Applied and Environmental Microbiology, Apr. 2007, p. 2405–2415 Vol. 73, No. 8
0099-2240/07/$08.00+0 doi:10.1128/AEM.02473-06

MINIREVIEW

Performance, Design, and Analysis in Microbial Source Tracking Studies▽

Donald M. Stoeckel[1] and Valerie J. Harwood[2]*

U.S. Geological Survey, Ohio Water Science Center, Columbus, Ohio,[1] and Department of Biology, University of South Florida, Tampa, Florida[2]

Microbial source tracking (MST) includes a group of methodologies that are aimed at identifying, and in some cases quantifying, the dominant source(s) of fecal contamination in resource waters, including drinking, ground, recreational, and wildlife habitat waters. MST methods can be grouped into two major types. Library-dependent methods are culture based and rely on isolate-by-isolate typing of bacteria cultured from various fecal sources and from water samples These isolates are matched to their corresponding source categories by direct subtype matching (41, 70) or by statistical means (23, 37, 40, 41, 80, 83, 102). In contrast, library-independent methods frequently are based on sample-level detection of a specific, host-associated genetic marker in a DNA extract by PCR (6, 11, 26, 88). Analyses of certain chemicals associated with sewage, including fecal sterols (29, 30, 47), optical brighteners (29, 30, 68), and host mitochondrial DNA (67), have also been utilized for what can be more broadly termed fecal source tracking; however, in this review we compare the performance of only fecal source tracking studies in which the target(s) is microbial.

Mounting pressure to determine the origin of nonpoint source fecal pollution, as exemplified by the U.S. Environmental Protection Agency's total maximum daily load program, has led to a steady increase in manuscript submissions and grant applications that include MST approaches. At the same time, resource managers concerned with water quality and regulatory pressures struggle with the choice of methodology in the face of requirements for immediate application. Although there has been significant progress in the MST field over the last 10 years, variability among performance measurements and validation approaches in laboratory and field studies has led to a body of literature that is very difficult to interpret, both for scientists and for end users (99).

In this review, we first consider the development and validation of MST methods in a historical context to describe the lessons learned in early studies. Next, uniform performance characteristics are introduced to allow comparison of method performance across MST studies (Tables 1 and 2), and this is followed by a discussion of considerations for field study design and implementation.

DEVELOPMENT OF EVALUATION AND VALIDATION STRATEGIES

Library-dependent MST owes its origins to the studies of *Escherichia coli* population biology (14, 71, 77) and antimicrobial resistance studies conducted 15 to 25 years ago (4, 5, 19, 54, 58). Many of the basic premises in early MST studies, such as the assumption that host-specific genetic or phenotypic characteristics are influenced by selective pressure (79, 102), reflected this understanding of population biology. Validation of isolate-by-isolate (generally library-dependent) and sample-level (generally library-independent) classification methods is based on various performance criteria, some of which are approach specific and some of which are common to all approaches. Many of these validation strategies have been refined as MST has advanced, which has resulted in increasingly rigorous assessments of the accuracy of MST methods (26, 41, 72, 74, 75, 97, 99).

Methods that rely on classification of isolates. Many early MST studies were based on construction of an isolate library, or host origin database, of indicator bacteria from known-source fecal material; thus, the methods are termed library dependent (23, 99). The fecal indicator microorganisms used have included *E. coli*, fecal coliforms, fecal streptococci, enterococci, and coliphages. Indicator organism isolates (bacteria or bacteriophages) were characterized by phenotypic (37, 40, 102) or genotypic (12, 23, 80) methods. In principle, fecal indicators isolated from water samples then could be classified according to subtype (also termed patterns, fingerprints, or operational taxonomic units) or group similarity based on corresponding characteristics of the library isolates. In the special case of F-specific RNA coliphage typing, the "library" is quite simple, consisting of two human-associated types and two non-human-associated types (28, 43). Thus, although based on isolate classification, MST methods based on coliphage typing are generally classified in the "library-independent" category (99).

Many strategies and calculations have been used to evaluate the accuracy of library-based, isolate-by-isolate classification. Discriminant analysis was the most frequently used statistical approach in early studies (80, 102). Evaluation of method accuracy in these studies was restricted to the resubstitution table, in which each library isolate was classified into a source category based on its similarity to the library isolates, while remaining in the library. This procedure, also known as a library self-cross (102), was used to calculate rates of correct classification (RCCs) for each source category (102). The early libraries analyzed by discriminant analysis/resubstitution had

* Corresponding author. Mailing address: Department of Biology, University of South Florida, Tampa, FL 33620. Phone: (813) 974-1524. Fax: (813) 974-3263. E-mail: vharwood@cas.usf.edu.
▽ Published ahead of print on 16 February 2007.

TABLE 1. Performance statistics for tests in which results were based on isolate-by-isolate classification into the various known-source categories[a]

Test[b]	Target	Sample type	Human source		Nonhuman source		Reference(s)
			Proportion of true-positive classifications	*n*	Proportion of true-positive classifications	*n*	
ARA	*E. coli*	Blind isolates	0.24	17	0.83	53	97
			0.27	44	0.86	133	72
ARA	Enterococci	Blind isolates	0.66	44	0.55	55	72
Carbon source utilization	*E. coli*	Blind isolates	0.12	17	0.98	126	97
BOX-PCR	*E. coli*	Blind isolates	0.31	16	0.95	133	97
BOX-PCR, HFERP	*E. coli*	Isolates from reference feces	0.54	210	0.94	1,321	53
REP-PCR	*E. coli*	Blind isolates	0.60	10	0.94	83	97
PFGE	*E. coli*, NotI	Blind isolates	0.67	6	0.91	34	97
Ribotyping	*E. coli*, HindIII	Isolates from reference feces	0.85	84	0.79	317	89
Ribotyping	*E. coli*, HindIII	Blind isolates	0.06	17	0.81	53	72
			0.50	44	0.92	130	97
F+ RNA coliphage	Types I through IV	Isolates from blind samples	0.54	28	0.00	1	75
F+ RNA coliphage	Types I through IV	Isolates from reference feces	0.76–1.00	17–25	0.26–0.87	21–1,088	18, 28, 43, 75, 87
F+ RNA coliphage	Types I through IV	Isolates from wastewater	0.83	403	0.88	297	28
			0.87	6,465	0.91	2,495	87

[a] For simplicity, only data from two-way splits (human or nonhuman) were compiled.
[b] ARA, antibiotic resistance analysis; REP-PCR, repetitive DNA element PCR with repetitive extragenic palindromic primers; BOX-PCR, repetitive DNA element PCR with BOX primers; PFGE, pulsed-field gel electrophoresis; HFERP, horizontal fluorophore-enhanced repetitive extragenic palindromic.

high RCCs that generally were in the range from 80 to 90% (12, 37, 80, 102); however, resubstitution estimates of RCC are now understood to be elevated relative to their predictive capabilities for isolates not in the library (41, 83, 104).

In order to correct for artificially elevated RCCs by resubstitution analysis, jackknife analysis (also known as hold-out analysis and cross-validation in various statistical packages and publications) was used in the next generation of MST studies (41, 42, 72, 104). In jackknife analysis, held-out isolates are compared with the residual library at each iteration; thus, each isolate cannot provide a match for itself, and the evaluation gains statistical independence compared with resubstitution analysis (83, 104). Because the subtypes isolated from a given sample tend to be more similar to each other than to isolates from separate samples (98), jackknife analysis is considered an internal measure of library accuracy and also tends to overestimate library accuracy compared to results obtained with challenge (proficiency) isolates (41, 42, 72, 104). One strategy for avoiding the bias inherent in internal measures of library accuracy is a form of jackknife analysis called the pulled-sample test (104), in which all the isolates from a given sample are held out of the library and treated as unknowns.

The practice of retaining multiple isolates of the same subtype from each feces sample used to build a library can inflate RCCs measured by jackknife analysis, unless precautions are taken (53). Subtypes that are indistinguishable from one another (within the discriminatory limits of the typing method) allow direct matching to their corresponding "sisters" from the same sample and may skew group centroids (meaning that the central tendency of the group becomes unduly weighted toward a particular subtype [10]). For this reason, indistinguishable isolates found in a given fecal or wastewater sample should be culled (42, 53, 99); this process has been termed "decloning" (99). Definition of the pattern similarity that constitutes a distinct subtype should not be done arbitrarily but should be established by determining the precision of the typing method by multiple trials with one or more control strains (49).

The history of evaluation strategies for isolate-by-isolation classification in library-dependent MST described above illustrates refinements that have been made to the process. The high RCCs initially generated by resubstitution analysis were interpreted by many workers to indicate that the methods were capable of accurately predicting the source of isolates from water samples; however, evaluation with challenge (proficiency) samples (see below) disproved this notion. Subsequent methodological improvements associated with construction, evaluation, and validation of libraries have alleviated some of these concerns. Our assessment is that it is premature to categorically abandon library-dependent methods; however, when workers conduct a laboratory validation of this class of methods, whether novel or existing, it is essential to incorporate refinements such as calculating RCCs based on cross-validation of a decloned library rather than the previously accepted method of library self-cross (53, 99). Only with rigorous validation can a library, regardless of size, be considered effective for application.

