human fecal samples and sewage. The BV assay showed the greatest sensitivity toward fecal samples (88%) compared with other assays [BT (65%), HF183 (12%), biADO (85%), and biCATg (46%)]. BV and both Bifidobacterium methods showed 100% sensitivity to sewage, while BT (39%) and HF183 (70%) did not perform as well. HF183 was 100% specific, while the performance of other methods was adequate, but not as strong (86% for BV, 84% for biADO, 87% for biCATg; BT not being tested).

End-point PCR assays using human-associated Bacteroidales HF134 and HF183 (Bernhard & Field, 2000b), the esp gene of Ent. faecium (Scott et al., 2005), ADO and Bifidobacterium dentium (DEN) (Bonjoch et al., 2004), and human mtDNA (Humito) (Martellini et al., 2005), as well as animal-specific markers, were evaluated singly and in combination with their ability to distinguish between human and nonhuman fecal samples (Balleste et al., 2010). Human sewage and cow, poultry, and pig slaughterhouse effluent and farm samples were tested from Spain, France, Sweden, the UK, and Cyprus. Predictive models identified combinations of methods that showed improved source identification of fecal samples compared with individual markers. The best predictive model gave 90.1% correct classification of fecal samples using five markers: ADO, DEN, HF134, Pomito, and Bomito (a pig- and bovine-specific mtDNA marker, respectively).

**Inter-laboratory comparisons**

Due to the paucity of data available on interlaboratory comparison of qPCR methods for MST, studies employing end-point PCR are included in this review. Where qPCR was used, it is specified in the text. A multilaboratory comparison was performed to evaluate the success of F+ colipage typing (Hsu et al., 1995) as well as host-specific Bacteroides markers including the HF183 human marker (Bernhard & Field, 2000b) at determining the source of fecal contamination in blind samples spiked with fecal material (Field et al., 2003; Griffith et al., 2003). Methods targeting adenoviruses and enteroviruses (Noble et al., 2003) were also evaluated as well as several genotypic and phenotypic library-dependent methods (Harwood et al., 2003; Myoda et al., 2003). The HF183 method correctly identified 100% of samples containing human fecal material, while the F+ colipage method correctly identified only 10% of samples containing individual human feces but 90% of samples containing sewage spikes (Griffith et al., 2003). The enterovirus and adenovirus assays also suffered high false-negative rates for individual fecal samples, but were detected in 60% and 40% of sewage-spiked samples, respectively.

Human sewage and cow, poultry, and pig slaughterhouse effluent and farm samples were tested from Spain (n = 132), France (n = 3), Sweden (n = 2), the UK (n = 2), and Cyprus (n = 5) using six human-associated markers including HF134 and HF 183 (Bernhard & Field, 2000b), esp (Scott et al., 2005), ADO and B. dentium (DEN) (Walters & Field, 2009), and human mtDNA (Humito) (Martellini et al., 2005; Balleste et al., 2010). None of the methods showed 100% sensitivity: ADO was most sensitive (96%), followed by Humito (84%), DEN (64%), HF183 (50%), HF134 (30%), and esp (4%). When an enrichment culture step was used for esp analysis [as specified in the original protocol of Scott et al. (Scott et al., 2005), sensitivity increased to 77%. The most specific assay was DEN (92%), followed by HF134 (81%), ADO (74%), HF183 (71%), Humito (41%), and esp (6% and 68% with enrichment). The fact that the HF183 marker showed lower sensitivity and specificity in this study compared to others described in this review reinforces the need to evaluate markers using both sewage and individual fecal samples from different geographic areas. Indeed, the original authors pointed out that HF183 has shown variable performance depending on geographic location.

Three markers [HF183 (Bernhard & Field, 2000b), M. smithii (Ufnar et al., 2006), and HPyVs (McQuaig et al., 2006)] were evaluated using end-point PCR (Harwood et al., 2009). Each method was tested against more than 40 human fecal samples representing individual feces, sewage influent, and septic system samples as well as over 300 animal fecal samples (dogs, cats, cows, birds, and wild animals) collected from three geographic areas surrounding the Gulf of Mexico. All assays were 100% sensitive when tested against 53, 44, and 41 samples, respectively. The HPyVs method was the most specific (100%), followed by M. smithii at 98%, and HF183 at
96%. Each method detected < 20 gene copies as per PCR. Sewage was spiked in environmental water and diluted to determine LOD in an environmental matrix. The HF183 method could detect the target in dilutions as low as $10^{-6}$, corresponding to 33 CFU $\times 100$ mL$^{-1}$ enterococci. The M. smithii and HPyVs methods were generally two orders of magnitude less sensitive, losing their ability to detect the target at dilutions of $10^{-3}$–$10^{-4}$.

A comparative interlaboratory study was conducted by the Southern California Coastal Water Research Project (SCCWRP) using 17 methods to assess the presence of human fecal contamination in waters spiked with known concentrations of sewage or gull guano (Griffith et al., 2009). End-point PCR assays were performed for four human-associated targets including HF183 (Bernhard & Field, 2000b), esp (Scott et al., 2005), M. smithii (Ufnar et al., 2006), and HPyVs (McQuaig et al., 2006) while qPCR assays were performed for two human-associated Bacteroidales markers (Bernhard & Field, 2000b; Kildare et al., 2007), and human-associated M. smithii using previously published primers (Ufnar et al., 2006). Several human pathogenic viruses were also tested, including adenoviruses, enteroviruses, Hepatitis A, and noroviruses. Marker performance was evaluated in blind duplicate environmental water samples that included samples spiked with sewage, gull guano, or nothing (control). Sewage was diluted to give ranges from 50 to $10^{4}$ CFU $\times 100$ mL$^{-1}$ for E. coli and enterococci, while gull guano was diluted to give average FIB concentrations between $>2 \times 10^{5}$ and $>2 \times 10^{5}$ CFU $\times 100$ mL$^{-1}$.

Among the methods based on bacteria, the Ent. faecium esp assay (Scott et al., 2005) was highly sensitive and specific; however, this method requires a 24-h culture step that is not conducive to rapid target detection. PCR and qPCR methods targeting HF183 were adequately sensitive (90% and 65%, respectively) and 100% specific; however, the qPCR method of Kildare, et al. (Kildare et al., 2007) had a high false-positive rate. End-point PCR for M. smithii (Ufnar et al., 2006) had a high false-negative rate (failed to identify sewage-spiked samples) while the qPCR method for this marker had a high false-positive rate (produced positive results for unspiked control samples). The HPyVs end-point assay (McQuaig et al., 2006) was the most accurate among the human viral methods (65% sensitivity and 100% specificity).

Reproducibility of results for duplicate samples was also assessed in the SCCWRP study. The end-point M. smithii method was 100% reproducible, and the HPyVs method was also highly reproducible (94%). Other methods, such as qPCR for M. smithii, were much less reproducible (67%). This study highlights the need to evaluate methods in different matrices, as spiking in different types of ambient water affected method sensitivity in some cases.

A multilaboratory method evaluation study called the ‘Source Identification Protocol Project’ (SIPP) was carried out in 2011 in which sample source was blinded to participating laboratories (Boehm et al., 2013). Results were published in a series of articles (Stewart et al., 2012; Harwood et al., 2013) that focused on various aspects of the study, with an overview article that captured the major findings of the study (Boehm et al., 2013). Forty-one MST methods were tested for performance criteria such as sensitivity, specificity, and LOD by 27 laboratories. Methods evaluated included those targeting human waste and that of a number of other host species or groups including cows, dogs, gulls, pigs, horse, and sheep. Participating laboratories chose the methods that each evaluated and used in-house protocols (i.e. they were not provided standard operating procedures). Each laboratory received 64 blind ‘challenge’ samples that contained either a single fecal source or mixture of two sources (1 : 10 ratio by volume) suspended in artificial freshwater (Boehm et al., 2013). Some of the challenge samples were prepared as 10-fold more dilute suspensions to test method sensitivity. Dilutions of fecal material were chosen with the goal of creating challenge samples containing c. 2000 CFU $\times 100$ mL$^{-1}$ enterococci; however, in practice, enterococci concentrations varied from $4.3 \times 10^{2}$ to $1.7 \times 10^{6}$ CFU $\times 100$ mL$^{-1}$. For most assays, laboratories received membrane filters through which 50 mL of each challenge sample had been passed. Specificity testing was carried out against human feces, septage and sewage, and animal fecal samples (cow, dog, deer, pig, chicken, pigeon, gull, horse, and goose). The results discussed in this review include assays (e.g. HF183, BacH, BacHum, BsteriF1, Btheta, gyrB, HumM2, nifH, BacCow, CowM2, CowM3, BacCan-UCD, Gull-2, and Pig-2-Bac) that have been described in earlier sections or in ‘Animal markers’.

