Technical Report 363
March 2002

Microbiological Source Tracking Workshop:

Workgroup Findings and Recommendations

Prepared for EPA

Brock B. Bernstein

John F. Griffith

Stephen B. Weisberg

Southern California Coastal Water Research Project

Microbiological Source Tracking Workshop:

Workgroup Findings and Recommendations

Brock B. Bernstein
John F. Griffith
Stephen B. Weisberg

March 2002

Introduction

Microbiological source tracking methods are potentially powerful tools that are increasingly being used to define the nature of water-quality problems in watersheds across the nation. While these techniques show much promise, most are still in the early stages of development. Many of these techniques have been tested in a limited number of watersheds and with a limited number of possible sources.

Public agencies, particularly those in California, are preparing to spend several million dollars in applying these techniques to identify sources of bacteriological water contamination. These expenditures are based upon the assumption that these methods provide a firm scientific foundation for regulatory or remediation decisions.

This workshop brought together nationally recognized experts in environmental microbiology, molecular biology, and microbial detection methods with the intention of better defining the current state of knowledge regarding source tracking techniques, standardizing field and laboratory methods for those approaches that are most well developed, and drafting a protocol for a National Source Tracking methods comparison study that can be used to inform local decision makers about the reliability of the different methods.

Sponsored by the U.S. Environmental Protection Agency, California State Water Resources Control Board, Southern California Coastal Water Research Project, and National Water Research Institute, this workshop represented a unique opportunity for cooperative interaction among those involved in all facets of microbiological water quality.

Workshop Structure

The Microbial Source Tracking Workshop was held February 5 – 7, 2002 in Irvine, California. The main intent of the Workshop Steering Committee (Appendix 1) was to summarize existing knowledge about source tracking methods and to then use this review as the basis for designing studies to compare, evaluate, and validate a wide range of such methods. The intended product of the workshop was the outline of an approach for carrying out such studies.

The first day of the workshop was a public session attended by 226 people. This session was comprised of 15 presentations on source tracking issues and methods, including ribotyping and other genetic fingerprinting approaches, antibiotic resistance profiles, biomarkers, and more specific indicator organisms (Appendix 2). The variety of methods presented can be broadly categorized according to two key sets of features (Table 1). First, there are methods that distinguish among bacterial and/or viral samples based directly on their genetic makeup (genotypic methods) and those that distinguish among samples based on secondary characteristics such as antibiotic resistance (phenotypic methods). Second, there are methods that require the development of a background library or database against which to compare a sample and those that do not need such a library for their completion. The presentations focused on recent testing or applications of these methodologies. Many presenters indicated that they are pursuing multiple techniques.

Table 1. Two-way classification of some of the more widely used source tracking methods in terms of their focus on genotypes or phenotypic characteristics and their relative dependence on a background library or database of genotypic or phenotypic characteristics.

	Library dependent	*Library independent*
Genotypic	ribotyping bacterial community fingerprinting (t-RFLP) repetitive intergenic DNA sequences (PCR)	F+ coliphage human pathogenic virus detection (PCR or RT-PCR) Bacteroides genotyping (PCR) Enterotoxin biomarkers (PCR)
Phenotypic	multiple antibiotic resistance carbon source profiling	Enterotoxin biomarkers F+ coliphage serotyping IgA Antibodies

The second and third days of the workshop involved small-group discussion sessions that were limited to 40 invited participants. There were three such sessions, which concentrated on: 1) identifying a set of criteria to use in evaluating the outcomes of methods comparison studies, 2) design of such studies, and 3) defining adequate library size for the library-dependent methods. The invited participants were divided into four breakout groups of 10 people each according to affiliation and expertise. All groups were asked to address the same issues. After each session, all 40 participants were reconvened to compare recommendations among groups. The following sections of this report summarize the recommendations that evolved from the small group sessions.

Methods Evaluation Criteria

Four subgroups of workshop participants discussed conceptual, technical, policy, logistical, and economic issues involved in improving the application of microbiological source tracking methods to the point where they could be used consistently, broadly, and with confidence in a range of real-world conditions. There was unanimous agreement that a rigorous methods comparison and standards development effort is needed because the current situation is characterized by a lack of uniform standards or reference materials and ad hoc rules of thumb for sampling and measurement designs. After reviewing the work of the four separate subgroups, the participants agreed that the following set of criteria should provide the basis for any methods comparison study. These are divided into three tiers (measurement reliability, management relevance, cost and logistics) that reflect distinct categories of issues identified in the discussion (Table 2).