Sample-level classification. As strategies for construction and analysis of MST libraries evolved, other methods that focused on detection of genetic markers associated with fecal contamination sources emerged. These methods are collectively termed library-independent methods (99), and they typ-

TABLE 2. Performance statistics for tests in which MST methods were tested with reference samples to determine the ability or failure to detect the sole source of fecal contamination

Test[a]	Target	Host category	Sample type	Sensitivity (*n*)[b]	Specificity (*n*)[c]	Reference(s)
Isolate-by-isolate classification						
ARA	*E. coli*	Human	Blind samples	1.00 (7)	0.80 (5)	41
ARA	Fecal coliforms	Human	Blind samples	0.43 (7)	1.00 (5)	41
ARA	Enterococci	Human	Blind samples	0.75 (4)	0.25 (8)	41
ARA	Fecal streptococci	Human	Blind samples	1.00 (4)	0.38 (8)	41
MAR	*E. coli*	Human	Blind samples	1.00 (7)	0.00 (5)	41
Carbon source utilization	*E. coli*	Human	Blind samples	1.00 (7)	0.20 (5)	41
Carbon source utilization	Fecal streptococci	Human	Blind samples	1.00 (4)	0.25 (8)	41
Coliphage, F+ RNA, genotyping or sequencing	Probes and primers described	Human	Blind samples	0.33, 0.67 (3)	0.75, 1.00 (4)	26
Coliphage, F+ RNA, genotyping or sequencing	Probes and primers described	Nonhuman	Blind samples	0.00, 0.00 (4)	0.33, 0.33 (3)	26
Ribotyping (EcoRI/PvuII) (two trials)	*E. coli*	Human	Blind samples	0.88, 1.00 (6, 8)	0.00, 0.50 (1, 4)	74
Ribotyping (PstI)	Enterococci	Human	Blind samples	1.00 (8)	0.00 (4)	74
PFGE (XbaI)	*E. coli*	Human	Blind samples	0.88 (8)	0.50 (4)	74
Box-PCR (three trials)	*E. coli*	Human	Blind samples	1.00 each (8)	0.00–0.50 (4)	74
Marker detection						
Bacteroides thetaiotaomicron PCR	B.thetaF/B.thetaR	Human	Individual feces	0.92 (25)	0.98 (241)	11
Bacteroides thetaiotaomicron PCR	B.thetaF/B.thetaR	Human	Wastewater	1.00 (20)	NR (NR)[d]	11
Bacteroides thetaiotaomicron PCR	Primers, two internal probes described	Human	Individual feces	0.78 (9)	0.76 (71)	57
Bacteroidales PCR (two trials)	HF183F, HF134F/Bac708R	Human	Blind samples	0.70, 1.00 (10, 14)	1.00, 1.00 (6, 7)	26
Bacteroidales PCR	HF183F/Bac708R	Human	Individual feces	0.20–0.85 (7–25)	0.85–1.00 (46–73)	6
Bacteroidales PCR	HF183F/Bac708R	Human	Wastewater	1.00 (41)	1.00 (75)	6, 9, 11, 91
Bacteroidales qPCR	HF183F/reverse primer described	Human	Individual feces	0.86 (7)	1.00 (19)	91
Bacteroidales qPCR	HF183F/reverse primer described	Human	Wastewater	1.00 (4)	NR (NR)	91
Bacteroidales PCR (two trials)	CF128F, CF193F/Bac708R	Ruminants and pseudoruminants	Blind samples	1.00 (7, 9)	0.89, 0.92 (9, 12)	26
Bacteroidales PCR	CF128F/Bac708R	Cattle	Individual feces	1.00 (19)	0.73 (40)	6
Bacteroidales PCR	CF128F/Bac708R	Ruminants and pseudoruminants	Individual feces	0.97, 1.00 (31, 20)	1.00, 1.00 (20, 28)	6, 27
Bacteroidales PCR	CF128F/Bac708R	Ruminants and pseudoruminants	Wastewater	1.00 (75)	0.93 (14)	9
Bacteroidales PCR	CF193F/Bac708R	Cattle	Individual feces	1.00 (19)	0.70 (40)	6
Bacteroidales PCR	CF193F/Bac708R	Ruminants and pseudoruminants	Individual feces	1.00 (31)	1.00 (28)	6
Bacteroidales PCR	DF475F/Bac708R	Dog	Blind samples	0.40 (15)	0.86 (7)	26
Bacteroides fragilis phage	Host strain HSP40	Human/nonhuman	Wastewater	1.00 (36)	0.90 (20)	98
Bacteroides fragilis phage	Host strain HSP40	Human/nonhuman	Fecal samples	0.13 (90)	1.00 (145)	33
Sorbitol-fermenting bifidobacteria	Phenotypic, human bifid sorbitol agar	Human	Wastewater	1.00 (4)	1.00 (2)	17
Bifidobacterium adolescentis PCR	Bi-ADO 1/2	Human	Wastewater	1.00 (22)	0.84 (60)	8
Bifidobacterium adolescentis colony hybridization	S-S-B.ado-0182-a-S-27	Human	Individual feces	0.92 (12)	1.00 (85)	63
Bifidobacterium adolescentis colony hybridization	S-S-B.ado-0182-a-S-27	Human	Wastewater	0.67 (3)	1.00 (3)	63
Enterococcus faecium enrichment, PCR	*esp* marker	Human	Septic systems	0.80 (10)	1.00 (59)	88
Enterococcus faecium enrichment, PCR	*esp* marker	Human	Wastewater	1.00 (55)	1.00 (43)	88
Escherichia coli toxin gene	STIb	Human	Blind samples	0.75 (15)	0.33 (7)	26
Escherichia coli toxin gene	LTIIa	Cattle	Wastewater	0.87 (31)	1.00 (207)	55
Escherichia coli enrichment, PCR	STII or STb	Swine	Wastewater	0.90 (31)	1.00 (217)	56
Escherichia coli enrichment, PCR	STII or STb	Swine	Individual feces	NR (NR)	1.00 (224)	56
Adenovirus, nested PCR	Primers described	Human	Wastewater	0.92 (12)	1.00 (31)	64
Adenovirus, nested PCR	Primers described	Human	Blind samples	0.50 (8)	1.00 (3)	75
Adenovirus, nested PCR	Primers described	Swine	Three to five individual feces	0.74 (23)	1.00 (20)	64
Adenovirus, nested PCR	Primers described	Cattle	Three to five individual feces	0.75 (8)	1.00 (35)	64
Enterovirus, RT-PCR	Primers described	Human	Blind samples	0.38 (8)	1.00 (4)	75
Enterovirus, RT-PCR	Primers described	Cattle	Individual feces	0.76 (95)	0.63 (54)	62
Enterovirus, RT-PCR	Primers described	Cattle, deer	Individual feces	0.63 (145)	0.75 (4)	62
Enterovirus, qRT-PCR	Primers and probe described	Cattle	Individual feces	0.78 (193)	0.42 (100)	51
Teschovirus, qRT-PCR	PTV-1 (329–394)	Swine	Wastewater	1.00 (6)	1.00 (unknown)	52

[a] ARA, antibiotic resistance analysis; MAR, multiple antibiotic resistance; Rep-PCR, repetitive DNA element PCR; PFGE, pulsed-field gel electrophoresis; RT-PCR, reverse transcriptase PCR; qRT-PCR, quantitative reverse transcriptase PCR.

[b] Sensitivity is the ability to detect a source when it is present and is calculated by dividing the number of true-positive results by the number of samples that should contain the target.

[c] Specificity is the ability to detect a source when it is not present and is calculated by dividing the number of true-negative results by the number of samples that should not contain the target.

[d] NR, not reported.

ically are used to classify a sample based on whether it contains or does not contain detectable fecal contamination from a particular source. Among the first efforts in this direction was the report that certain *Bacteroides* spp. were frequently associated with human feces but not with animal feces (57). PCR assays targeting the 16S rRNA genes of host-associated members of the order *Bacteroidales* have been developed so that they target sequences of bacteria specifically associated with ruminants or humans (6, 60), swine, or horses (21). Similarly, toxin genes of enteropathogenic *E. coli* were proposed as markers for cattle and swine feces (55, 56). Human-pathogenic viruses were suggested as useful source identifiers (50, 81), as were human-associated bacteriophages infecting *Bacteroides fragilis* HSP40 (98). Other bacterial targets include sorbitol-fermenting bifidobacteria for human contamination (8, 66, 82), as well as *Rhodococcus coprophilus* for grazing animals (30, 65, 78, 86).