The data from the SIPP study were evaluated in both binary (i.e. all assays were treated as presence/absence) and quantitative format. All qPCR results were normalized to enterococci CFU, and the method was considered sensitive if it provided a median of 50 gene copies as per enterococci CFU (Boehm et al., 2013). Alternative normalization schemes (i.e. DNA mass or copy number of the general marker of fecal pollution, GenBac3) were also considered. A limited number of assays met the minimum specificity and sensitivity requirements of 80% that were arbitrarily set forth by the investigators. When data were treated in binary fashion, these assays included HF183 (end-point assay, SYBR Green), CowM2 and CowM3, BacCan-UCD, and Gull-2 (SYBR Green). When concentration of a given marker in target and nontarget hosts was considered, the number of assays meeting the...
criteria was reduced to HF183 (TaqMan), BacH, and Pig-2-Bac.

When data were considered in binary fashion, reported sensitivity for the HF183 end-point assay was generally higher than 80%, with specificity ranging from 92 to 100%. The SYBR Green version of the HF183 assay had even higher sensitivity (> 87%), but specificity was lower (> 85%; one of the participating laboratories detected the marker in all types of nontarget host feces) (Boehm et al., 2013). Other human-associated assays (BacHum, BsteriF1, HF183-TaqMan, Btheta, BacH, and HumM2) exhibited high sensitivities, but showed lower specificity (e.g. below 80%). The gyrB marker met the specificity criteria, but had lower sensitivity (not detected in some sewage and septage samples) while nifH did not meet either criterion. Among the animal MST markers, CowM2, CowM3, Gull2 (SYBR format), and BacCan, all had sensitivity and specificity in excess of 80%, while Pig-2-Bac showed sufficient sensitivity but inadequate specificity (cross-reacted with dog and human sources).

When data were considered in a quantitative format, among HAM BacH, Btheta, gyrB, HF183 (TaqMan), and HumM2 showed 100% specificity, while BacHum, BsteriF1, HF183 (SYBR Green), and nifH cross-reacted with 7–19% of samples containing feces from nontarget hosts (Boehm et al., 2013). In addition, nifH, Btheta, gyrB, and HumM2 had imperfect sensitivity (i.e. < 80%), while BacH, BacHum, BsteriF1, and HF183 (SYBR Green and TaqMan) all passed sensitivity criteria. Overall, the only two HAM that met both specificity and sensitivity criteria were BacH and HF183 (TaqMan). Among animal MST markers, CowM2, CowM3, and Gull2 (both SYBR and TaqMan) failed to satisfy both sensitivity and specificity requirements. The BacCan assay was deemed sensitive, but it cross-reacted with samples from septage, goose, and cow feces. Pig-2-Bac was the only animal MST marker that satisfied both requirements. Alternative normalization schemes did not change specificity of the assays with the exception of BacCan, which passed the criteria when results were normalized to DNA mass (Boehm et al., 2013). While virus markers tended to be highly specific, they were not sensitive enough under the study conditions to detect fecal contamination in most samples (Harwood et al., 2013).

Survival and persistence studies

The data on survival and persistence of various markers of human fecal contamination have been predominantly generated through mesocosm studies utilizing surface waters (freshwater and marine) inoculated with sewage, human feces, or pure cultures of organisms of interest (Kreader, 1998; Seurinck et al., 2005; Okabe et al., 2007; Bae & Wuertz, 2009; Walters & Field, 2009; Balleste et al., 2010; Dick et al., 2010; Green et al., 2011a, b; Sokolova et al., 2012). The majority of the existing literature has focused on MST methods targeting Bacteroidales and the effect of selected environmental parameters (e.g. sunlight exposure, temperature, salinity, protozoa predation) on their decay.

One of the first such studies investigated persistence of the end-point PCR signal from Bacteroides distasonis at different temperatures and in the presence/absence of eukaryotic predators indigenous to Ohio River water (Kreader, 1998). Persistence in unfiltered river water was inversely related to temperature, where B. distasonis DNA was detectable for up to 14, 5, and 2 days at 4, 14, and 24 ºC, respectively (Kreader, 1998). Absence of eukaryotic predators extended persistence at 24 ºC by at least a week, indicating that grazing by bacterivorous protozoa is also an important mechanism influencing survival (Kreader, 1998).

Balleste, et al. investigated the effect of environmental parameters on survival of B. fragilis, BT, and environmental Bacteroides spp. via culture and molecular techniques (Balleste et al., 2010). The decay of culturable B. fragilis was strongly influenced by the combined effect of elevated temperatures and grazing predators, while culturable BT and environmental Bacteroides spp. were more affected by dissolved oxygen concentrations (Balleste et al., 2010). In general, culturable environmental strains persisted longer than either B. fragilis or BT, but less than culturable fecal coliforms and enterococci (Balleste et al., 2010). At elevated temperatures representative of summer, persistence of qPCR and culturable cells of both B. fragilis and BT was comparable; however, at lower, winter temperatures DNA persisted for significantly longer than the culturable subset of the population (Balleste et al., 2010). A strong effect of temperature was also noted by Seurinck, et al., who found that the HF183 marker persisted for up to 24 days at 4 and 12 ºC and for up to 8 days at 28 ºC in fresh river water when measured by qPCR (Seurinck et al., 2005).

The persistence of the qPCR signal of four different, feces-derived Bacteroides MST markers (total, human, cow and pig) was investigated at varying temperatures (4, 10, 20, and 30 ºC) and salinities (0, 10, 20, and 30 ppt) (Okabe et al., 2007). The response of different markers to changing conditions was similar; persistence was the longest at lower temperatures and higher salinities. This observation was attributed to the lack of or decreased activity of eukaryotic grazers under the less favorable conditions (Okabe et al., 2007). Decay of total and fecal coliforms in the same study differed considerably; concentrations of both groups of organisms remained largely unchanged for the first 4 days in the temperature
trial experiments (Okabe et al., 2007). However, decay of total coliforms, but not fecal coliforms, was considerably enhanced at higher salinities indicating existence of a more resistant subpopulation (Okabe et al., 2007). Greater persistence in marine waters compared to freshwater was noted for several Bacteroides markers (BsterrF1, BuniF2, GenBac3, HF183, HF134, and HumM2) by qPCR (Green et al., 2011a, b). Interestingly, the general fecal pollution marker (GenBac3) persisted longer than the host-associated fraction indicating differential decay profiles between the two groups (Green et al., 2011a, b).

Decay of both culturable and qPCR signals of enterococci was assessed in the same study. While there was no appreciable difference in enterococci decay between the two methods, enterococci in general persisted longer than any Bacteroides qPCR signal (Green et al., 2011a, b).

Walters et al., measured persistence of two human (HF183, HF 184) and two ruminant (CF128 and CF193) Bacteroidales markers by end-point PCR in freshwater mesocosms inoculated with fresh human or cattle feces and incubated at 13 °C (Walters & Field, 2009). In general, cow-associated markers persisted significantly longer than their human counterparts (Walters & Field, 2009). Exposure to light did not affect the decay of human-associated MST markers or CF193, but did increase the decay rate of CF128 (Walters & Field, 2009). Another study also found that sunlight exposure in freshwater matrices had no appreciable effect on decay of human- and ruminant-associated Bacteroides markers [BacH and BacR assessed by qPCR (Reischer et al., 2006, 2007)] (Sokolova et al., 2012).

A similar experimental design was used to compare the effect of light exposure on BacUni-UCD, BacHum-UCD, BacCow-UCD, and BacCan-UCD (qPCR) markers and enterococci (culturable and qPCR) in seawater mesocosms (Bae & Wuertz, 2009). Decay rates of host-associated Bacteroidales markers were remarkably similar, irrespective of the marker type (i.e. target host species) or treatment conditions (i.e. sunlight exposed vs dark), while culturable enterococci decayed significantly more rapidly than the corresponding qPCR signal (Bae & Wuertz, 2009). The general trend of similar persistence in light exposed and dark mesocosms was observed for other human-associated Bacteroides qPCR markers (Green et al., 2011a, b). Dissenting results were reported by Walters et al. (2009) who noted that the time required for 90% decrease in HF183 and BacHum qPCR marker concentrations was almost five times faster in light-exposed treatments compared to dark controls (Walters et al., 2009). Decay of enterococci (culturable and qPCR signal) was assessed in the same study; in general, culturable concentrations decreased significantly faster than the corresponding qPCR signal (Walters et al., 2009). Comparisons between decay rates of Bacteroidales DNA (qPCR) and intact cells (assessed by binding of propidium monoazide) in marine waters indicate that sunlight exposure has a significant effect on persistence of cells (i.e. extended in dark compared to light-exposed treatments), but not on the persistence of DNA (Bae & Wuertz, 2009, 2012).