The ordering of the tiers reflects an inherent prioritization of the evaluation criteria. The most fundamental requirement of any method is that it provides accurate and repeatable results (Tier 1). Without this, other criteria, such as ease of communication to the public or a low cost per sample, would be meaningless. Likewise, adequate performance on the management criteria (Tier 2) is more important than cost and logistical issues (Tier 3). For example, a small library size or minimal training requirement would be less important if test results were not related to public health outcomes of interest or could not be readily communicated to key management audiences.

Table 2. Method evaluation criteria agreed on by the workshop participants. Criteria are divided into three categories, or tiers, that reflect different aspects of performance.

Category of criteria	Specific evaluation criteria
Tier 1: Measurement reliability	• Reproducibility of results within and across laboratories • Accuracy of classification of isolates into the correct group of sources (for library dependent methods) • Confidence that an identified indicator is from the presumed source (for library independent methods) • Level of resolution, or ability to discriminate among sources (i.e., human vs. non-human, livestock vs. wildlife, non-human species level, cattle from separate farms) • Matrix stability (in what matrices, e.g., saltwater, freshwater, turbid water, humic acid environments, is the method applicable?) • Geographical stability (over what area is the method applicable?) • Temporal stability (over what time frame is the method applicable?) • Confirmation by peer review
Tier 2: Management relevance	• Relationship to actual source(s) of contamination • Relationship to public health outcomes • Relationship to commonly used water quality indicators • Ease of communication to the public • Ease of communication to management audiences
Tier 3: Cost and logistics	• Equipment and laboratory facilities required • Training required • Library size required (for library dependent methods) • Library development effort per "unit" required (for library dependent methods) • Implementation time • Cost of ensuring results are legally defensible • Cost per sample, including all operations and maintenance overhead • Sample turnaround time

Tier 1 focuses on measurement issues that affect the accuracy and repeatability of test results. While accuracy and repeatability are important for any analytical method, there are several aspects of this that are specific to microbiological source tracking methods. For all such methods, the lack of widely accepted standardized techniques makes the issue of reproducibility of results both within and across laboratories more crucial than it might be for methods with a longer history of development and application. Another measurement issue involves how accurately different methods identify the source(s) of a particular isolate or indicator from a sample. While the manner is which this issue would be evaluated differs somewhat depending on whether the method depends on a library of isolates, the core question of accurate identification of isolates and/or indicators is basic to the acceptance and wider use of any method.

The working groups identified additional Tier 1 criteria related to a set of "stability" issues that are associated with repeatability. These issues include geographic, temporal, and matrix stability. Geographical stability refers to the size of the area over which a particular method (and the library or indicators it is dependent on) is applicable. Species of gut flora in a particular host species may vary spatially for a variety of reasons, and even the same bacterial or virus species

may vary genetically from one area to another, sometimes significantly so. Temporal stability refers to a method's applicability over time. The gut flora community of a host species, or even a host individual, has the potential to shift over time. Methods or indicator species that are robust to such changes are preferable to methods which require more frequent library renewal. Matrix stability refers to a method's performance in the presence of interferences that may obstruct or bias the method. For example, saltwater can inhibit culturing of some organisms. Organic compounds, such as humic acid, can impede PCR because it binds to DNA present in the sample, making the DNA unavailable to the enzymes used in the method. Depending on the particular kind of interference, methods will vary in their sensitivity to these factors.

Tier 2 focuses on management issues that affect the acceptance and usefulness of test results in decision making. The more directly a method targets actual sources of contamination, such as differentiating among all animal types as compared to only differentiating human from non-human sources, the more valuable it will be for decision making. A particular method is similarly strengthened the more it can be linked, through epidemiological studies or other means, to public health outcomes of concern. On the other hand, workshop participants understood that currently used water quality indicators, despite their weaknesses in these two respects, are widely used, have a long time series of data, and form the basis for regulatory policy. Microbiological source tracking methods that have a demonstrated relationship to the currently used water quality indicators would be preferable. Finally, preferable methods should readily understood by both public and management audiences. Ease of communication will help promote their widespread acceptance and use.

Tier 3 focuses on cost and logistical issues directly related to implementation. These are relatively straightforward and include items such as the expense of necessary measurement equipment and the degree of staff training required.

Recommended Study Design

The participants at the workshop unanimously agreed on the need for rigorous studies to compare methods on the basis of the criteria described above. The participants agreed on a study design structured around four phases that advance through successively more complex questions about method performance, beginning with single-source, within-laboratory studies and ending with diverse, multi-source watershed evaluations. The cost efficiency of the study could be maximized if the sources and/or locations involved in the early phases were continued through the later phases. The following sections describe each of the four phases:

Phase 1: Repeatability

The core question to be addressed in Phase 1 is: How repeatable is the measurement technique? More specifically, the participants wanted to determine if the a laboratory can produce the same results for the same samples in multiple runs of the same method. Moreover, they thought it was important to find out if multiple investigators in different laboratories could produce the same results for the same samples.