The validation of these library-independent markers has been simpler in concept than the validation of library-dependent methods, since only the sample-level presence or absence of the marker in target and nontarget fecal samples is assessed. Sensitivity and specificity are commonly used parameters for sample-level classification (6, 87, 88). In quantitative terms, the sensitivity of a marker can be defined as the proportion of positive control samples in which the marker is detected (true-positive rate), and the specificity of a marker is 1 minus the proportion of negative-control samples in which the marker is detected (false-positive rate). For example, if an assay for a human fecal source detected a human-associated marker in 95 of 100 human feces samples, its sensitivity would be 0.95. If the assay were tested with 100 nonhuman feces samples and 2 of the samples were positive as determined by the assay, its specificity would be 0.98. Analogous calculations that provide slightly different information are positive predictive value and negative predictive value, which can be used for methods that are based on isolate-by-isolate classification (41, 72), as well as the methods that are based on sample-level classification.

COMPARATIVE EVALUATIONS OF METHOD PERFORMANCE

Refinement of validation criteria has resulted in a reduction in the reported accuracy for library-dependent methods; the lowest (most conservative) estimates of method accuracy have been calculated when methods were challenged with blinded isolates and/or samples collected independently from the library (26, 35, 72, 74, 75, 96, 97). In an attempt to provide an unbiased assessment of performance for the many available MST methods, below we present a comprehensive set of performance criterion values based on published studies.

Trends in accuracy estimates over time. Evaluation of accuracy trends across studies and over time requires normalization to the number of classification categories, because the probability of being correct by chance alone increases as the number of classification categories decreases. For example, in the study of Stoeckel et al. (97), the average rate of correct classification (ARCC) for eight categories was 27%, whereas the two-category ARCC was 55%. Although the two-way split appears to have a better ARCC, the ARCC actually was not significantly greater than would be expected for random classification to source (50%), while the ARCC for the eight-way split was twice that expected by chance alone (12.5%). This inconsistency could be addressed by comparing only studies in which the same level of discrimination was attempted, but in many reports only the results of high- or low-level discrimination are presented (23, 72). In order to circumvent this difficulty, we propose a calculation termed the benefit over random: the difference between the observed RCC and the RCC that would be expected from classification by chance alone. For the example above, the benefit over random for the eight-way classification is 14.5% (27% − 12.5%) and for the two-way classification it is only 5% (55% − 50%), even though the correct classification rate appears to be higher for the two-way split.

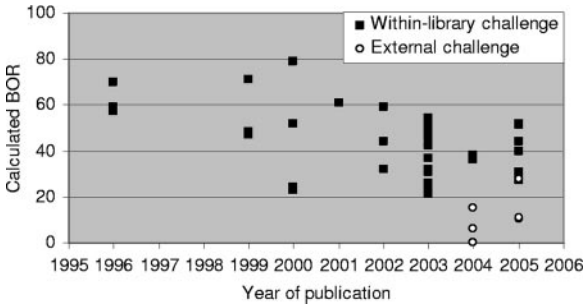

FIG. 1. Benefit over random (BOR) calculated based on ARCCs reported in 21 MST studies published between 1996 and 2005 (1, 12, 13, 16, 23, 34, 36, 37, 40–42, 53, 72, 80, 89, 93, 97, 100, 102–104). Isolates were classified into between two and six source categories in the various studies. Only results for the lowest classification split were included in the analysis.

The benefit over random was used to compare ARCCs calculated in 21 studies published between 1996 and 2005 (1, 12, 13, 16, 23, 34, 36, 37, 40–42, 53, 72, 80, 89, 93, 97, 100, 102–104). The general decrease in accuracy for library-dependent MST methods over the past decade is illustrated in Fig. 1. The apparent decrease in accuracy can be explained by refinement of internal validation measures over time (see above). The benefit over random tends to be much lower when the accuracy is estimated based on external validation, such as validation with challenge samples that are blinded to the analytical laboratory and collected independent of the library, and demonstrates that internal correct classification rates are artificially elevated.

Head-to-head comparison studies. Head-to-head comparisons have been done both on an isolate-level basis and with spiked samples. Only the studies in which blinded (to the analyst) isolates or samples were used are discussed in this section. General observations about performance across the multiple methods are presented here; the reader is referred to the original reports for specific comparisons between methods and to the next section for method-specific performance statistics.

In a study conducted by the U.S. Geological Survey, with multiple collaborators, library-dependent methods employing *E. coli* libraries were compared (97). In this study, investigators received identical sets of *E. coli* isolates for library construction and internal validation. A collection of challenge isolates was submitted double blinded to each participant. In this method

comparison study 13 to 90% accuracy of source identification was found when classification at the level of host species was attempted. A trade-off was evident, in that methods with low accuracy attempted to identify 100% of the isolates, while the method with the highest accuracy attempted to identify only 5% of the isolates. Thus, depending on the approach used, the reported data would include either a large percentage of incorrectly classified isolates or a large percentage of unclassified isolates.

Three aspects of the investigation probably contributed to the unexpectedly poor performance of the methods evaluated: the library was small (each host species was represented by 70 to 200 isolates, some of which were indistinguishable); the challenge isolates were collected in a different time frame than the library isolates (the library isolates were collected in the fall, and challenge isolates were collected in the spring); and the fundamental assumption that *E. coli* subtypes were host specific was not upheld. One of these three reasons, library size, is most commonly addressed in attempts to enhance classification accuracy (104). However, insufficient library size alone does not explain the high error rate observed in library-dependent methods. In one study (72), 27 to 28% of blinded challenge isolates were correctly classified by use of a library of *E. coli* isolates (116 to 159 isolates per host species) whose size was comparable to the size of the U.S. Geological Survey library. Increasing the size of the library (to 380 to 690 isolates per host species) did not increase the classification accuracy. In certain studies of the population biology of *E. coli* (3, 31, 39, 77) workers have identified and discussed the so-called "cosmopolitan" (99) or non-host-specific distribution of many *E. coli* strains, which is probably a major factor that contributes to errors in library-dependent methods.

So far, a series of articles arising from a multi-investigator study led by the Southern California Coastal Water Research Project (SCCWRP) and the U.S. Environmental Protection Agency (26, 35, 41, 74, 75, 83, 96) comprise the only published head-to-head comparison of library-dependent and library-independent methods. In the SCCWRP study, identical collections of reference fecal material were provided to investigators to build libraries and/or optimize their methods. Test isolates and seeded water samples derived from fecal reference material, whose source was blinded to the investigators, were analyzed by 12 methods. This study provided conceptually simple method validation, in that the blinded "challenge" materials (isolates and seeded water samples) were derived from the same samples used to build libraries and for method optimization. Each of the methods had certain strengths and weaknesses (35); in general, library-independent methods had higher false-negative rates for feces from individuals (75) because of patchy distribution in the host population, while the library-dependent methods tended to have higher false-positive rates (a source was identified when it was not actually present) (41). Furthermore, the library-dependent methods correctly identified the dominant source of fecal contamination in about 75% of the seeded samples, but the percentage of either *E. coli* or enterococci attributed to each fecal source varied considerably from the actual seeding load (S. Weisberg, personal communication).

Library-dependent methods also have been compared using antibiotic resistance analysis of *E. coli* and enterococci and ribotyping of *E. coli* (68). The libraries, which were not decloned, had high to adequate rates of correct classification, but they were generally unable to identify the sources of proficiency isolates submitted in single-blind fashion to the analysts. The proficiency isolates were derived from fecal samples collected in parallel with library fecal samples, but they were not represented in the libraries. The library composed of *Enterococcus* antibiotic resistance patterns performed better than the *E. coli* libraries, as the correct classification rate for proficiency isolates in four source categories (45%) was equivalent to the library's correct classification rate (48%). None of the methods was deemed accurate enough for determination of fecal sources in a field setting (68).

Method-specific performance statistics. In other reviews of MST workers have described library-dependent and library-independent methods (24, 25, 69, 90, 92, 94, 95, 99). A U.S. Environmental Protection Agency guide document (99) and a book (85) have recently been published; however, direct comparison of method performance was not a focus of these previous publications, due in part to the different measurements and terms of validation employed. Below, we present a comprehensive review of performance evaluations based on consistent measurements and terms (Tables 1 and 2). For consistency and clarity of comparison, data for only two-way splits are presented. A small percentage of studies could not be included in the tables (e.g., the studies of Dick et al. [21] and Wiggins et al. [104]), because data necessary for our calculations were not available in the publications. For most studies, however, sufficient data were gleaned from the results to calculate either the proportion of isolates correctly classified into human or nonhuman source categories (Table 1) or the sensitivity and specificity for detection of various host-associated markers in feces or contaminated water samples (Table 2). Note that although the term sensitivity often is used to describe the detection limit of a method, in Table 2 the statistical definition of sensitivity is used (sensitivity calculated by determining the percentage of true-positive detections among all positive samples).