The decay of qPCR signal for three Bacteroidales markers (AllBac, HF183 and BacHum) was investigated in freshwater mesocosms inoculated with sewage (Dick et al., 2010). The treatment variables in this particular study included exposure to artificial sunlight, temperature (15 and 25 °C), and predation. Reduced temperature and predation resulted in longer persistence of targets, while exposure to sunlight had no significant effect, which is consistent with findings from many other studies (Kreader, 1998; Seurinck et al., 2005; Bae & Wuertz, 2009; Walters et al., 2009; Green et al., 2011a, b; Sokolova et al., 2012). In general, the AllBac marker persisted longer than other targets, indicating the presence of a persistent subpopulation (Dick et al., 2010).

Other studies
The persistence of the B. thetaiotaomicron α-1,6-mannanase marker (B. theta α) through wastewater and septage treatment (Yampara-Iquique et al., 2008) was compared to that of E. coli and enterococci (Srinivasan et al., 2011). BT cell equivalents were greater in sewage compared to E. coli or enterococci, and all microorganisms were found at similar levels in septage. All three microorganisms showed significant differences in concentrations before and after disinfection by chlorine (P < 0.05). However, cultivable E. coli and enterococci showed marked reduction after disinfection, while qPCR signals from all markers exhibited much lower declines in concentration. In sewage, regression analysis of pooled data from untreated and treated samples showed strong correlations between E. coli and B. theta α qPCR markers (r² = 0.87), as well as between enterococci and B. theta α qPCR markers (r² = 0.79). Similarly, septage samples showed strong correlations between B. theta α and E. coli (r² = 0.91) and enterococci qPCR markers (r² = 0.92).

McQuaig, et al. tested persistence of several FIB (fecal coliforms, E. coli, and enterococci), human-associated MST markers (HPyV-qPCR, HF183, and M. smithii – end-point PCR) and pathogens (Adenoviruses-end-point PCR) over 28 days by incubating sewage at two temperatures (25 and 35 °C). Fecal coliforms, E. coli, and enterococci persisted at significantly higher concentrations at 25 °C as compared to 35 °C, while HPyV concentrations showed no significant decrease at either temperature until day 21 (McQuaig et al., 2009). Adenovirus and M. smithii...
PCR signals were detected for the duration of the study at both temperatures, while the HF183 marker persisted for 14 and 28 days at 25 and 35 °C, respectively (McQuaig et al., 2009).

Leskinen, et al. evaluated whether concentration of 10 liter surface water samples by hollow-fiber ultrafiltration (HFUF) would improve sensitivity of detection of three MST markers in samples from an estuarine beach and a stream contaminated by a recent sewage main break in Florida (Leskinen et al., 2010). HPyVs (McQuaig et al., 2009), HF183 (Bernhard & Field, 2000b), and the nifH gene of _M. smithii_ (Ufnar et al., 2006; Harwood et al., 2009) were measured by end-point PCR in concentrated samples and 500 mL of surface water (Leskinen et al., 2010). All three markers were detected following HFUF in a sample from the stream collected 5 days after the sewage spill, even though the _M. smithii_ marker was not detected using the standard filtration procedure. Forty-five days after the spill, the HFUF treatment, but not the standard filtration, allowed detection of the _M. smithii_ and HPyVs markers. Similar results were reported in another study, where the _M. smithii_ marker was detectable for 24 days in surface waters following a sewage leak (Ufnar et al., 2006).

Hamza, et al. investigated stability of HPyV, human adenoviruses (HAdV), pepper mild mottle virus (PMMoV) and TTV spiked into river water at two different temperatures (4 and 25 °C) over 21 days by quantitative reverse-transcriptase PCR. Overall, all targets persisted longer at lower temperature, while PMMoV decreased the least at 25 °C over 21 days (1.1 log_{10}), followed by TTV (3.0 log_{10}), HAdV (3.7 log_{10}), and HPyV (4.2 log_{10}) (Hamza et al., 2011).

To determine the best methods by which to concentrate viruses (JC polyomaviruses, adenoviruses, and noroviruses), (Albinana-Gimenez et al., 2009) evaluated five concentration methods using river water and spiked tap water. Methods included filtration through glass wool with and without prior acidification to a pH of 3.5, ultrafiltration, electropositive filtration, and centrifugation to concentrate viral particles from 10 to 50 L of tap water or volumes of river water ranging from 42 mL to 50 L. Glass wool filtration with preacidification and ultrafiltration showed similar recovery efficiencies for JC polyomavirus and adenovirus, but only the glass wool method allowed for detection of noroviruses in tap water samples (10 L). The glass wool concentration method (without acidification, which was reportedly impossible for large sample volumes) was determined to be the most effective at concentrating viruses from larger volumes (50 L). Recovery efficiencies for all methods and viral targets were low (< 33% recovery). When comparing different volumes of river water, 1 L was determined to be the optimal sample volume, and the authors suggest that using larger volumes also concentrates PCR inhibitors, negating any benefit from larger sample volumes.

**Selected field studies**

Field studies were included in this section if they measured FIB levels or pathogens in conjunction with MST analysis by end-point PCR or qPCR. Studies that only included MST marker detection or quantification were omitted with the rationale that information about correlation with regulatory parameters and/or pathogens advances our understanding of the ultimate usefulness of MST much more than stand-alone measurements of MST markers.

**Human-associated Bacteroidales**

End-point PCR methods for human-associated Bacteroidales have been widely field tested in different matrices and in a variety of geographic regions, for example, end-point PCR assays for Bacteroidales including the human-associated HF134 marker and a general Bacteroidales marker (Bernhard & Field, 2000b) were employed in a study in British Columbia (Jokinen et al., 2010). Water samples were analyzed for fecal coliforms and the bacterial pathogens _Salmonella_, _Campylobacter_, and _E. coli_ 0157:H7. The HF134 and general Bacteroidales targets were found at all four sites, although the general marker was more prevalent. Correlations between the HAM and the concentration of each pathogen were not evaluated; however, the frequency of _Campylobacter_ detection was negatively correlated with the presence of the general _Bacteroidales_ marker as well as with fecal coliform concentrations. _Salmonella_ and _E. coli_ 0157:H7 detection rates were more strongly correlated with fecal coliform levels than with detection of the general _Bacteroidales_ marker.

The human-associated _Bacteroides_ marker HF183 (Bernhard & Field, 2000b) was used in an end-point PCR assay to assess the presence of human fecal contamination in Lake Michigan (Bower et al., 2005). The marker was detected in all WWTP samples collected over a 9-month period from both urban and rural facilities, and the marker was detected in samples with _E. coli_ concentrations ranging from 0.2 to 82 CFU × 100 mL⁻¹. The marker was detected in Lake Michigan up to 9 days following a combined sewer overflow and was consistently detected at sites impacted by a sanitary sewer overflow when _E. coli_ concentrations were > 200 CFU × 100 mL⁻¹. Furthermore, the marker was detected during ten sampling events when _E. coli_ concentrations were less than the regulatory level of 235 CFU × 100 mL⁻¹.
The end-point HF183 assay (Bernhard & Field, 2000b) was also used to characterize fecal pollution reaching an urban freshwater beach on Lake Ontario in Canada (Edge et al., 2010). Suspected impacts included combined sewer outfalls, as well as abundant gulls and Canada geese. The most probable source of fecal contamination, which was based on E. coli levels, the frequency of HF183 detections, and library-dependent MST methods, varied from human sewage near the river to bird droppings at the other end of the beach. HF183 detection was not positively correlated with E. coli levels, but in certain cases was negatively correlated.

Parker, et al. also used end-point PCR for the HF183 marker to assess beach contamination, but this study focused on stormwater runoff impacting a marine beach in North Carolina (Parker et al., 2010). Coastal sites affected by stormwater runoff outfalls were evaluated for HF183, a BT-like marker termed ‘fecal Bacteroides’ (Converse et al., 2009), E. coli, Enterococcus spp., and total coliforms under various flow conditions. The authors proposed an ‘action threshold’ for fecal Bacteroides of 5000 cells X100 mL\(^{-1}\) that would trigger further MST analysis. One site of three tested was frequently positive for HF183 (44%), and these positive results were always accompanied by high levels of fecal Bacteroides. No correlation analysis was shown.

A study conducted in the Great Lakes (Lake Michigan) utilized end-point PCR for HF183 and measurements of E. coli to determine pollution sources to a popular recreational beach (Kinzelman & McLeLlan, 2009). A stormwater outfall and other runoff were identified as the major contributors of E. coli to the beach. Three of five storm event samples from the outfall contained HF183; however, the authors noted that E. coli levels in excess of \(10^8\) MPN-100 mL\(^{-1}\) were associated with detection of HF183. A bovine marker was not detected in any sample. Several improvements were made to the beach including retrofitting of the storm drain. E. coli concentrations in nearshore waters of the beach decreased from a mean of 232 MPN-100 mL\(^{-1}\) (2000–2004) to 76 MPN-100 mL\(^{-1}\) following the remediation actions.