In this initial phase, multiple replicate samples from just a few unique sources would be distributed to a series of different laboratories to assess repeatability of each method both within- and between-laboratories. Some of the samples would be distributed as "knowns" associated

with a specific source, while others would be distributed as "blind" samples to assess the investigator's ability to match patterns with the known samples. In either case, samples in this phase would be distributed as isolates or in the simplest, most replicable format associated with that particular method. Since one purpose of this phase is to quantify repeatability among laboratories using the same method, considerable method standardization would be required prior to testing.

Several of the breakout groups suggested that this phase could be extended to also address temporal stability. Two types of temporal stability testing were identified. First, selected samples could be retained in the laboratory and processed at several successive time intervals. Second, fecal material from the same animal could be sampled over time to assess how repeatable or applicable a library developed at one time is to an environmental sample collected at a later time.

Phase 2: Accuracy – Laboratory samples

The core question to be addressed in Phase 2 is: Can methods accurately detect the source(s) in laboratory-created "blind" samples? This phase would be the next step in a continuum of increasingly sample complexity and would differ from Phase 1 in three ways. First, test samples would be distributed in an aqueous matrix to simulate the manner in which environmental samples would be received in a typical application. Second, a subset of the samples would include feces from multiple sources, again to more closely simulate the conditions encountered in typical applications. Third, the tests could be extended to incorporate matrix interferences (e.g. turbidity, salt) into the test samples.

This phase would proceed in two steps. In the first step, participating laboratories would be provided with the raw fecal material and/or isolates (from individual sources) required to build libraries needed to carry out the tests. Participating laboratories would be told the sources of these samples and participants generally agreed that laboratories should be allowed to develop libraries of whatever size they deemed appropriate. Participants also agreed that these samples should include the sources used in Phase 1, as well as those likely to be used in Phases 3 and 4, in order to promote efficiency by reducing the amount of repetitive library development that would otherwise be required through all four study phases. In addition, the sources used in Phases 1 and 2 should reflect sources that are of primary concern in watershed characterization studies and TMDL development. This will not only include an important aspect of realism in the study, but will also help promote interest in, and support for, the further application of microbial source tracking methods.

In the second step of Phase 2, participating laboratories would be provided a series of "unknowns" to identify. These samples would represent both single sources and various mixtures of sources created in the laboratory and placed into a water sample. For example, if four sources are used, laboratories might receive samples comprised of single sources and combinations of two, three, and all four sources. In addition, these samples could also be provided in different matrices, with the most likely ones mentioned by workshop participants being salt water, fresh water, turbid water and water containing humic acid. This would be intended to assess robustness of the methods when potential interferences are present.

Several participants expressed concern about the difficulty of combining different sources into a common sample, given the different pathogen density per unit fecal material among source types, but several other participants suggested that this could be accomplished by accurate quantification

of indicator bacteria density in the source material. Other participants suggested that requiring all laboratories to process the full set of samples (i.e., multiple combinations of sources in all matrices) would involve a large effort for each laboratory, which would be exacerbated if individual laboratories were to test more than one method. While no consensus was reached on the issue of laboratory workload, options were identified for randomizing the assignment of samples to laboratories that would allow for individual laboratories to test only a portion of the full complement of potential combinations of sources, matrices, and methods.

Phase 3: Accuracy – Ambient samples

The core question to be addressed in Phase 3 is: Can methods accurately detect source(s) in simple, dominant-source watersheds?

Phase 3 would extend the study to the testing of ambient samples (e.g., of water or soil) collected directly from the environment. However, in order to fit with the overall study strategy of proceeding in incremental steps of increasing complexity, the workshop participants agreed these samples would be drawn from a relatively simple watershed dominated by one or only a few sources. While this watershed might include the same sources used in Phase 2, samples could be drawn, for example, from runoff directly below known sources (i.e., dairy farm, leaking septic system) and would thus constitute a performance test of methods in a real-world situation, albeit a relatively simplified one. This approach will extend the Phase 2 tests by including soil bacteria, potential mixtures of sources, and other real-world interferences that cannot be accurately mimicked in laboratory-created samples. Thus, while the dominant source in the test samples will be known to the test managers, the actual makeup of the samples will be unknown both to the test managers and to the laboratories themselves. Depending on the results of Phases 1 and 2, test managers could calibrate the complexity and difficulty of Phase 3 samples by selecting test samples at different distances downstream of major sources within the watershed. This assumes that test samples collected further downstream will include more potential interferences.