As described above, refinement of validation over time for library-dependent methods means that many early reports of accuracy would be considered inflated according to current validation standards. Therefore, data in Tables 1 and 2 were compiled from a subset of reports in which the researchers measured accuracy by procedures consistent with current (2006) procedures, and by necessity the results of several groundbreaking MST reports were excluded. Thus, only external tests of accuracy (e.g., challenge samples collected independent of the library) were compiled from each of these reports. Data for only human versus nonhuman classifications are presented in Table 1, and these classifications are comparable to the host-versus-nonhost classifications that are presented in Table 2.

In general, the accuracy with which isolates are classified in source categories (Table 1) varies widely depending on the method, the library size and representativeness (for library-dependent methods), the decision about whether to reject uncertain classifications, and the analysis laboratory. The various classification strategies employed (e.g., maximum similarity and discriminant analysis) can contribute to the observed variability in method performance (59, 83). Therefore, validation

of the performance characteristics in the time and space for each study is critical for evaluation of results regardless of whether the analytical method has performed well in other studies.

The data presented in Table 2 are compiled results for method performance for proficiency samples that consisted of contaminated water samples or fecal samples. Because some of the library-independent markers had nonhuman targets, sensitivity and specificity measurements were compiled in terms of host versus nonhost instead of human versus nonhuman (as in Table 1). With few exceptions, the host-associated markers evaluated were detectable in nearly all proficiency samples that represented fecal material from many individuals, but their distribution was more patchy among individuals. In general, the markers rarely were detected in nonhost sources (specificity, >0.75). The sensitivity and specificity varied greatly depending on the geographic location and research lab, highlighting the need for local validation of method performance.

The type of proficiency sample chosen for a study affects the ability to compare method performance with the performance in other studies. In most library-dependent studies the validation material was isolates; in the SCCWRP method comparison study, however, the performance validation samples were contaminated water samples (35). For this reason, data obtained with library-dependent methods in the SCCWRP study are presented in Table 2, which is dominated by library-independent methods, rather than in Table 1, which is dominated by library-dependent methods. Most isolate-based proficiency evaluations indicated that few isolates are generally correctly classified for human sources, whereas many isolates are generally correctly classified for nonhuman sources (Table 1). However, because of high false-positive error rates commonly encountered in library-dependent methods, in the SCCWRP study the sensitivity (the ability to detect a source when it is present) for human fecal contamination was very high, but the specificity (the ability to rule out a source when it is absent) for human fecal contamination was relatively low (Table 2).

Although useful for comparisons of method performance, the proportion of true-positive classifications (on an isolate basis) or the sensitivity and specificity of detection (on a sample basis) are not the only important performance measures for MST methods. Other performance measures that should be considered when an MST method is selected for application include the efficiency of recovery of the target from environmental waters, the analytical detection limit of the test, the reproducibility of analytical results, and the suitability of marker detection (or quantification) to meet study-specific objectives. These considerations have been reviewed previously (99).

CONSIDERATIONS FOR FIELD STUDY DESIGN

A primary goal in most applications of MST methods to field settings is to provide data that allow reliable interpretation of sources of fecal contamination in bodies of water (inland or coastal recreational water or drinking source water) that do not meet designated use criteria based on indicators of fecal contamination. In this section, we provide specific information for MST researchers, data users, water resource managers, and journal referees and editors about credible data and the reliability of interpretation in MST field studies. The information presented covers both criteria for a successful study design, according to the current state of our knowledge, and validation steps that should enhance the credibility of interpretations.

Validation in space and time. The performance of any given MST method can vary with geographic area, time, and analytical laboratory (Tables 1 and 2). Therefore, it is critical to validate the accuracy of isolate-by-isolate classification and/or the specificity and sensitivity of host-associated marker detection in samples in the time and space of each study.

(i) Isolate-by-isolate classifications. The library used for classification of test isolates must be validated by use of challenge isolates collected from animals in the watershed during the investigation. Various studies have demonstrated that small known-source libraries (i.e., fewer than 2,000 isolates) collected in an area have higher accuracy in the local area than when they are applied in other areas (39, 104). Local isolates from known sources may improve the classification accuracy and/or the percentage of isolates classified in a source category for both small libraries and so-called "superlibraries" (96). In a study of surface waters in Virginia (46), 559 distinct ribotypes of *E. coli* isolated from known sources were added to a superlibrary (containing approximately 50,000 ribotypes obtained from across the United States). Of the 559 ribotypes collected locally, nearly 30% were not represented in the superlibrary and were not classified in a host group because the analysis laboratory required an exact match. The ability to classify *E. coli* isolated from local water bodies was improved from 52 to 65% by adding a small proportion (about 1%) of local isolates to the library; however, the accuracy of the classifications was not tested.

Proficiency isolates should be collected at the same time that isolates of unknown origin are collected from water samples. The extent to which characteristics used for MST classification in library-dependent methods shift over time is not well understood (96, 99). One study (104) indicated that the accuracy of classification by use of antibiotic resistance analysis was stable for at least 1 year. However, in a study in which challenge isolates were collected in a different season (spring) than the known-source isolates (fall), the classification accuracy for seven protocols generally was low (97). Temporal variability remains one of the possible factors contributing to errors observed in studies.

(ii) Sample-level classification. The use of host-associated markers to classify the sources of fecal contamination at the sample level (compared to isolate-level classification) has recently emerged as a widespread strategy for MST studies; therefore, the geographic distribution of such markers has not been comprehensively described. In some cases, however, distribution can be implied. For example, the various studies in which *Bacteroidales* markers HF183 and CF128 were used suggest that there is wide distribution within the target host populations in North America and Europe (Table 2). Similarly, preliminary data for the *Enterococcus faecium* human-associated *esp* marker indicate that the host distribution in North America is wide, as the marker has been detected in municipal wastewater from at least 19 states in the United States and over 200 samples in Canada (T. Scott, personal communication, 2006). Because the distributions of most markers have not been thoroughly characterized, the expectation that host-asso-

ciated markers are present in a specific host population and not present in nonhost populations should be confirmed by collecting and testing reference feces in the time and space relevant to the study.

The distribution of some host-associated markers within the host population may be affected by seasonal or other ecological factors. The distribution of library-independent markers, particularly those related to virulence factors (detection of host-pathogenic viruses and *E. coli* toxin genes), may vary over short and long time scales, as shown by seasonal patterns of occurrence for some pathogens (20, 45, 61). Furthermore, the relationships between the survival rates of indicator bacteria and the survival rates of pathogens may vary by season (44). Longer-term trends in marker prevalence in host populations have not been described yet.

Detection limit. Detection limit is a critical measure for interpretation of MST results. The detection limit can indicate either the minimum detectable percentage for credible interpretation of the presence of a fecal source (for isolate-level classification) or the threshold level of fecal contamination above which host-associated markers can be detected (for sample-level classification).

(i) Isolate-by-isolate classification. Library-dependent isolate classification methods frequently generate false-positive results in challenge samples (41, 72); for example, in two method comparison studies (35, 97), many library-dependent methods detected many or all potential sources in proficiency samples that were contaminated with only one to three sources. In field applications, it is therefore difficult to determine whether a given source category should be considered a contributor when a low percentage of isolates are classified in that category. One way to control for low-frequency false-positive results is to implement a detection threshold (41, 74, 100, 104). Whitlock et al. (100) proposed that the observed frequency of isolate misclassification could be used to set the minimum level at which a given source would be considered detected. This calculation was termed the minimum detectable percentage by Wiggins et al. (104) and was determined to be 25% in a large (6,587 isolates) library of enterococci which were subtyped by antibiotic resistance analysis and classified into three source categories. A minimum detectable percentage of 15% was applied to the water samples analyzed in the SCCWRP MST method comparison study. As a result, the percentage of false-positive results for seeded water samples was greatly reduced for most phenotypic, library-based methods (41).

Isolate-by-isolate classification by means of coliphage typing generally is used to determine whether human source fecal contamination is present in a sample, not the relative proportions of human contamination and nonhuman contamination (99). This method might prove to be useful in coliphage-based analyses to use the specificity of each type (as in Table 1) as a minimum detectable threshold for considering a source present in a sample.

(ii) Sample-level classification. Sample-level classification methods include both cultivation-based methods (enrichment cultures for *E. coli* and *E. faecium* virulence factors) and cultivation-independent methods. A crucial consideration for both methods is that a threshold level of fecal contamination is acceptable for regulatory purposes; therefore, a detection limit generally should be reported in terms of the regulatory fecal indicator concentration(s). Such a calculation has been carried out for assays of *Bacteroidales* markers from human sources using the PCR (9, 91), showing that the PCR assays are as sensitive as or more sensitive than conventional, culture-based assays for fecal indicator bacteria. For enrichment culture-based tests, the detection limit is determined in terms of the incidence of the marker in the cultivated population of fecal indicator bacteria. In other words, the detection limit indicates the number of target cells (such as *E. coli* or enterococci) from the host source that must be screened in order to have a reasonable chance of detecting the marker. Such calculations have been done in some cases; for example, the *esp* gene was consistently detected when at least 100 *Enterococcus* colonies from sewage were tested, suggesting that there was ~1% frequency among human origin enterococci (88). The detection rates for various host-associated toxin genes in *E. coli* ranged from 1 in 4 cells to 1 in 185 cells (15).