Human-associated Bacteroidales qPCR methods also have been widely field-tested in different geographic areas and matrices. In Japan, the Human-Bac1 qPCR results showed no correlation with FIB in water samples (Okabe et al., 2007). The HF183 qPCR method (Seurinck et al., 2005) was field tested in Hawaii in streams and marine waters with varying levels of sewage impact (Betancourt & Fujisaka, 2006). The marker was detected in environmental waters with enterococci and E. coli concentrations of < 1 and 15 CFU \(\times\) 100 mL\(^{-1}\), respectively. HF183 was detectable for up to 6 days by qPCR following a sewage spill.

The same HF183 qPCR method (Seurinck et al., 2005) was used to assess the source of fecal contamination in the Santa Ana River, CA (Litton et al., 2010), which is heavily influenced by disinfected, tertiary-treated effluent from twelve WWTPs. Total coliform, E. coli and enterococci levels were low within the WWTP and at its discharge site and increased significantly with distance from the discharge. By comparison, sediment concentrations of cultured FIB were relatively constant over the sampling sites. The HF183 marker was detected only in the Santa Ana River, in 60% and 25% of samples from two sites. Based on several lines of evidence, including Enterococcus speciation, the authors concluded that enterococci were naturalized and growing in the river sediment.

The previously mentioned study in Bangladesh (Ahmed et al., 2010) used qPCR for HF183 to analyze water samples from 20 locations along a polluted urban lake. The CF128 ruminant-associated marker and culturable enterococci were also measured. HF183 was found in 70% of lake samples, generally at concentrations only 1–2 orders of magnitude lower than raw sewage or feces, while the ruminant marker was found in 35% of samples, including three of the six samples that did not contain HF183. The very high concentrations of enterococci (\(1.1 \times 10^5\)–1.9 \(\times\) 10^6 CFU \(\times\) 100 mL\(^{-1}\)) combined with a high detection rate for the HF183 marker indicate that the lake suffers from high levels of human fecal pollution.

The HF183 qPCR assay (Seurinck et al., 2005) was used to investigate fecal pollution sources in a coastal shellfish harvesting area of France that experienced high E. coli concentrations (Gourmelon et al., 2010a). Samples from stream and storm drain sites in urban areas and a cattle farming area were evaluated for the HF183 marker, general and ruminant-specific Bacteroidales markers, F\(^+\) RNA coliphage, and cultured E. coli. Positive correlations were found between the general Bacteroidales marker and F\(^+\) coliphage, between the general Bacteroidales marker and E. coli, and between E. coli and F\(^+\) coliphage. HF183 was detected in all water samples from four urban sites, and most of these samples also contained a high percentage (57–100%) of group II F\(^+\) RNA coliphage. Two other urban sites showed lower frequency of detection of both of these two HAM s. The ruminant-associated marker was only detected at two urban sites at low frequency and concentration, but it was detected in 17% of samples taken adjacent to the cattle farming area, while HF183 was not detected in any samples from this location.

A qPCR method for human-associated Bacteroidales was used to investigate the source of fecal pollution in streams draining to a recreational beach in Wales (Wyer et al., 2010). Several methods were used to evaluate changes in water quality during a rainfall event, including hydrology tools, a microbial tracer experiment using a specific
microorganism in each of four streams, several FIB, qPCR for HF183 following the method of Stapleton et al. (Stapleton et al., 2009), as well as general and ruminant-associated Bacteroidales qPCR assays (Layton et al., 2006; Reischer et al., 2007). Tracer microorganism levels released at the start of a rainfall event showed that two streams contributed freshwater rapidly (in < 40 min) to this location. FIB levels were highest (including fecal coliform levels above the European Directive 76/160/EEC imperative level of 2000 CFU × 100 mL−1) when the microbial tracers from these two streams were most abundant. The relative frequency of detection and concentration of general and ruminant MST markers compared to human markers led the authors to conclude that the primary source of beach fecal pollution was cattle from dairy farms.

Another coastal study (Boehm et al., 2009) evaluated water quality and FIB contamination using qPCR for the Bacteroides MST marker BacHum-UCD (Kildare et al., 2007). F+ and somatic coliphages, enteroviruses, and adenoviruses, and several FIB were also measured over a 72-h period in seawater at Avalon Beach in California (Boehm et al., 2009). Adenoviruses were not detected during the study. Coliphages were correlated with FIB, but not with enteroviruses; however, the occurrence of the BacHum-UCD marker was positively correlated with enteroviruses, was negatively correlated with F+ coliphages, and was more frequently detected than F+ coliphages or enteroviruses. The concentration of the BacHum-UCD marker was negatively correlated with E. coli and enterococci concentrations.

Schriewer, et al. utilized the BacHum-UCD qPCR assay (Kildare et al., 2007) to investigate the relationships among Bacteroidales markers (human, general, dog, and cow), FIB, and pathogens in rivers and estuaries of the Monterey Bay region of California (Schriewer et al., 2010). In general, the strongest correlations were found between FIB (e.g. total coliforms and fecal coliforms); however, human Bacteroidales concentrations correlated with total coliform, fecal coliform, and enterococci concentrations. Only one significant relationship was found between Bacteroidales markers and pathogens. Specifically, the general Bacteroidales marker correlated with Cryptosporidium spp. using a predictive qualifier (PQ) parameter, a weighted measure that predicts pathogen detection relative to threshold cutoff levels of FIB or markers. BacHum-UCD did not predict pathogen occurrence, but general Bacteroidales and FIB showed similar PQ performance when compared for their ability to predict Vibrio cholerae, Cryptosporidium spp., and Giardia spp. This field test illustrated a relationship between general Bacteroidales marker detection and certain pathogens at this study location and showed that the HAM was able to identify sites with likely human fecal contamination, though a connection was not made between the human marker and pathogen detection.

The HF183 marker was also used to determine the presence of human fecal contamination at 45 stormwater outfalls over a 4-year period (Sauer et al., 2011). Samples positive for HF183 by end-point PCR were further analyzed by qPCR using the BacHum-UCD assay. The frequency of detection of HF183 at various stormwater outfalls ranged from 11 to 100%. Human-associated Bacteroides gene concentrations (BacHum-UCD) were not correlated with E. coli concentrations or with enterococci measured by culture or qPCR, but were correlated with E. coli measured by qPCR. High levels of enteric viruses were measured at an outfall that also had high concentrations of the BacHum-UCD target.

Multiple targets for markers of human fecal sources

Several MST field studies have utilized multiple assays directed at HAMs to evaluate fecal pollution sources. Abdelzaher, et al. explored the relationships among the HAMs HPyVs and Ent. faecium esp gene, FIB, and several viral, bacterial, and protozoan pathogens at a beach in South Florida (Abdelzaher et al., 2010). FIB levels exceeded recreational water-quality criteria in only one instance in which HPyVs, Giardia, and Cryptosporidium were also detected.

Korajkic, et al. used end-point PCR methods for HPyVs (McQuaig et al., 2006) and the Ent. faecium esp gene (Scott et al., 2005) to evaluate the impact of remediation at a Florida beach (Korajkic et al., 2011). Prior to remediation, fecal coliforms and enterococci exceeded state standards in 58% and 50% of samples, respectively. The HPyVs marker was detected in 17% of impacted beach samples, and the esp marker was detected in 42% of samples. After sewer main repairs and removal and relocation of beach restrooms, a significant reduction in FIB occurred; only 7% of fecal coliforms and 11% of enterococci samples exceeded standards. The frequency of detection of the esp marker also declined to 18% of samples; however, the HPyVs frequency of detection remained the same. Fecal coliform levels and esp marker detection were correlated, but no correlation was observed between FIB and HPyVs. Salmonella spp. and enteroviruses were measured after remediation, and Coxsackie B4 virus was detected at one site on one date. HPyVs were codetected at this site/date, which coincided with a nearby sewage spill. The differences in relationship between FIB and MST markers and the codetection of an enterovirus with HPyVs, but not esp marker, suggest that use of multiple MST markers are helpful in interpreting field studies.
Relationships among HPyVs, HF183, the nifH gene of *M. smithii*, and human adenoviruses were examined at Doheny State Beach and Avalon Beach in California (McQuaig et al., 2012). Although the locations differ in hydrology and potential fecal inputs, enterococci levels exceeded regulatory standards in c. 30% of samples at both sites. The human-associated *Bacteroides* HF183 marker was detected most frequently, and the *M. smithii* marker was detected with low frequency at both locations. The concentration (by qPCR) and frequency (by end-point PCR) of HPyVs were significantly higher at Avalon Beach, while adenovirus was detected only at Doheny Beach. The presence of adenovirus correlated with both HPyVs concentration and HF183 detection at Doheny Beach, which supports the utility of these markers in identifying water that poses a human health risk. HF183 correlated weakly with FIB at both beaches, while *M. smithii* correlated only with FIB at Doheny Beach. Human MST markers and adenovirus were detected routinely in samples that met FIB regulatory standards. The complexity of relationships among the human markers and FIB underscores the importance of using multiple human markers in MST field studies.