Phase 3 will require careful selection of the test watershed. Workshop participants emphasized using a single source catchment. Some participants felt that having characterization data available on the test watershed might be valuable. Such information might include source identification from sanitary surveys, census data, wildlife surveys, and other studies, and should include detailed GIS coverages that enable mapping of test results if needed. This could enable test managers to develop a more rigorous set of expectations against which to evaluate the performance of participating laboratories.

Once again, several workshop participants recognized that using the same sources identified in Phase 2 will increase cost-efficiency, particularly for the library-dependent methods, and will ensure that geographic variability is minimized at the same time. Other participants suggested that an assessment of the geographic stability of methods could be incorporated into Phase 3 by expanding the design to include additional populations and/or watersheds.

Phase 4: Comparability in complex watersheds

The core question to be addressed in Phase 4 is: Do different methods produce comparable results in more complex systems with unknown sources?

Phase 4 would essentially mimic Phase 3, but with samples drawn from more complex watersheds with a wider variety of sources. While Phase 4 assesses whether different methods provide the same answer, it does not allow for easy explanation of differences among methods because underlying source materials and potential interferences are unknown. The workshop participants devoted somewhat less attention to this phase of the study, under the assumption that more detailed design decisions would be dependent on the results of the first three phases. However, they did stress that it should include, as a subset, the sources and watershed(s) used in Phases 1 – 3.

Depending on the number of sites, watersheds, and/or populations used, Phase 4 will permit a more thorough assessment of the geographic stability criterion as well as a start toward an assessment of the temporal stability criterion under the Tier 1 set of evaluation criteria, depending on the length of time the evaluation study continues.

Evaluating management, cost, and logistical criteria

Data from Phases 3 and 4 will also permit an evaluation of the Tier 2 set of evaluation criteria related to management, particularly the relationship of test results to actual sources of contamination and commonly used water quality indicators. Evaluation of the relationship of test results to public health outcomes will depend on the availability of adequate epidemiological data.

Each of the four phases will provide quantitative information needed to evaluate the cost and logistical criteria. This information can be combined with the broader experience of practitioners to develop a comprehensive assessment of the cost and logistical constraints and implications of each method.

Two of the breakout groups identified that part of evaluating cost-efficiency of a method involved quantifying multiple sources of variability inherent in applying the method. Methods with smaller cumulative variability within a source type will be more cost-effective, as they will require less, or less frequent, library development. Additionally, understanding these sources of variability provides the basis for developing a formal statistical model and would help furnish a basis for decision making about specific study design issues such as the necessary number of replicate samples. Sources of variability were grouped into four major categories:

- Laboratory
  - Within laboratory
  - Across laboratories
- Space
  - Within isolate
  - Across isolates within animal or sample
  - Across animals within herd or restricted population
  - Across herds or restricted populations of the same species within a watershed
  - Across watersheds
  - Across larger systems
- Time
  - Short-term
  - Seasonal
  - Year-to-year

- Matrix
  - Freshwater
  - Saltwater
  - Turbid water
  - Humic environments

Library Development

Practitioners have been making decisions about library size for the library-dependent methods based on logistical and cost constraints and/or on ad hoc rules of thumb developed through practical experience. No studies, however, have yet rigorously evaluated the minimum or optimal library size in terms of the kinds of evaluation criteria listed above. The workshop participants agreed unanimously on the need for a thorough study of this issue.

Library size is directly related to the accuracy of classification, matrix stability, level of resolution, and geographic and temporal stability criteria under the Tier 1 set of evaluation criteria. Participants agreed that this is an optimization problem in which the goal is to find a point of diminishing returns after which extra sampling to build a larger library adds less and less to a method's performance on the relevant evaluation criteria. A key question for both managers and practitioners is the degree to which a library developed for a particular watershed or region will apply to other watersheds or regions. A useful approach to this question recommended during the workshop was to define a set of nested watersheds, selected in coordination with Phases 3 and 4 of the study design described above, and determine the optimal library size for successively larger geographic areas.

The workshop participants also reached several broad conclusions about the issue of library development. First, they agreed that there is not enough knowledge at present to determine optimal library size for the library dependent methods. Second, they agreed that the optimal size would, to some extent, probably be method and watershed specific. Given the way that libraries are used in different methods, as well as differences in genetic variability across target organisms, it is highly unlikely that a single size library will prove optimal in all situations. Third, they agreed that statistical analyses of the large amount of existing data, if properly integrated, could provide useful answers about the sources of variability listed above and therefore about optimal library size. Fourth, they agreed that such analyses would not replace the need for a rigorously designed study that would specifically address the issue of library size. The ultimate goal of such efforts would be to develop an ability to model or predict the size of a needed library based on watershed characteristics such as the identity of sources and the loading processes that move contaminants through a system.