Quantification. As noted above, the ability of any MST method to quantitatively determine the relative contributions of fecal contamination in a water sample has not been convincingly demonstrated yet. Despite this fact, researchers continue to report quantitative results for MST methods (70). Indeed, because total maximum daily load assessments require allocation of fecal contamination loads among potential sources, it seems likely that quantitative assessments of fecal contamination in water samples will continue to be requested by resource managers. One method of providing convincing evidence of quantification would be to correctly approximate the proportional contribution of fecal contamination from multiple sources to a blinded spiked sample, as was attempted in the SCCWRP study (35). The proportional contribution could be calculated on the basis of fecal indicator bacterium concentrations for each source or on the basis of the mass (dry weight) of feces from each source.

(i) Isolate-by-isolate classification. Library-dependent methods offer the potential for quantification of source contributions; however, many confounding factors must be considered, including differential survival of indicator bacterial subtypes in secondary habitats such as water (2, 32) and the high error rates observed for classification of validation isolates and samples (26, 41, 72, 74, 75, 97).

(ii) Sample-level classification. Most library-independent methods based on PCR currently lack quantitative capability, although some quantitative PCR (qPCR) assays have been developed that eventually may allow proportional source allocation. These qPCR assays include assays for total and human-associated *Bacteroidales* (22, 91), *Rhodococcus coprophilus* (86), human adenoviruses and enteroviruses (75), bovine enteroviruses (51), and porcine teschoviruses (52). The use of a series of qPCR assays for every sample in a study may well be impractical; however, the use of a decision tree to determine which analyses would be most beneficial at each site may be useful (99).

Experimental design. In addition to the considerations discussed above (validation of classification accuracy, measurement of detection limits, and quantifying contributions from various sources), a researcher also must determine how to collect representative samples, make choices about the number

and composition of source categories to detect, and decide whether to apply tiered and toolbox approaches.

(i) Representative sampling. A necessary first step in study design is to ensure representative sampling of bodies of water. Fecal contamination is not distributed evenly in receiving water, and care must be taken to represent the cross-section of flowing streams, longitudinal variability in the body of water, and short-term temporal variability that might result from intercepting a contamination slug. The complexity of these issues makes a prescriptive recommendation impossible; for each study area and time frame, diurnal patterns of contamination events, transport pathways, and growth or inactivation of organisms in the environment should be considered in the sample design. The spatial and temporal variability of indicator bacterium concentrations at swimming beaches has been described, and *E. coli* concentrations were observed to be as much as three times higher in the morning than later in the day (101). Indicator bacterium concentration data also can be generated from sediment resuspension (although on the longer temporal scale the bacteria generally are thought to originate from fecal contamination) (48, 73). On a seasonal scale, pathways of contamination can vary with climate or hydrology, and as a result, different sources of fecal contamination can be implicated by MST depending on the season (46, 100) and hydrology (38).

(ii) Number and choice of source categories. Discrimination requirements for MST methods are dictated by study objectives, which are often aligned with pollution control strategies. These requirements in turn influence the choice of methods used, because various MST methods are applicable to a limited range of sources (e.g., human or nonhuman sources; human, domestic, or wildlife sources). The goals of the study must be weighed carefully in order to determine whether a broad split (e.g., human versus nonhuman) is informative, whether a pet-livestock-human split is sufficient, or whether classification to the host species level is necessary (99). Despite the political realities and management level convenience of source definitions such as "wildlife" and "livestock," for some markers it may not be scientifically defensible to use such artificial definitions. For example, the specificity of *Bacteroidales*-based marker CF193 extends across the livestock-wildlife boundary to include all ruminants (6). As previously noted, while library-dependent methods may attempt to detect contamination from many different sources, this level of discrimination is generally not defensible in light of the demonstrated accuracy of the method.

(iii) Use of "toolbox" and "tiered" approaches. Each of the MST methods described in this review has the potential to provide faulty data and generate misleading interpretations. The consequences of incorrect source identification can be substantial; e.g., it may lead to expensive infrastructure improvements that do not improve environmental water quality. Because of these considerations, a current theme in MST recommendations is to use a "toolbox" approach, in which multiple methods allow more accurate interpretations of the data (68, 96, 99). An obvious disadvantage of using multiple methods is increased expense. In two recent studies (7, 76) the workers have instead utilized a tiered approach, in which problem areas were first identified by use of fecal indicator bacterium distribution patterns, putative sources were identified by use of MST markers, and then additional MST marker data were collected to clarify ambiguous results.

CONCLUDING STATEMENTS

The positive results of previous MST studies (12, 23, 37, 40, 80, 102) have led many scientists and water resource managers to believe that existing methodologies for tracking fecal contamination to its source are accurate, reliable, and readily deployed in field investigations. On the other hand, the publication of recent method comparison studies has resulted in recommendations for validation strategies with greater rigor and more applicability to the ultimate use of MST tools (84, 96, 99). The failure of many current MST studies to include adequate validation has led to exaggerated expectations for method performance and has hindered the transfer of MST methods from the research realm to the applied realm. The description of MST validation strategies in this review should allow end users and evaluators (e.g., water resource managers, reviewers, and editors) to more effectively gauge the credibility of MST studies. Validation and improvement of MST methods are ongoing processes. In particular, the decrease in reported accuracy observed in recent studies (Fig. 1) reflects improvements in MST validation practices, including library decloning, construction of larger libraries representing more diverse host populations, and realistic tests of method accuracy, such as tests that use independently collected proficiency isolates and samples. Although end users are eager for recommendations on the comparative accuracy of MST methods, the fact is that the field has not yet reached the state where any one method can be discarded or universally recommended.

Application of MST methods to field settings and interpretation of data should ideally be undertaken only when a number of requirements have been satisfied. (i) Libraries and markers should be validated in the time and space of the study by using proficiency (preferably blind) samples that are independent of calibration data sets. (ii) The minimum detectable percentage for library-dependent methods and other potentially quantitative methods must be defined to minimize false-positive results, and method detection limits for library-independent methods must be calibrated using a relevant regulatory standard, such as indicator bacterium concentrations. (iii) Quantitative library-independent methods (such as qPCR) must be tested for contamination during processing, DNA recovery efficiency, and PCR amplification efficiency in order to relate detected concentrations to original concentrations in environmental samples. (iv) Field application requires not only validation of the accuracy of a method but also collection of representative samples, measurement of seasonal and spatial variability, and collection of sufficient information to allow knowledge of host sources to be translated into specific sources and flow paths of contamination.

Although the limitations identified in method comparison studies may have dampened the promise of early MST studies, the past decade of research has resulted in advances in the field, providing a solid foundation for further improvements. The use of common performance measures and validation strategies in the many studies that will be conducted over the next decade should facilitate rapid progress in this area, as we continue to work toward availability of reliable analyses, clas-

sification approaches, and interpretation strategies for tracking fecal contamination to its sources by use of MST tools.

ACKNOWLEDGMENTS

Among the many sources of information drawn upon for this work, Steve Weisberg, Troy Scott, and Linda Dick provided invaluable assistance. V.J.H. thanks George Simmons, Bruce Wiggins, and Charles Hagedorn for their stimulating discussions, assistance, and collaborative spirit.

D.M.S. received support from the U.S. Geological Survey Office of Water Quality.