Water quality at a Florida marine beach impacted by nonpoint source pollution was studied using several methods for quantification of enterococci and MST markers directed toward human, dog, and gull sources (Shibata et al., 2010). The two human source markers used were BacHum-UCD (Kildare et al., 2007) and HF183 (Bernhard & Field, 2000b). Little correlation was noted among the FIB and MST markers with the exception of a weak relationship between BacHum-UCD and qPCR for enterococci.

Coastal and tidally influenced creek waters in Mississippi were collected over a 2-year period and assessed for the presence of HF183 and the *M. smithii* nifH gene by end-point PCR, as well as enterococci concentrations by membrane filtration (Flood et al., 2011). Neither HAM correlated with enterococci concentrations.

### Correlation with pathogens and risks of waterborne illness

#### Epidemiology

In addition to establishing the performance of MST methods in terms of specificity, sensitivity, and limit of detection in ambient waters, an important next step is to conduct epidemiological studies to determine their ability to predict human health risks due to waterborne pathogens. Numerous epidemiological studies have assessed the relationships among point source or nonpoint source pollution, FIB levels, and human health outcomes, and recent meta-analyses provide extensive reviews of this topic [e.g. (Wade et al., 2003; Zmirou et al., 2003)], and more recent studies have focused on the effect of nonpoint source contamination on human health (Colford et al., 2007; Fleisher et al., 2010; Soller et al., 2010a; Wade et al., 2010). However, because many of the MST methods previously described have been developed recently, few studies to date have investigated the ability of these methods to predict the risk of illness. Table 2 summarizes four epidemiological studies published to date that include MST methods.

The first study was conducted at two freshwater beaches in the U.S. Great Lakes that were impacted by WWTP effluent and combined sewer overflows (CSOs) (Wade et al., 2006). A cohort of beachgoers were surveyed on the sampling day and in follow-up interviews (n = 5667) to determine how rates of gastrointestinal (GI) illness correlated with water-quality analyses, including qPCR for *Enterococcus* (Haugland et al., 2005) and human *Bacteroides* (HF183) (Dick & Field, 2004). Results were reported as adjusted odds ratios (aOR), which represent the increase in risk of illness for each log unit increase of FIB concentration. *Enterococcus* copy numbers were significantly associated with GI illness, and the authors also noted that the association between *Enterococcus* and GI illness strengthened with increased exposure time. In contrast, no overall trend was observed with human *Bacteroides*, although a nearly significant positive association was detected at the Lake Erie beach.

A second epidemiological study took place at multiple beaches within Mission Bay in coastal California (Colford et al., 2007). This study attempted to determine correlations between indicators of water quality and rates of swimming-related illnesses at a site thought to be impacted by primarily nonhuman and nonpoint sources of pollution. A cohort of beachgoers was surveyed on the day of sampling and 2.2 weeks later (n = 8797) to determine rates of a number of illnesses by tracking GI, respiratory, dermatological, and other symptoms. Human-associated *Bacteroides* (HF183) was quantified via qPCR (Bernhard & Field, 2000b; Bernhard et al., 2003), somatic and F+ coli-phages were quantified, and MPN estimates of adenovirus and norovirus concentrations were determined via reverse-transcriptase PCR (Colford et al., 2005). However, of all the measured health outcomes, only skin rash and diarrhea were significantly more frequent in swimmers compared with nonswimmers. For these two symptoms, there were no correlations between rates of illness and FIB levels or rates of detection of human *Bacteroides* or somatic coliphage. Viruses were not detected often enough to determine any relationship.

A third study was conducted (Wade et al., 2008) as a follow-up to previous work (Wade et al., 2006) and was
Table 2. Correlations observed in studies between MST markers and risks for various types of waterborne illnesses

<table>
<thead>
<tr>
<th>MST marker</th>
<th>GI illness</th>
<th>Respiratory illness</th>
<th>Skin infections</th>
<th>Eye infections</th>
<th>Ear infections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenovirus*</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
</tr>
<tr>
<td>Bacteroides</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
</tr>
<tr>
<td>Human CE (qPCR)</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
</tr>
<tr>
<td>Enterococcus CE (qPCR)</td>
<td>(+)²</td>
<td>(+)²</td>
<td>(+)²</td>
<td>(+)²</td>
<td>(+)²</td>
</tr>
<tr>
<td>F+ coliphage*</td>
<td>(+)¹</td>
<td>(+)¹</td>
<td>(+)¹</td>
<td>(+)¹</td>
<td>(+)¹</td>
</tr>
<tr>
<td>Norovirus*</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
</tr>
<tr>
<td>Somatic coliphage*</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
<td>n.d.1</td>
</tr>
<tr>
<td>Staphylococcus aureus</td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
</tr>
</tbody>
</table>

A plus (+) sign indicates that a significant correlation was found; a (−) sign indicates that it was tested, but no significant correlation was found; (n.d.) = no data, which indicates that data were collected, but the results were not published due to methodological difficulties or sampling limitations. References are given for each case.

*Indicates presence/absence test.
1Calford et al. (2007).
2Wade et al. (2006).
3Sinigalliano et al. (2010).
4Wade et al. (2008).

designed to include nonenteric illnesses as well as analysis by age group. This study was also conducted in the Great Lakes region of the United States at beaches that were impacted by effluent from WWTPs and CSOs. Beachgoers were surveyed on the sampling day and in follow-up interviews (total n = 21 015) to determine rates of illness. Enterococcus spp. and human-associated Bacteroides HF183 concentrations were measured via qPCR (Dick & Field, 2004), but low sensitivity associated with the Bacteroides assay prohibited the use of these data. The incidence of GI illness was consistently correlated with Enterococcus concentrations by qPCR, and the risk was significantly higher among children below the age of ten.

The fourth study was conducted at a subtropical marine beach in Florida that is affected by nonpoint sources of pollution (Sinigalliano et al., 2010). This study utilized a prospective randomized exposure design. Adult participants were randomly grouped into bathers and nonbathers (n = 1303) and were surveyed on the day of sampling and 7 days later to track rates of GI tract illnesses, upper respiratory ailments, and skin illnesses. Six qPCR MST methods were employed: two human-associated Bacteroides methods (HF183 and BacHum-UCD), two methods to detect Enterococcus spp. (Haugland et al., 2005; Siefring et al., 2008), one dog-associated Bacteroidales method, and one gull-associated method to detect Catelllicoccus marimalllum (Lu et al., 2008). Enterococci levels determined by membrane filtration were positively associated with skin illness and 24 h of antecedent rainfall; however, no other correlations were noted. According to a separately published epidemiological report for this same study [the BEACHES Study, (Fleisher et al., 2010)], GI illness, skin rash, and respiratory illness were significantly more frequent in bathers compared to nonbathers.

Correlation with pathogens

Some studies have offered additional insight into MST marker effectiveness for water-quality and human health risk assessment by correlating markers and human pathogens (see Table 3 for summary of studies in surface waters). For example, in separate studies conducted in California, two groups (Boehm et al., 2003; Noble et al., 2006) analyzed the presence of human-associated Bacteroides HF183 by end-point PCR (Bernhard & Field, 2000b) and the presence of enteroviruses via quantitative reverse-transcriptase PCR in surface water samples. The study sites were vastly different: Boehm, et al. analyzed marine samples from Catalina Island, where flushing and dilution of pollutants is relatively great, while Noble et al. worked in the urban, highly impacted Ballona Creek watershed. Boehm, et al. found relatively poor agreement between the detection of HF183 and enteroviruses; in contrast, Noble et al. were able to detect HF183 in nearly every sample in which enteroviruses were detected. HF183 was also frequently detected in the absence of enteroviruses in many of the Ballona Creek samples (Noble et al., 2006). In a more recent study at Avalon Beach, however, Boehm, et al. observed a correlation between human-associated...
Table 3. Correlations between detection or concentration of MST markers and the detection of various waterborne pathogens in surface waters