Thus, near-term decisions about library size will necessarily have to be based on best judgment and practical experience. There was general support among workshop participants for the idea that, in the near future, it would be possible to develop some preliminary guidelines based on statistical analyses of existing data. SCCWRP volunteered to begin such analyses as soon as researchers provide data. Two analyses discussed included a preliminary assessment of geographic variability in isolates and an analysis of the breakeven point at which the number of unique patterns identified per number of isolates begins to level off. Over the longer term, the results of a more rigorous evaluation would improve on these preliminary guidelines. Finally, while participants believed that developing a national library of geographically stable isolates would be of benefit, they agreed that more local and regional work should be carried out first.

Conclusion

Workshop participants were enthusiastic about the rapid advances in source tracking techniques that had been made in the last several years. They also felt that many of the techniques presented at the workshop would fare well in the tests described above and such tests would dramatically improve the rigor of microbial source tracking methods. There was recognition, though, that because of the diversity of animal types, as well as intrinsic evolutionary processes in microbial organisms, these methods will never be 100% accurate. The challenge for practitioners is to recognize that there will always be some patterns that cannot be classified and to determine the point at which increased effort devoted to refining methods and building libraries begins to produce diminishing returns.

The workshop participants were also unanimous in agreeing on the value of the workshop. Recognizing that the field will continue to advance rapidly over the next several years, participants agreed on the need for additional periodic workshops in the future, perhaps every one or two years.

Appendix 1: Steering Committee

Gerard Stelma, Ph.D.
U.S. Environmental Protection Agency

Alfred Dufour, Ph.D.
U.S. Environmental Protection Agency

Fred Genthner, Ph.D.
U.S. Environmental Protection Agency

Fred Kopfler, Ph.D.
U.S. Environmental Protection Agency

Don Reasoner, Ph.D.
U.S. Environmental Protection Agency

Stephen A. Schaub, Ph.D.
U.S. Environmental Protection Agency

Stephen B. Weisberg, Ph.D.
Southern California Coastal Water Research Project

Appendix 2: Presented Papers

Lessons Learned and Questions Unanswered from 5 Years of Bacterial Source Tracking
Valerie J. Harwood, Ph.D.
Assistant Professor, Department of Biology
University of South Florida

Microbial Source Tracking: Principles and Practice
Mansour Samadpour, Ph.D.
Assistant Professor, Department of Environmental Health
University of Washington

Ribotyping Enterococci
Peter G. Hartel, Ph.D.
Associate Professor, Department of Crop & Soil Sciences
University of Georgia

Fecal Source Tracking with Bacteroides
Katherine G. Field, Ph.D.
Assistant Professor, Department of Microbiology
Oregon State University

Urbanization and Coastal Water Quality: What Can Molecular Fingerprinting Tell Us?
Michael G. LaMontagne, Ph.D., and Patricia A. Holden, Ph.D.
The Donald Bren School of Environmental Science & Management
University of California, Santa Barbara

Comparison of Environmental and Clinical Isolates of Escherichia Coli Using Various Genetic Fingerprinting Methods
Cindy H. Nakatsu, Ph.D., Byoung-Kwon Hahm, Ph.D., Jennifer Fiser, Steven Yeary, and Arun Bhunia, Ph.D.
Purdue University

Bacterial Endemism and Co-speciation
James T. Staley, Ph.D.
Professor, Department of Microbiology
University of Washington

Source Tracking Fecal Bacteria in the Environment Using rep-PCR DNA Fingerprinting: Prospects and Problems
LeeAnn K. Johnson, Mary B. Brown, Priscilla E. Dombek, Ph.D. and Michael J. Sadowsky, Ph.D.
Department of Soil, Water, and Climate
University of Minnesota

Microbial Source Tracking Using Antibiotic Resistance Analysis
Bruce A. Wiggins, Ph.D.
Professor, Department of Biology
James Madison University

Animal Source Tracking: A Complement to Microbial Source Tracking
R.D. Ellender, Ph.D., B.L. Middlebrooks, Ph.D., Shiao Wang, Ph.D., Dawn Rebarchik, and D. Jay Grimes, Ph.D.
College of Science and Technology
University of Southern Mississippi

Carbon Source Profiles, Pulsed-field Gel Electrophoresis Patterns, and Antibiotic Resistance Analysis
Charles Hagedorn, Ph.D.
Professor, Department of Crop and Soil Environmental Sciences
Virginia Polytechnic Institute and State University