REFERENCES

1. **Albert, J. M., J. Munakata-Marr, L. Tenorio, and R. L. Siegrist.** 2003. Statistical evaluation of bacterial source tracking data obtained by rep-PCR DNA fingerprinting of *Escherichia coli*. Environ. Sci. Technol. **37:**4554–4560.
2. **Anderson, K. L., J. E. Whitlock, and V. J. Harwood.** 2005. Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. Appl. Environ. Microbiol. **71:**3041–3048.
3. **Anderson, M. A., J. E. Whitlock, and V. J. Harwood.** 2006. Diversity and distribution of *Escherichia coli* genotypes and antibiotic resistance phenotypes in feces of humans, cattle, and horses. Appl. Environ. Microbiol. **72:**6914–6922.
4. **Bell, J. B., G. E. Elliott, and D. W. Smith.** 1983. Influence of sewage treatment and urbanization on selection of multiple resistance in fecal coliform populations. Appl. Environ. Microbiol. **46:**227–232.
5. **Bell, R. B.** 1978. Antibiotic resistance patterns of fecal coliforms isolated from domestic sewage before and after treatment in an aerobic lagoon. Can. J. Microbiol. **24:**886–888.
6. **Bernhard, A. E., and K. G. Field.** 2000. A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Appl. Environ. Microbiol. **66:**4571–4574.
7. **Boehm, A. B., J. A. Fuhrman, R. D. Mrše, and S. B. Grant.** 2003. Tiered approach for identification of a human fecal pollution source at a recreational beach: case study at Avalon Bay, Catalina Island, California. Environ. Sci. Technol. **37:**673–680.
8. **Bonjoch, X., E. Balleste, and A. R. Blanch.** 2004. Multiplex PCR with 16S rRNA gene-targeted primers of *Bifidobacterium* spp. to identify sources of fecal pollution. Appl. Environ. Microbiol. **70:**3171–3175.
9. **Bower, P. A., C. O. Scopel, E. T. Jensen, M. M. Depas, and S. L. McLellan.** 2005. Detection of genetic markers of fecal indicator bacteria in Lake Michigan and determination of their relationship to *Escherichia coli* densities using standard microbiological methods. Appl. Environ. Microbiol. **71:**8305–8313.
10. **Brown, M. T., and L. R. Wicker.** 2000. Discriminant analysis, p. 209–234. *In* H. Tinsley and S. Brown (ed.), Handbook of applied multivariate statistics and mathematical modeling. Academic Press, San Diego, CA.
11. **Carson, C. A., J. M. Christiansen, H. Yampara-Iquise, V. W. Benson, C. Baffaut, J. V. Davis, R. R. Broz, W. B. Kurtz, W. M. Rogers, and W. H. Fales.** 2005. Specificity of a *Bacteroides thetaiotaomicron* marker for human feces. Appl. Environ. Microbiol. **71:**4945–4949.
12. **Carson, C. A., B. L. Shear, M. R. Ellersieck, and A. Asfaw.** 2001. Identification of fecal *Escherichia coli* from humans and animals by ribotyping. Appl. Environ. Microbiol. **67:**1503–1507.
13. **Carson, C. A., B. L. Shear, M. R. Ellersieck, and J. D. Schnell.** 2003. Comparison of ribotyping and repetitive extragenic palindromic-PCR for identification of fecal *Escherichia coli* from humans and animals. Appl. Environ. Microbiol. **69:**1836–1839.
14. **Caugant, D. A., B. R. Levin, and R. K. Selander.** 1984. Distribution of multilocus genotypes of *Escherichia coli* within and between host families. J. Hyg. **92:**377–384.
15. **Chern, E. C., Y. L. Tsai, and B. H. Olson.** 2004. Occurrence of genes associated with enterotoxigenic and enterohemorrhagic *Escherichia coli* in agricultural waste lagoons. Appl. Environ. Microbiol. **70:**356–362.
16. **Choi, S., W. Chu, J. Brown, S. J. Becker, V. J. Harwood, and S. C. Jiang.** 2003. Application of enterococci antibiotic resistance patterns for contamination source identification at Huntington Beach, California. Mar. Pollut. Bull. **46:**748–755.
17. **Cimenti, M., N. Biswas, J. K. Bewtra, and A. Hubberstey.** 2005. Evaluation of microbial indicators for the determination of bacterial groundwater contamination sources. Water Air Soil Pollut. **168:**157–169.
18. **Cole, D., S. C. Long, and M. D. Sobsey.** 2003. Evaluation of F+ RNA and DNA coliphages as source-specific indicators of fecal contamination in surface waters. Appl. Environ. Microbiol. **69:**6507–6514.
19. **Cooke, M. D.** 1976. Antibiotic resistance among coliform and fecal coliform bacteria isolated from sewage, seawater, and marine shellfish. Antimicrob. Agents Chemother. **9:**879–884.
20. **Cukor, G., and N. R. Blacklow.** 1984. Human viral gastroenteritis. Microbiol. Rev. **48:**157–179.
21. **Dick, L. K., A. E. Bernhard, T. J. Brodeur, J. W. Santo Domingo, J. M. Simpson, S. P. Walters, and K. G. Field.** 2005. Host distributions of uncultivated fecal *Bacteroidales* bacteria reveal genetic markers for fecal source identification. Appl. Environ. Microbiol. **71:**3184–3191.
22. **Dick, L. K., and K. G. Field.** 2004. Rapid estimation of numbers of fecal *Bacteroidetes* by use of a quantitative PCR assay for 16S rRNA genes. Appl. Environ. Microbiol. **70:**5695–5697.
23. **Dombek, P. E., L. K. Johnson, S. T. Zimmerley, and M. J. Sadowsky.** 2000. Use of repetitive DNA sequences and the PCR to differentiate *Escherichia coli* isolates from human and animal sources. Appl. Environ. Microbiol. **66:**2572–2577.
24. **Field, K. G.** 2004. Faecal source identification, p. 349–366. *In* J. A. Cotruvo, A. Dufour, G. Reese, J. Bartram, R. Carr, D. O. Cliver, G. F. Craun, R. Fayer, and V. P. J. Gannon (ed.), Waterborne zoonoses: identification, causes, and control. IWA Publishing, London, England.
25. **Field, K. G., and T. J. Brodeur.** 2003. Molecular approaches to microbiological monitoring: fecal source detection. Environ. Monit. Assess. **81:**313–326.
26. **Field, K. G., E. C. Chern, L. K. Dick, J. Fuhrmann, J. Griffith, P. Holden, M. G. LaMontagne, J. Le, B. Olson, and M. T. Simonich.** 2003. A comparative study of culture-independent, library-independent genotypic methods of fecal source tracking. J. Water Health **1:**181–194.
27. **Fogarty, L. R., and M. A. Voytek.** 2005. Comparison of *Bacteroides-Prevotella* 16S rRNA genetic markers for fecal samples from different animal species. Appl. Environ. Microbiol. **71:**5999–6007.
28. **Furuse, K., A. Ando, S. Osawa, and I. Watanabe.** 1981. Distribution of ribonucleic acid coliphages in raw sewage from treatment plants in Japan. Appl. Environ. Microbiol. **41:**1139–1143.
29. **Gilpin, B., T. James, F. Nourozi, D. Saunders, P. Scholes, and M. Savill.** 2003. The use of chemical and molecular microbial indicators for faecal source identification. Water Sci. Technol. **47:**39–43.
30. **Gilpin, B. J., J. E. Gregor, and M. G. Savill.** 2002. Identification of the source of faecal pollution in contaminated rivers. Water Sci. Technol. **46:**9–15.
31. **Gordon, D. M.** 2001. Geographical structure and host specificity in bacteria and the implications for tracing the source of coliform contamination. Microbiology **147:**1079–1085.
32. **Gordon, D. M., S. Bauer, and J. R. Johnson.** 2002. The genetic structure of *Escherichia coli* populations in primary and secondary habitats. Microbiology **148:**1513–1522.
33. **Grabow, W. O. K., T. E. Neubrech, C. S. Holtzhausen, and J. Jofre.** 1995. *Bacteroides fragilis* and *Escherichia coli* bacteriophages: excretion by humans and animals. Water Sci. Technol. **31:**223–230.
34. **Graves, A. K., C. Hagedorn, A. Teetor, M. Mahal, A. M. Booth, and R. B. Reneau, Jr.** 2002. Antibiotic resistance profiles to determine sources of fecal contamination in a rural Virginia watershed. J. Environ. Qual. **31:**1300–1308.
35. **Griffith, J. F., S. B. Weisberg, and C. D. McGee.** 2003. Evaluation of microbial source tracking methods using mixed fecal sources in aqueous test samples. J. Water Health **1:**141–151.
36. **Guan, S., R. Xu, S. Chen, J. Odumeru, and C. Gyles.** 2002. Development of a procedure for discriminating among *Escherichia coli* isolates from animal and human sources. Appl. Environ. Microbiol. **68:**2690–2698.
37. **Hagedorn, C., S. L. Robinson, J. R. Filtz, S. M. Grubbs, T. A. Angier, and R. B. Reneau, Jr.** 1999. Determining sources of fecal pollution in a rural Virginia watershed with antibiotic resistance patterns in fecal streptococci. Appl. Environ. Microbiol. **65:**5522–5531.
38. **Hartel, P. G., E. A. Frick, A. L. Funk, J. L. Hill, J. D. Summer, and M. B. Gregory.** 2004. Sharing of ribotype patterns of *Escherichia coli* isolates during baseflow and stormflow conditions. Scientific Investigations Report 2004-5004. U.S. Geological Survey, Reston, VA.
39. **Hartel, P. G., J. D. Summer, J. L. Hill, J. V. Collins, J. A. Entry, and W. I. Segars.** 2002. Geographic variability of *Escherichia coli* ribotypes from animals in Idaho and Georgia. J. Environ. Qual. **31:**1273–1278.
40. **Harwood, V. J., J. Whitlock, and V. Withington.** 2000. Classification of antibiotic resistance patterns of indicator bacteria by discriminant analysis: use in predicting the source of fecal contamination in subtropical waters. Appl. Environ. Microbiol. **66:**3698–3704.
41. **Harwood, V. J., B. Wiggins, C. Hagedorn, R. D. Ellender, J. Gooch, J. Kern, M. Samadpour, A. C. H. Chapman, B. J. Robinson, and B. C. Thompson.** 2003. Phenotypic library-based microbial source tracking methods: efficacy in the California collaborative study. J. Water Health **1:**153–166.
42. **Hassan, W. M., S. Y. Wang, and R. D. Ellender.** 2005. Methods to increase fidelity of repetitive extragenic palindromic PCR fingerprint-based bacterial source tracking efforts. Appl. Environ. Microbiol. **71:**512–518.
43. **Hsu, F. C., Y. S. Shieh, J. van Duin, M. J. Beekwilder, and M. D. Sobsey.** 1995. Genotyping male-specific RNA coliphages by hybridization with oligonucleotide probes. Appl. Environ. Microbiol. **61:**3960–3966.
44. **Hutchison, M. L., L. D. Walters, S. M. Avery, and A. Moore.** 2005. Decline of zoonotic agents in livestock waste and bedding heaps. J. Appl. Microbiol. **99:**354–362.
45. **Hutchison, M. L., L. D. Walters, S. M. Avery, F. Munro, and A. Moore.**