<table>
<thead>
<tr>
<th>MST marker</th>
<th>Campylobacter</th>
<th>Clostridium perfringens</th>
<th>Cryptosporidium</th>
<th>Giardia</th>
<th>Enteric viruses*</th>
<th>E. coli (pathogenic)</th>
<th>Leptospira</th>
<th>Salmonella</th>
<th>Shigella</th>
<th>Staphylococcus aureus</th>
<th>Vibrio cholerae</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteroidales</td>
<td>(+)¹</td>
<td>(-)²</td>
<td></td>
<td>(+)³</td>
<td>(-)¹</td>
<td>(-)¹</td>
<td>(-)¹</td>
<td>(-)¹</td>
<td>(-)¹</td>
<td>(-)¹</td>
<td>(-)¹</td>
</tr>
<tr>
<td>Human (+/-)</td>
<td>(-)¹</td>
<td>(-)²</td>
<td></td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
</tr>
<tr>
<td>Bacteroidales Ruminant (+/-)</td>
<td>(-)²</td>
<td>(-)²</td>
<td></td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
<td>(-)²</td>
</tr>
<tr>
<td>Bacteroidales Human (qPCR)</td>
<td>(-)⁶</td>
<td>(-)⁷</td>
<td>n.d.⁹</td>
<td>(+)¹⁰</td>
<td>(+)¹</td>
<td>(+)¹⁰</td>
<td>(+)¹¹</td>
<td>(-)⁸</td>
<td>(-)⁸</td>
<td>(-)⁸</td>
<td>(-)⁸</td>
</tr>
<tr>
<td>Human polyomavirus Ent. faecium esp gene (qPCR)</td>
<td>(-)¹²</td>
<td></td>
<td></td>
<td>(+; negative)⁷</td>
<td>(-)⁶</td>
<td>(-)⁶</td>
<td>(-)⁶</td>
<td>(-)⁶</td>
<td>(-)⁶</td>
<td>(-)⁶</td>
<td>(-)⁶</td>
</tr>
</tbody>
</table>

A plus (+) sign indicates a significant correlation (designated ‘negative’ for a negative correlation); a (−) sign indicates that no significant correlation was found.

* (+/-) indicates conventional (end-point) PCR.

† Includes adenoviruses.

¹ Walters et al. (2007).
² Fremaux et al. (2009).
³ Noble et al. (2006).
⁴ McQuaig et al. (2012).
⁵ Boehm et al. (2002).
⁶ Schriewer et al. (2010).
⁷ Viau et al. (2011).
⁸ Savichtcheva et al. (2007).
⁹ Jenkins et al. (2009).
¹⁰ Schriewer et al. (2010).
¹² Wong et al. (2009).
Bacteroides BacHum-UCD (Kildare et al., 2007) measured by qPCR and enteroviruses (Boehm et al., 2009).

More recently, qPCR methods have been used to correlate pathogens and MST markers more precisely. In a large study encompassing 22 tropical streams on O‘ahu, Hawai‘i, Viau, et al. (Viau et al., 2011) analyzed water samples collected during dry weather for four types of enteric viruses (adenovirus, enterovirus, norovirus GI, and GI2) and the human Bacteroides marker BacHum-UCD (Kildare et al., 2007). In addition, they obtained historical data from the same samples that were analyzed for C. perfringens, Campylobacter, and Salmonella to include additional pathogen data. BacHum-UCD was positively correlated with the presence of Campylobacter (P = 0.02), but, interestingly, a negative correlation was found between BacHum-UCD and two of the viruses, adenovirus and norovirus GI (P = 0.02 for each). No other significant correlations were found. In a companion study, a positive correlation was found between BacHum-UCD and Leptospira (P = 0.004) in the same samples (Viau & Boehm, 2011).

Savichtcheva, et al. sampled freshwater and WWTPs in Hokkaido, Japan to determine whether the concentration of Bacteroidales, as determined by qPCR (Okabe et al., 2007), correlated with the presence or absence of a variety of human pathogens, including pathogenic strains of E. coli, Salmonella spp., Shigella spp., Clostridium perfringens, Staphylococcus aureus, and V. cholerae (Savichtcheva et al., 2007). In this case, human-associated Bacteroidales markers showed significant correlations with E. coli O157: H7, Salmonella, and enterotoxigenic E. coli. They also concluded that culture-dependent methods for FIB enumeration were inferior predictors of pathogen presence when compared with the MST markers.

Similarly, Walters, et al. analyzed the ability of the general Bacteroidales marker Bac32F (Bernhard & Field, 2000a), ruminant-specific markers CF128 and CF193 (Bernhard & Field, 2000b), human-specific markers HF134 and HF183 (Bernhard & Field, 2000b), and the pig-specific marker PF163 (Dick et al., 2005) to detect the presence of E. coli O157:H7, Salmonella spp., and Campylobacter spp. in water samples from the Oldman River Basin in Canada (Walters et al., 2007). The ruminant-specific markers were particularly good predictors of E. coli O157:H7 presence, and, to a lesser extent, Salmonella. The HAMs showed a strong association with the presence of Campylobacter (OR = 10). These results suggest not only that the markers for Bacteroidales can be successful predictors of the presence of pathogens but that host-specific markers may give insight into which host-specific pathogens are most likely to be present.

In a different approach, a modification of the standard qPCR methods for Enterococcus and Bacteroidales was used that included pretreatment with propidium monoazide prior to DNA extraction to limit detection to only viable organisms (Varma et al., 2009). In this case, reductions in both of these markers were tracked through a WWTP and found to correlate well with reductions in infectious viruses, suggesting their utility as indicators of pathogen presence.

Walters, et al. used seawater microcosms seeded with sewage to further explore the relationship between persistence of MST markers and persistence of pathogens (Walters et al., 2009). They measured persistence of the human-associated Bacteroides markers HF183 by endpoint PCR (Bernhard & Field, 2000b) and Bac-Hum by qPCR (Kildare et al., 2007), Enterococcus spp. via culturable and qPCR methods (Haugland et al., 2005), and infectious enteroviruses using plaque assays and RT-PCR. They found differential survival of Enterococcus spp., the Bacteroidales marker, and enterovirus genetic markers, with Bacteroidales persisting for less time than the other two. However, in the mesocosms exposed to sunlight (as is the case in recreational waters), the Bacteroidales markers had the most similar persistence dynamics to infectious enteroviruses (as determined by the plaque assays), suggesting that Bacteroidales may be the best indicator of health risks associated with enteroviruses, at least under the conditions of their study.

The presence of human polyomaviruses (HPyVs) has also been found to correlate with human pathogens. McQuaig, et al. examined correlations between HPyVs, human-associated Bacteroides HF183 (Bernhard & Field, 2000b), and adenoviruses in marine water at two beaches in California, USA (McQuaig et al., 2012). HF183 detection and HPyVs concentration were significantly correlated with adenovirus detection at Doheny Beach, but not at Avalon Beach. In previous work, McQuaig, et al. analyzed correlations between concentrations of HPyVs via qPCR and the presence of human-associated Bacteroides HF183 and human adenoviruses in a variety of samples from the human waste stream, including individual septic tanks, septic tank pump trucks, lift stations, WWTP influent, and WWTP effluent (McQuaig et al., 2009). They found strong correlations between both MST markers and the detection of adenoviruses in all samples except for the septic tanks. In the individual septic tank samples, both human-associated Bacteroides and HPyVs were detected in the majority of samples, but adenoviruses were never detected. Furthermore, in 28-day laboratory microcosms, the authors concluded that the persistence of adenoviruses in the water was better mimicked by HPyVs than by culture-based bacterial indicators such as E. coli and Enterococcus spp., which both declined rapidly even though adenoviruses were detectable for much longer periods of time.
It is important to note, however, that correlations between MST markers and pathogens have not been found universally. Fremaux, et al. found no correlations between the detection of Bacteroides marker HF183 by end-point PCR and the presence of pathogens including Campylobacter, Salmonella, and Shiga toxin-producing E. coli in Saskatchewan, Canada, using odds ratio analyses (Fremaux et al., 2009). They did, however, find the presence of ruminant-associated Bacteroidales marker (CF128) to be a significant predictor of the presence of Salmonella (OR = 12.2). Similarly, the Ent. faecium esp marker (Scott et al., 2005) was evaluated as a marker for adenovirus, enterovirus, and rotavirus at two Great Lakes beaches (Wong et al., 2009), but no significant correlations were found. Interestingly, however, physical data (wind speed, temperature, pH, turbidity, etc.) were shown to be good predictors of pathogen presence.

As part of a method validation study, Jenkins, et al. compared the presence of Cryptosporidium spp. to detection of the human-associated Bacteroides HF183 and BacHum-UCD markers by qPCR in river samples from Kenya (Jenkins et al., 2009) (see 'Further evaluation of methods'). Cryptosporidium was detected in 78% of water samples, and only three of these were also positive for human Bacteroides by either assay (two with BacHum-UCD and one with HF183). A ruminant-specific marker BacCow-UCD was codetected with Cryptosporidium spp. in 10 of 19 samples. Ruminant and HAMs were codetected with Cryptosporidium spp. in two samples. Overall, the ruminant-associated marker was more closely associated with Cryptosporidium presence than the HAMs.