Coliphage Tracking to Identify Sources of Fecal Contamination
Mark D. Sobsey, Ph.D.
Professor, Department of Environmental Sciences
University of North Carolina

Source Tracking of Fecal Waste Material in Environmental Waters Using a Biomarker Based on Enterotoxin Genes in E. Coli
Betty H. Olson, Ph.D., Leila A. Khatib, and Eunice C. Chern
Department of Environmental Analysis
University of California, Irvine

Detection of Enteroviruses Using PCR-based Techniques for Source Identification and Assessment of Microbiological Water Quality
Rachel T. Noble, Ph.D.
Assistant Professor, Institute of Marine Science
University of North Carolina

Adenovirus as an Index of Human Viral Contamination
Sunny C. Jiang, Ph.D.
Assistant Professor, Environmental Analysis and Design
University of California, Irvine

Application of a methods to identify coliform pollution sources using multiple antibiotic resistance, selected molecular techniques and GIS spatial analysis
Geoffrey Scott
NOAA/NOS, Center for Coastal Environmental Health and Biomolecular Reasearch

Appendix 3: Abstracts of Presentations

Lessons Learned and Questions Unanswered from 5 Years of Bacterial Source Tracking

Valerie J. Harwood
Assistant Professor, Department of Biology
University of South florida

For over a decade, the source of fecal indicator bacteria (such as fecal coliforms, Escherichia coli, and enterococci) has been a pressing question in water-quality assessment. Accurate risk analysis, effective remediation efforts, and valid total maximum daily load assessments all depend upon knowledge of the source (i.e., human, dog, cattle, wild animals) of contamination. Reports of fecal indicator organisms isolated from cold-blooded vertebrates and plants, and the steadily growing realization that the long-term survival and growth of indicator bacteria outside their host is not uncommon, have provided considerable impetus to the effort to better characterize the various sources and population dynamics of indicator organisms in natural waters. Ultimately, the goal of these efforts is to develop the tools required to define the sources of fecal pollution in surface waters and groundwater.

Bacterial source tracking (BST) is a term that was first coined by Hagedorn and Wiggins in the website http://www.bsi.vt.edu/biol_4684/BST/BST.html to include the various sub-typing methods, such as antibiotic resistance analysis, pulsed-field gel electrophoresis, ribotyping, and restriction fragment length polymorphism, that are being developed for discriminating between sources of indicator bacteria in natural waters. The term can be expanded to microbial source tracking, which includes methods based on sub-typing viruses or bacteriophages.

Implicit in the development of all of these sub-typing methods is the hypothesis that the sub-types of an indicator organism population are unequally distributed among host populations. Succinctly phrased, we hypothesize that indicator subtypes display some degree of host specificity. That this host specificity is not absolute has become increasingly clear in the last decade, shifting the hypothetical paradigm from an epidemiological one (we can clearly assign a particular subtype to one host and no other) to a population biology viewpoint (Subtypes A, B, and D are more frequently found in Host 1, while Subtypes C, E, and F are more often found in Host 2).

Most BST studies have certain similarities. A group of host animals whose feces could potentially impact the waters under study are identified. The usual suspects include sewage, wild animals, dogs, and livestock, but can vary greatly depending upon the characteristics of the watershed(s) in question. Feces from individual host animals are collected, and a library of isolates (bacterial or viral) obtained from the feces is sub-typed by the chosen BST method(s). The patterns, or "fingerprints", of the isolates are analyzed statistically, and the extent to which the isolates from each source can be discriminated from the isolates from other sources (the "internal consistency" of the library) is evaluated. While cluster analysis can demonstrate the extent to which isolates from different sources form discreet groups, multivariate analyses such as discriminant analysis

Correspondence should be addressed to:

Valerie J. Harwood, Ph.D.
Assistant Professor, Department of Biology
University of South Florida
4202 East Fowler Avenue, Tampa, FL 33620-5550
Tel: 813-974-1524 Fax: 813-974-3263 E-mail: vharwood@chumal.cas.usf.edu

are more useful for predicting membership in a group, which is necessary when isolates from water samples (unknown source) are analyzed. A common expression of the internal consistency of a library is the average rate of correct classification. The library is "self-crossed" for this analysis; that is, the isolates in the library are used as both the calibration dataset (the standards) and the test dataset (the unknowns). The number of isolates assigned to the correct categories is divided by the total number of isolates in the library, and is given as a percentage.