2005. Analyses of livestock production, waste storage, and pathogen levels and prevalences in farm manures. Appl. Environ. Microbiol. **71:**1231–1236.
46. **Hyer, K. E., and D. L. Moyer.** 2004. Enhancing fecal coliform total maximum daily load models through bacterial source tracking. J. Am. Water Resour. Assoc. **40:**1511–1526.
47. **Isobe, K. O., M. Tarao, M. P. Zakaria, N. H. Chiem, Y. Minh Le, and H. Takada.** 2002. Quantitative application of fecal sterols using gas chromatography-mass spectrometry to investigate fecal pollution in tropical waters: western Malaysia and Mekong Delta, Vietnam. Environ. Sci. Technol. **36:**4497–4507.
48. **Jamieson, R. C., D. M. Joy, H. Lee, R. Kostaschuk, and R. J. Gordon.** 2005. Resuspension of sediment-associated *Escherichia coli* in a natural stream. J. Environ. Qual. **34:**581–589.
49. **Jenkins, M. B., P. G. Hartel, T. J. Olexa, and J. A. Stuedemann.** 2003. Putative temporal variability of *Escherichia coli* ribotypes from yearling steers. J. Environ. Qual. **32:**305–309.
50. **Jiang, S., R. Noble, and W. Chu.** 2001. Human adenoviruses and coliphages in urban runoff-impacted coastal waters of Southern California. Appl. Environ. Microbiol. **67:**179–184.
51. **Jimenez-Clavero, M. A., E. Escribano-Romero, C. Mansilla, N. Gomez, L. Cordoba, N. Roblas, F. Ponz, V. Ley, and J. C. Saiz.** 2005. Survey of bovine enterovirus in biological and environmental samples by a highly sensitive real-time reverse transcription-PCR. Appl. Environ. Microbiol. **71:**3536–3543.
52. **Jimenez-Clavero, M. A., C. Fernandez, J. A. Ortiz, J. Pro, G. Carbonell, J. V. Tarazona, N. Roblas, and V. Ley.** 2003. Teschoviruses as indicators of porcine fecal contamination of surface water. Appl. Environ. Microbiol. **69:**6311–6315.
53. **Johnson, L. K., M. B. Brown, E. A. Carruthers, J. A. Ferguson, P. E. Dombek, and M. J. Sadowsky.** 2004. Sample size, library composition, and genotypic diversity among natural populations of *Escherichia coli* from different animals influence accuracy of determining sources of fecal pollution. Appl. Environ. Microbiol. **70:**4478–4485.
54. **Kaspar, C. W., J. L. Burgess, I. T. Knight, and R. R. Colwell.** 1990. Antibiotic resistance indexing of *Escherichia coli* to identify sources of fecal contamination in water. Can. J. Microbiol. **36:**891–894.
55. **Khatib, L. A., Y. L. Tsai, and B. H. Olson.** 2002. A biomarker for the identification of cattle fecal pollution in water using the LTIIa toxin gene from enterotoxigenic *Escherichia coli*. Appl. Microbiol. Biotechnol. **59:**97–104.
56. **Khatib, L. A., Y. L. Tsai, and B. H. Olson.** 2003. A biomarker for the identification of swine fecal pollution in water, using the STII toxin gene from enterotoxigenic *Escherichia coli*. Appl. Microbiol. Biotechnol. **63:**231–238.
57. **Kreader, C. A.** 1995. Design and evaluation of *Bacteroides* DNA probes for the specific detection of human fecal pollution. Appl. Environ. Microbiol. **61:**1171–1179.
58. **Krumperman, P. H.** 1983. Multiple antibiotic resistance indexing of *Escherichia coli* to identify high-risk sources of fecal contamination of foods. Appl. Environ. Microbiol. **46:**165–170.
59. **Lasalde, C., R. Rodriguez, and G. A. Toranzos.** 2005. Statistical analyses: possible reasons for unreliability of source tracking efforts. Appl. Environ. Microbiol. **71:**4690–4695.
60. **Layton, A., L. McKay, D. Williams, V. Garrett, R. Gentry, and G. Sayler.** 2006. Development of *Bacteroides* 16S rRNA gene TaqMan-based real-time PCR assays for estimation of total, human, and bovine fecal pollution in water. Appl. Environ. Microbiol. **72:**4214–4224.
61. **Learmonth, J. J., G. Ionas, K. A. Ebbett, and E. S. Kwan.** 2004. Genetic characterization and transmission cycles of *Cryptosporidium* species isolated from humans in New Zealand. Appl. Environ. Microbiol. **70:**3973–3978.
62. **Ley, V., J. Higgins, and R. Fayer.** 2002. Bovine enteroviruses as indicators of fecal contamination. Appl. Environ. Microbiol. **68:**3455–3461.
63. **Lynch, P. A., B. J. Gilpin, L. W. Sinton, and M. G. Savill.** 2002. The detection of *Bifidobacterium adolescentis* by colony hybridization as an indicator of human faecal pollution. J. Appl. Microbiol. **92:**526–533.
64. **Maluquer de Motes, C., P. Clemente-Casares, A. Hundesa, M. Martin, and R. Girones.** 2004. Detection of bovine and porcine adenoviruses for tracing the source of fecal contamination. Appl. Environ. Microbiol. **70:**1448–1454.
65. **Mara, D. D., and J. I. Oragui.** 1981. Occurrence of *Rhodococcus coprophilus* and associated actinomycetes in feces, sewage, and freshwater. Appl. Environ. Microbiol. **42:**1037–1042.
66. **Mara, D. D., and J. I. Oragui.** 1983. Sorbitol-fermenting bifidobacteria as specific indicators of human faecal pollution. J. Appl. Bacteriol. **55:**349–357.
67. **Martellini, A., P. Payment, and R. Villemur.** 2005. Use of eukaryotic mitochondrial DNA to differentiate human, bovine, porcine and ovine sources in fecally contaminated surface water. Water Res. **39:**541–548.
68. **McDonald, J. L., P. G. Hartel, L. C. Gentit, C. N. Belcher, K. W. Gates, K. Rodgers, J. A. Fisher, K. A. Smith, and K. A. Payne.** 2006. Identifying sources of fecal contamination inexpensively with targeted sampling and bacterial source tracking. J. Environ. Qual. **35:**889–897.
69. **Meays, C. L., K. Broersma, R. Nordin, and A. Mazumder.** 2004. Source tracking fecal bacteria in water: a critical review of current methods. J. Environ. Manag. **73:**71–79.
70. **Meays, C. L., K. Broersma, R. Nordin, A. Mazumder, and M. Samadpour.** 2006. Spatial and annual variability in concentrations and sources of *Escherichia coli* in multiple watersheds. Environ. Sci. Technol. **40:**5289–5296.
71. **Milkman, R.** 1973. Electrophoretic variation in *Escherichia coli* from natural sources. Science **182:**1024–1026.
72. **Moore, D. F., V. J. Harwood, D. M. Ferguson, J. Lukasik, P. Hannah, M. Getrich, and M. Brownell.** 2005. Evaluation of antibiotic resistance analysis and ribotyping for identification of faecal pollution sources in an urban watershed. J. Appl. Microbiol. **99:**618–628.
73. **Muirhead, R. W., R. J. Davies-Colley, A. M. Donnison, and J. W. Nagels.** 2004. Faecal bacteria yields in artificial flood events: quantifying in-stream stores. Water Res. **38:**1215–1224.
74. **Myoda, S. P., C. A. Carson, J. J. Fuhrmann, B.-K. Hahm, P. G. Hartel, H. Yampara-Iquise, L. Johnson, R. L. Kuntz, C. H. Nakatsu, M. J. Sadowsky, and M. Samadpour.** 2003. Comparison of genotypic-based microbial source tracking methods requiring a host origin database. J. Water Health **1:**167–180.
75. **Noble, R. T., S. M. Allen, A. D. Blackwood, W. Chu, S. C. Jiang, G. L. Lovelace, M. D. Sobsey, J. R. Stewart, and D. A. Wait.** 2003. Use of viral pathogens and indicators to differentiate between human and non-human fecal contamination in a microbial source tracking comparison study. J. Water Health **1:**195–207.
76. **Noble, R. T., J. F. Griffith, A. D. Blackwood, J. A. Fuhrman, J. B. Gregory, X. Hernandez, X. Liang, A. A. Bera, and K. Schiff.** 2006. Multitiered approach using quantitative PCR to track sources of fecal pollution affecting Santa Monica Bay, California. Appl. Environ. Microbiol. **72:**1604–1612.
77. **Ochman, H., T. S. Whittam, D. A. Caugant, and R. K. Selander.** 1983. Enzyme polymorphism and genetic population structure in *Escherichia coli* and *Shigella*. J. Gen. Microbiol. **129:**2715–2726.
78. **Oragui, J. I., and D. D. Mara.** 1983. Investigation of the survival characteristics of *Rhodococcus coprophilus* and certain fecal indicator bacteria. Appl. Environ. Microbiol. **46:**356–360.
79. **Parveen, S., R. L. Murphree, L. Edmiston, C. W. Kaspar, K. M. Portier, and M. L. Tamplin.** 1997. Association of multiple-antibiotic-resistance profiles with point and nonpoint sources of *Escherichia coli* in Apalachicola Bay. Appl. Environ. Microbiol. **63:**2607–2612.
80. **Parveen, S., K. M. Portier, K. Robinson, L. Edmiston, and M. L. Tamplin.** 1999. Discriminant analysis of ribotype profiles of *Escherichia coli* for differentiating human and nonhuman sources of fecal pollution. Appl. Environ. Microbiol. **65:**3142–3147.
81. **Pina, S., M. Puig, F. Lucena, J. Jofre, and R. Girones.** 1998. Viral pollution in the environment and in shellfish: human adenovirus detection by PCR as an index of human viruses. Appl. Environ. Microbiol. **64:**3376–3382.
82. **Rhodes, M. W., and H. Kator.** 1999. Sorbitol-fermenting bifidobacteria as indicators of diffuse human faecal pollution in estuarine watersheds. J. Appl. Microbiol. **87:**528–535.
83. **Ritter, K. J., E. Carruthers, C. A. Carson, R. D. Ellender, V. J. Harwood, K. Kingsley, C. Nakatsu, M. Sadowsky, B. Shear, B. West, J. E. Whitlock, B. A. Wiggins, and J. D. Wilbur.** 2003. Assessment of statistical methods used in library-based approaches to microbial source tracking. J. Water Health **1:**209–223.
84. **Rochelle, P.** 2006. Workshop on source tracking of pathogens. Water Intelligence Online **5**:200608WF03HHE3.
85. **Santo Domingo, J. W., and M. J. Sadowsky.** 2007. Microbial source tracking. ASM Press, Washington, DC.
86. **Savill, M. G., S. R. Murray, P. Scholes, E. W. Maas, R. E. McCormick, E. B. Moore, and B. J. Gilpin.** 2001. Application of polymerase chain reaction (PCR) and TaqMan PCR techniques to the detection and identification of *Rhodococcus coprophilus* in faecal samples. J. Microbiol. Methods **47:**355–368.
87. **Schaper, M., J. Jofre, M. Uys, and W. O. Grabow.** 2002. Distribution of genotypes of F-specific RNA bacteriophages in human and non-human sources of faecal pollution in South Africa and Spain. J. Appl. Microbiol. **92:**657–667.
88. **Scott, T. M., T. M. Jenkins, J. Lukasik, and J. B. Rose.** 2005. Potential use of a host associated molecular marker in *Enterococcus faecium* as an index of human fecal pollution. Environ. Sci. Technol. **39:**283–287.
89. **Scott, T. M., S. Parveen, K. M. Portier, J. B. Rose, M. L. Tamplin, S. R. Farrah, A. Koo, and J. Lukasik.** 2003. Geographical variation in ribotype profiles of *Escherichia coli* isolates from humans, swine, poultry, beef, and dairy cattle in Florida. Appl. Environ. Microbiol. **69:**1089–1092.
90. **Scott, T. M., J. B. Rose, T. M. Jenkins, S. R. Farrah, and J. Lukasik.** 2002. Microbial source tracking: current methodology and future directions. Appl. Environ. Microbiol. **68:**5796–5803.
91. **Seurinck, S., T. Defoirdt, W. Verstraete, and S. D. Siciliano.** 2005. Detection and quantification of the human-specific HF183 *Bacteroides* 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater. Environ. Microbiol. **7:**249–259.
92. **Seurinck, S., W. Verstraete, and S. D. Siciliano.** 2005. Microbial source