A recent study by Schriewer, et al. in Monterey Bay, California also found no correlations between detection of the BacHum-UCD Bacteroides marker (Kildare et al., 2007) and the presence of pathogens including Campylobacter spp., Salmonella spp., E. coli 0157:H7, Vibrio spp., Cryptosporidium spp., or Giardia spp (Schriewer et al., 2010). They did, however, find that a general Bacteroidales marker had a comparable or better ability to predict pathogens compared with FIB using a weighted measure called a PQ (see ‘Selected Field Studies’ and Table S1).

**Animal markers**

Human health risk from animal fecal contamination is generally assumed to be less severe than from human sources, in part because of the host-specific nature of viruses; however, waterborne zoonotic infections caused by pathogens shed in domestic/agricultural animal feces, such as Salmonella, E. coli O157:H7, Campylobacter jejuni, Giardia spp., Cryptosporidium spp., and hepatitis E virus still pose a definite health risk (Field & Samadpour, 2007; Soller et al., 2010b). Recently, a quantitative microbial risk assessment (QMRA) study estimated that under some scenarios, cattle feces represent as great a risk to human health as human sewage (Soller et al., 2010b). Although a review of all MST methods directed toward animal sources is beyond the scope of this work, here, studies employing rapid MST methods for animal sources are briefly reviewed and evaluated based on their sensitivity to target fecal contamination, specificity against nontarget contamination, limits of detection and quantification, and field-testing results (Table S2).

**Pig markers**

Two TaqMan methods for pig feces (Pig-1-Bac and Pig-2-Bac) target the 16S rRNA gene of porcine-associated Bacteroidales (Mieszkin et al., 2009). Pig-1-Bac and Pig-2-Bac were highly sensitive (98 and 100%, respectively) and specific (100%). Field testing was performed using river water samples (n = 24) from 14 sites, and markers were detected in 25% (Pig-1-Bac) and 62.5% (Pig-2-Bac) of samples collected around pig farms, but not in areas influenced by cattle grazing or human fecal contamination. Concentrations of these markers ranged from 3.6 to 4.1 log10 target copies x 100 mL^-1. The persistence of the Pig-2 marker was subsequently evaluated in seawater and freshwater microcosms (Solecki et al., 2011) in which the marker persisted between 20 and 27 days, respectively. Meanwhile concentrations of E. coli and enterococci remained detectable by culture-based methods for 55 days in both water types.

A TaqMan method targeting porcine adenovirus (PADV) was evaluated using pig slurries, urban and slaughterhouse sewage, and river water (Hundesa et al., 2009). The method showed 87% sensitivity to pooled pig fecal samples and 100% sensitivity to slaughterhouse sewage. One hundred percent specificity was observed when the PADV assay was tested against both human- and bovine-specific adenovirus serotypes. In addition, urban sewage samples tested negative for this marker. Six river water samples collected downstream of a pig-rearing area were positive, suggesting that this marker may be useful in identification of fecal contamination from swine in environmental waters.

**Dog markers**

A dog marker (BacCan-UCD) and TaqMan qPCR method were developed targeting the 16S rRNA gene from host-associated Bacteroidales (Kildare et al., 2007). This method was 100% sensitive toward dog fecal material and 86% specific, cross-reacting with human feces, WWTP influent, and cat feces. In a field-validation study in which contamination sources were blinded to the
researchers, ten samples spiked with varying amounts of target and nontarget fecal material were tested. Only samples amended with dog feces showed amplification with this marker (Kildare et al., 2007). A subsequent field study in California (Schriewer et al., 2010) suggested that fecal contamination from dogs made up <1% of the total Bacteroidales load in coastal waters.

A second dog marker was developed targeting mtDNA (NADH dehydrogenase subunit 5, ND5) that was reported to be 100% specific against individual fecal samples (Caldwell & Levine, 2009). Sensitivity was not directly assessed, although amplicons were reported to have 100% identity to canine ND5 sequences via NCBI BLAST. The marker was further tested against 24 sewage samples, all of which were positive.

**Cow markers**

Using the previously developed CF128 forward primer (Bernhard & Field, 2006b), a TaqMan qPCR method specific for cow feces was developed (BacCow-UCD) (Kildare et al., 2007). This assay was 100% sensitive to cow fecal material and showed 95% specificity, cross-reacting with horse fecal samples. A field study in which ten water samples were amended with varying concentrations of target and nontarget sewage found that the concentration of BacCow-UCD marker detected was proportional to the amount of cow feces added. BacCow-UCD was also detected in samples amended with horse feces, confirming the cross-reactivity of the assay.

The BoBac TaqMan qPCR assay also targeted the 16S rRNA gene of cattle-associated Bacteroidales (Layton et al., 2006). This assay showed 100% sensitivity to a small number of bovine fecal samples and cross-reacted with only one dog fecal sample (overall specificity of 93%). Field testing was performed in rural and resort areas that were expected to be impacted by cattle and humans, respectively, and the BoBac assay showed that a high proportion of fecal contamination in the rural watershed was attributable to cattle. This assay was also evaluated by Kildare, et al. in conjunction with the BacCow-UCD assay (Kildare et al., 2007) where it performed comparably, showing 100% sensitivity and 97% specificity but cross-reacted with two (25%) human fecal samples.

Two other bovine-specific TaqMan methods were developed based on sequences identified by metagenomic analysis of the bacterial assemblage in cattle feces (Shanks et al., 2006, 2008). CowM2 and CowM3 assays target genes encoding proteins involved in energy metabolism and a secretory protein, respectively. Both CowM2 and CowM3 were ≥98% sensitive and 100% specific (Shanks et al., 2008). The performance of these methods was subsequently compared with that of the BoBac qPCR method (Shanks et al., 2010b). The CowM2 assay was detected in seven of the eleven cattle populations tested, CowM3 was detected in nine, and BoBac was present in every population. Both CowM2 and CowM3 showed 100% specificity to cattle waste, while BoBac was only 47% specific.

**Gull and other bird markers**

The Gull-2 marker, which targets the 16S rRNA gene of *Catellicoccus marimammalium*, found in gull feces, was developed in an end-point PCR and a SYBR Green qPCR format (Lu et al., 2008). These assays showed 71% and 74% sensitivity, respectively, to gull feces collected from around the United States and 100% host specificity when tested against a wide variety of individual and composite nontarget DNA samples (Table S2). Seventy-two water samples were collected from Canada, Ohio, and the Great Lakes, and amplification via the Gull-2 assays was only observed in waters expected to be impacted by gull fecal contamination (Lu et al., 2008). No amplification was observed in waters impacted by other species of waterfowl or other animal sources. The Gull-2 assay was adapted to a TaqMan format, and the marker was evaluated to determine if its presence could predict illness following exposure to subtropical, marine recreational waters; however, no significant relationship was found between marker concentration and human illness (Sinigalliano et al., 2010).

SYBR Green, Gull-3, and TaqMan, Gull-4, assays targeting 16S rRNA gene sequences from *Streptococcus* spp. and *C. marimammalium*, respectively, were developed and tested against the Gull-2 assay (Ryu et al., 2012b). The presence of each marker was evaluated in gull feces ($n = 255$), feces from six nonavian species ($n = 180$), 15 other non-gull avian species, and in environmental water samples (gull-impacted and non-gull-impacted). More than 86% of gull feces tested positive for the Gull-2 and Gull-4, while fewer than 30% of samples gave positive results for Gull-3. All three assays showed low cross-reactivity with nonavian feces (0.6–15%) and low to moderate cross-reactivity with other, non-gull avian feces (13–31%). In waters presumed to be impacted by gull feces ($n = 349$), 86%, 59%, and 91% of samples were positive for Gull-2, Gull-3, and Gull-4, respectively. In contrast, c. 5% (Gull-2 and Gull-4) and 21% (Gull-3) of non-gull-impacted water samples ($n = 239$) tested positive for these markers. Results from this study indicate that Gull-2 and Gull-4 markers may be more conservative indicators of gull fecal pollution (Ryu et al., 2012b).

Subtractive hybridization of gull 16S rRNA genes was used to identify other potentially gull- or avian-specific targets, and two SYBR green qPCR assays were developed...
targeting *C. marinomammalium* and unclassified *Helicobacter* spp. (GPC and GFD, respectively) (Green et al., 2011a, b). Assays were evaluated against DNA from 635 fecal samples representing 31 host species to determine sensitivity and specificity toward avian species and gulls, specifically. GFC showed greater sensitivity and specificity to gulls while GFD broadly targeted bird fecal samples. Assays were further evaluated against environmental water samples, and markers were detected in all waters expected to be impacted by gulls (GFC) or birds (GFD).