Our experience with antibiotic resistance analysis suggests that there is relatively little difference in discrimination between sources, whether fecal coliforms, *E. coli*, or enterococci are used as the indicator. In a study carried out in Florida, the average rate of correct classification for a fecal coliform library (6,144 isolates; six source groups) was 63.9 percent, and the corresponding average rate of correct classification for a fecal streptococcus library (4,619 isolates; six source groups) was 62.3 percent (Harwood et al., 2000). Another set of Florida libraries with six source groups constructed using a different antibiotic scheme yielded correct classification rates of 57.1 percent for fecal coliforms (3,295 isolates), 61.5 percent for *E. coli* (2,565 isolates), and 52.6 percent for enterococci (1,889 isolates). Preliminary data from a California study shows similar results.

What is clear from our experience with antibiotic resistance analysis is that very small libraries are not very useful for BST. Generally, small libraries (less than 300 isolates per source) have relatively high correct classification rates when analyzed by discriminant analysis; however, frequently, there is an element of artifact in the "accuracy" of these small libraries, which can be tested by two methods. First, the patterns contained in the library can be randomly reassigned to sources, and the statistical analysis of choice carried out. If the "correct classification rate" is still high for the randomly assigned libraries, it is evident that artificial grouping of the patterns is occurring, but if the "correct classification rate" approaches the probability of the pattern being assigned to a given group by chance (i.e., 0.20, or 20 percent, in a library with five source categories), it strongly suggests that the correct classification rates in the "real" library are due to the presence of coherent source groups.

The second method for testing the statistical significance of the assignment of patterns to source groups is the F ratio, which is calculated as F=[n1n2(n1=n2-p-1)/(n1=n2(n1-n2-2)p]D2. Pair-wise comparisons of the distance between source categories are made, and F is compared to a statistically derived critical value; thus, a probability that the observed discrimination between sources would have occurred by chance can be calculated. Small libraries with high average rates of correct classifications (>85 percent) may "fail" the F test (P>0.05 between one or more pair-wise comparisons of source categories), while large libraries (500 to 1,000 isolates per source) with relatively low average rates of correct classifications (55 to 65 percent with five source categories) show statistically significant discrimination between sources. Patterns derived from molecular typing methods, such as ribotyping, can also be analyzed by randomizing data sets and by the F test.

An inherent disadvantage of molecular typing methods, such as pulsed-field gel electrophoresis and, in particular, ribotyping, is their low throughput and expense compared to antibiotic resistance analysis. The expense and feasibility of library construction become critical issues for the agencies or municipalities funding the study; thus, application of the most rigorous methods possible to determine appropriate library sizes is crucial to developing the BST field.

A related BST issue is the appropriate number of isolates to analyze from water samples whose patterns are, ultimately, compared to those of the library for source identification. In practice, a somewhat arbitrary approach to determining the number of isolates that should be analyzed has

been taken by most investigators (including this one), and has been based on what seems like a reasonable number. In the case of antibiotic resistance analysis, that number is generally 48 to 96 isolates per sample, which fit neatly into the microtitre plate. In the case of ribotyping, that number is likely to be five to 10. Although it seems, intuitively, that one should subtype more isolates from a water sample that contains higher numbers of the indicator to obtain a representative sample, the logistics of implementing such a practice are daunting (considering laboratory workload and BST study cost-estimates). Nonetheless, the issue is too important to allow it to be determined completely by logistical considerations.

To address this question, we hypothesized that a regular relationship between indicator organism number and diversity (defined here by the number of different patterns observed per 30 patterns) in water samples might exist, and that such a relationship could be used to determine the number of isolates that should be sub-typed from a given water sample. To test that hypothesis, samples from three small tributaries to the Hillsborough River in Florida with fecal coliform counts ranging from 113 colony-forming units (CFU)/100 milliliter (mL) to 25,000 CFU/100 mL were analyzed by antibiotic resistance analysis and ribotyping. Thirty MUG-positive, oxidase-negative fecal coliform isolates (presumed to be *E. coli*) were sub-typed by each method from each water sample. Intriguingly, diversity was highest in the water sample with the lowest fecal coliform number, as assessed by both antibiotic resistance analysis and ribotyping. The number of patterns identified by antibiotic resistance analysis was inversely correlated with the log-transformed fecal coliform count ($r2 = 0.857$); however, one ribotype sample differed from the trend. The water sample with the median fecal coliform count (850 CFU/100 mL) displayed a clonal ribotype; that is, the ribotypes of the 30 isolates were identical. Obviously, much more work along these lines is necessary to draw conclusions about the diversity of indicator subtypes in water and how these parameters may be exploited to better assess sources of fecal contamination, but it seems plausible that a diversity index could be useful for identifying certain types of point source pollution.