tracking for identification of fecal pollution. Rev. Environ. Sci. Biotechnol. **4:**19–37.
93. **Seurinck, S., W. Verstraete, and S. D. Siciliano.** 2003. Use of 16S-23S rRNA intergenic spacer region PCR and repetitive extragenic palindromic PCR analyses of *Escherichia coli* isolates to identify nonpoint fecal sources. Appl. Environ. Microbiol. **69:**4942–4950.
94. **Simpson, J. M., J. W. Santo Domingo, and D. J. Reasoner.** 2002. Microbial source tracking: state of the science. Environ. Sci. Technol. **36:**5279–5288.
95. **Sinton, L., R. Finlay, and D. Hannah.** 1998. Distinguishing human from animal faecal contamination in water: a review. N. Z. J. Mar. Freshw. Res. **32:**323–348.
96. **Stewart, J. R., R. D. Ellender, J. A. Gooch, S. Jiang, S. P. Myoda, and S. B. Weisberg.** 2003. Recommendations for microbial source tracking: lessons from a methods comparison study. J. Water Health **1:**225–231.
97. **Stoeckel, D. M., M. V. Mathes, K. E. Hyer, C. Hagedorn, H. Kator, J. Lukasik, T. L. O'Brien, T. W. Fenger, M. Samadpour, K. M. Strickler, and B. A. Wiggins.** 2004. Comparison of seven protocols to identify fecal contamination sources using *Escherichia coli*. Environ. Sci. Technol. **38:**6109–6117.
98. **Tartera, C., F. Lucena, and J. Jofre.** 1989. Human origin of *Bacteroides fragilis* bacteriophages present in the environment. Appl. Environ. Microbiol. **55:**2696–2701.
99. **U.S. Environmental Protection Agency.** 2005. Microbial source tracking guide. Document EPA/600/R-05/064. U.S. Environmental Protection Agency, Washington, DC.
100. **Whitlock, J. E., D. T. Jones, and V. J. Harwood.** 2002. Identification of the sources of fecal coliforms in an urban watershed using antibiotic resistance analysis. Water Res. **36:**4273–4282.
101. **Whitman, R. L., and M. B. Nevers.** 2004. *Escherichia coli* sampling reliability at a frequently closed Chicago Beach: monitoring and management implications. Environ. Sci. Technol. **38:**4241–4246.
102. **Wiggins, B. A.** 1996. Discriminant analysis of antibiotic resistance patterns in fecal streptococci, a method to differentiate human and animal sources of fecal pollution in natural waters. Appl. Environ. Microbiol. **62:**3997–4002.
103. **Wiggins, B. A., R. W. Andrews, R. A. Conway, C. L. Corr, E. J. Dobratz, D. P. Dougherty, J. R. Eppard, S. R. Knupp, M. C. Limjoco, J. M. Mettenburg, J. M. Rinehardt, J. Sonsino, R. L. Torrijos, and M. E. Zimmerman.** 1999. Use of antibiotic resistance analysis to identify nonpoint sources of fecal pollution. Appl. Environ. Microbiol. **65:**3483–3486.
104. **Wiggins, B. A., P. W. Cash, W. S. Creamer, S. E. Dart, P. P. Garcia, T. M. Gerecke, J. Han, B. L. Henry, K. B. Hoover, E. L. Johnson, K. C. Jones, J. G. McCarthy, J. A. McDonough, S. A. Mercer, M. J. Noto, H. Park, M. S. Phillips, S. M. Purner, B. M. Smith, E. N. Stevens, and A. K. Varner.** 2003. Use of antibiotic resistance analysis for representativeness testing of multiwatershed libraries. Appl. Environ. Microbiol. **69:**3399–3405.