Two TaqMan PCR assays targeting the 16S rRNA gene of *Bacteroidales* specific to Canada goose feces have also been developed (CGOF1-Bac and CGOF2-Bac) (Fremaux et al., 2010). Sensitivity and specificity were evaluated using raw sewage samples and fecal samples representing 17 species. Ninety-four percent of Canada goose fecal samples tested positive for the All-Bac maker for detection of general *Bacteroidales* (Layton et al., 2006). Sensitivity of the markers toward individual Canada goose fecal samples positive for All-Bac was 61% and 54% for CGOF1-Bac and CGOF2-Bac, respectively, and 57% and 50% when all samples (including those negative for All-Bac) were considered. CGOF1-Bac cross-reacted with only one nontarget sample (pigeon) among nontarget samples positive for All-Bac while CGOF2-Bac showed no cross-reactivity toward nontarget fecal DNA giving both a specificity of > 99%. LOD for both assays was determined to be < 10 target copies per reaction. These markers were further field tested on freshwater samples collected from eight sites during 3 months. CGOF1-Bac was detected in 87% of samples and CGOF2-Bac was detected in 79%. The highest mean concentrations for both markers were observed at sites known to be frequented by Canada geese, suggesting the effectiveness of these markers at identifying Canada goose fecal contamination in freshwater environments.

A Canada goose marker targeting mtDNA (NADH dehydrogenase subunit 2, ND2) was also developed and showed 100% specificity (Caldwell & Levine, 2009). Similar to the dog marker developed in the same study (see above), sensitivity was evaluated by NCBI Blast analysis and amplicons showed 100% identity to Canada goose ND2 sequences. This marker was also detected in 29% (n = 7) wastewater samples.

In addition to wild bird assays, a SYBR Green qPCR assay for poultry feces and litter (LA35) has also been developed targeting the 16S rRNA gene of *Brevibacterium avium* (Weidhaas et al., 2010). The marker was validated using fecal DNA from poultry litter samples, individual chickens, and nontarget species including cattle, swine, ducks, geese, and humans. The marker was detected in all soiled poultry soils and litter and showed a sensitivity of 61% toward individual chickens. The qPCR assay was 95% specific, yielding false-positive results with goose and sewage samples. LOD was 30 gene copies per reaction. The LA35 assay was further tested via blind sampling at a second laboratory, which tested target (n = 13) and nontarget (n = 32) fecal DNA as well as environmental samples. The second laboratory confirmed 100% sensitivity and specificity and saw amplification from environmental samples only when samples were collected from waters expected to be impacted by poultry litter.

Most recently, a marker to identify fecal pollution from sandhill cranes (Crane1) was developed targeting the 16S rRNA gene of unclassified *Lactobacillales* (Ryu et al., 2012a). The marker showed 63% sensitivity to individual sandhill crane samples and 95% specificity but did not cross-react with nontarget avian samples. Field testing was conducted at sites along the Platte River in the Southwestern United States believed to be impacted by crane fecal contamination, and 88% of 16 samples were positive for the Crane1 marker. No samples collected in the absence of abundant crane populations tested positive. The marker was further tested against 138 tropical water samples not expected to be impacted by cranes, and none of these samples tested positive.

**Other animal markers**

To assess the contribution of fecal contamination from muskrats in Eastern Ontario Canada, a TaqMan qPCR assay specific for this species (MuBa01) was developed targeting the 16S rRNA gene of *Bacteroidales* (Marti et al., 2011). The marker was evaluated against a set of fecal samples collected from wildlife, livestock, and human sources as well as 22 water samples collected from two sites where muskrats were observed. The marker was 66% sensitive toward muskrat fecal samples that tested positive for general (‘universal’) *Bacteroidales* by the BacUni assay (Kildare et al., 2007) (n = 9). No cross-reactivity (100% specificity) was observed toward all nontarget samples tested, and all nontarget samples tested positive via the BacUni assay. MuBa01 was detected in 2/11 water samples from one site and 1/11 from the other suggesting an ability to detect muskrat fecal contamination in environmental waters.

mtDNA qPCR assays targeting NADH dehydrogenase subunit 5 (ND5) and cytochrome b of cat and white-tailed deer, respectively, were developed to characterize domestic and light industrial wastewater influents from two municipal treatment plants in Southwestern North Carolina (Caldwell & Levine, 2009). Markers were tested for cross-reactivity against avian, mammalian, and piscine feces and found to be specific; however, percent specificity and number of nontarget fecal samples were not reported. Each marker was detected in one of 24 sewage...
Conclusions

A plethora of new MST assays targeting human sewage and other animal fecal sources has been developed in recent years, largely because of several challenges that make finding the ‘perfect’ method elusive. These challenges may be roughly divided into the areas encompassing (1) intrinsic method performance, for example, sensitivity and specificity toward sewage and fecal samples, including geographic range, (2) method performance in the field (sensitivity toward dilute samples, effects of PCR inhibition, and the efficiency of DNA recovery from environmental matrices), and (3) knowledge about ecology of the organisms and persistence of the markers in the environment, and correlation with FIB and pathogens in environmental waters.

In our view, the use of multiple methods to identify any one source is valuable because while most of the PCR methods in use have a sensitive analytical limit of detection for the PCR itself, ranging from 1 to 10 gene copies, each one has particular strengths and weaknesses with regard to other issues that can limit the usefulness of MST. For example, the most thoroughly vetted marker, *Bacteroides* HF183, is not completely specific for human waste; however, it has the advantages of being broadly distributed among human populations and of a relatively high concentration in sewage. The more strongly human-associated microorganisms, such as HPyVs and pathogenic viruses, are less concentrated in sewage and are therefore difficult to detect in dilute samples, even at levels where pathogens may be present; however, their specificity engenders great confidence in the finding of human sewage pollution when they are detected. Better concentration methods that allow high-throughput sampling of large volumes with quantitative recovery need to be developed for these viral targets and others with lower concentrations in fecal matter.

While markers exist for domestic animals such as cattle, poultry, horses, pigs, and dogs, because of the focus on human sources, the distribution and performance of animal markers are not as well understood. In particular, markers have not been developed for many domestic animals that are thought to be important contributors to fecal loading in surface waters. Even fewer markers for wild animals have been developed, in spite of the fact that many are implicated as significant sources of fecal pollution to environmental waters.

In addition to developing new markers and better concentration methods for diluted samples, the challenges of implementing PCR and qPCR methods in environmental waters are common to all the methods. Inhibition of the PCR must be routinely assessed, and DNA purification methods that allow quantitative recovery of DNA and relief from inhibition must be developed. Within a geographic area and on a given day, the extent to which PCR inhibition creates methodological problems can be highly variable (Harwood et al., 2011), and this variability is magnified many times over across wider geographic ranges and different water types.

Standardization of methods and units will be important, particularly as certain methods move toward acceptance in the regulatory realm. Relatively few studies have attempted interlaboratory studies of a given method, but those that have (Harwood et al., 2009; Shanks et al., 2012) emphasized the importance of standard operating procedures and training of personnel. Details as basic as how limits of detection and units are expressed make comparison of method performance difficult among studies. Furthermore, the strategy for normalizing reported gene copy numbers is not standardized and can lead to confusion when comparing among studies, that is, ‘Should one normalize to sample volume, to DNA recovered, to some microbial value such as enterococci concentration?’

The promise of MST should not be underestimated, even in light of the challenges outlined above. The process associated with vetting and optimizing MST methods is common to all developing areas of scientific endeavor, that is, one anticipates as many issues as possible and devises strategies to circumvent them, but inevitably, unanticipated hurdles arise that must be overcome to move the field forward. This review has summarized the current state of the science; however, we anticipate that the field will rapidly evolve to surpass the capabilities presented here.

Acknowledgements

The funding for this work was provided by the Water Environment Research Foundation (PATH3C09), the Florida Stormwater Association, and the Florida Department of Environmental Protection.
Disclaimer
The United States Environmental Protection Agency through its Office of Research and Development partially funded and collaborated in the research described here. It has been subjected to Agency review and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References


Lund V (1996) Evaluation of *E. coli* as an indicator for the presence of *Campylobacter jejuni* and *Yersinia enterocolitica*.


Srinivasan S, Aslan A, Xagoraraki I, Alocilja E & Rose JB (2011) Escherichia coli, enterococci, and Bacteroides...
Molecular methods for fecal source determination


Wong M, Kumar L, Jenkins TM, Xagoraraki I, Phanikumar MS & Rose JB (2009) Evaluation of public health risks at recreational beaches in Lake Michigan via detection of...


**Supporting Information**

Additional Supporting Information may be found in the online version of this article:

**Table S1.** Selected field studies employing PCR methods to detect human-associated MST markers and correlation of MST markers with FIB levels.

**Table S2.** Studies employing qPCR methods for the detection of animal-associated fecal source tracking markers.

---

© 2013 Federation of European Microbiological Societies Published by John Wiley & Sons Ltd. All rights reserved