In an ideal world, the patterns displayed by indicator organisms that reside within a host or host population would be very stable for the purposes of BST; however, it is known that environmental influences, including diet changes and antibiotic treatment, can affect the antibiotic resistance patterns of bacterial inhabitants of the gastrointestinal tract. The stability of the antibiotic resistance phenotype and the genotype (by ripotyping) of *E. coli* in the feces of one dog was tracked before the animal received antibiotic treatment, at the end of treatment (Clavamox [clavulinic acid + amoxicillin] for 2 weeks followed by enrofloxacin for 3 weeks), and several weeks after treatment. We hypothesized that both the antibiotic resistance patterns and the ribotypes of the bacterial population would shift with antibiotic treatment, which proved to be the case. Ribotyping indicated that the population of *E. coli* in the dog's feces was clonal and stable before antibiotic treatment. During antibiotic treatment, several ribotypes that were not previously observed were noted, and although the "original" ribotype was still represented, it appeared at similar frequency to many of the other patterns. Four weeks after the cessation of antibiotic treatment, a new clonal pattern emerged and was stable with repeated sampling over time. These data indicate that one should not assume that the genotypic patterns of indicator organisms within an individual are more stable than the phenotypic patterns.

Issues of library accuracy and representativeness, reproducibility of the sub-typing method, expense, laboratory sophistication required, and sample size are critical parameters that must be addressed as BST methods are developed and tested. Perhaps the most crucial element for rapid progress in the BST arena is agreement among investigators as to the most useful paradigm for approaching these questions. Identification of the source(s) of fecal inputs to natural waters can be approached from at least three distinctly different directions. The "Holy Grail" approach

would seek to define an indicator that is restricted to one species (usually human), that is ubiquitously distributed among the host population, and that is fairly numerous and readily measured. To date, the Holy Grail remains elusive. The second approach could be considered one of exclusivity, in which the only data that are included in analysis are those subtypes that seem to be restricted to certain host species. More broadly distributed patterns are considered noise and are not used in the data analysis. This approach begs an important question: couldn't the "unique" patterns be unique in the data set only because the true population has been inadequately sampled? It also excludes from analysis all isolates from sampled water that belong to broadly distributed subtypes, even if they comprise the dominant pattern among the sampled isolates. The third approach depends upon the principles of population biology, where exclusivity of membership in a source category is not the most important characteristic of a subtype. Instead, the composition of populations is much more important than that of individual isolates, and all of the data are useful rather than a restricted segment. Although the latter approach demands sophisticated statistical analysis and does not provide the yes/no, black/white answers sought by some, it ultimately has the greatest potential to legitimately answer questions about microbial sources in natural waters.

References

Harwood, V.J., J. Whitlock, and V.H. Withington (2000). "Classification of the antibiotic resistance patterns of indicator bacteria by discriminant analysis: Use in predicting the source of fecal contamination in subtropical Florida waters." *Applied and Environmental Microbiology*, 66: 3,698-3,704.

VALERIE HARWOOD is an Assistant Professor in the Department of Biology at the University of South Florida, where she has taught since 1998. Prior, she was an Assistant Professor in the Department of Natural Sciences at the University of North Florida. Harwood's research interests include the microbial ecology of fecal indicator organisms in natural waters, bacterial/microbial source tracking, and the prevalence and human health implications of antibiotic resistance in fecal indicator bacteria. She is also interested in protein biochemistry. Harwood received a B.A. in French from Iowa State University, a B.A. in Biology from State University of New York at Plattsburgh, and a Ph.D. in Biomedical Sciences from Old Dominion University and Eastern Virginia Medical School, with post-doctoral work at the University of Maryland Center of Marine Biotechnology.

Microbial Source Tracking: Principles and Practice

Mansour Samadpour
Assistant Professor, Department of Environmental Health
University of Washington

One of the major challenges in environmental microbiology is to identify and track the sources of fecal microbial pollution that impact bodies of water. Over the years, a number of methods have been proposed with very little success (Feachem, 1975; Kaspar et al., 1990). Over the past 12 years, we have developed the microbial source tracking (MST) method, which can be used to identify the sources of microbial pollution that impact bodies of water. The method can also be used to quantify the impact of each one of the sources. The MST method has been developed on the basis of the principles of microbiology, epidemiology, molecular epidemiology, sanitary engineering, and hydrogeology.

The ability to identify the sources of microbial pollution of fecal origin depends on whether or not there are host-specific lineages (*Eco*types) within a given species of bacteria (Maynard Smith, 1996), particularly among bacterial species like *Escherichia coli*, which are used as indicators of the microbiological quality of water. The existence of *Eco*types would then allow the principles of epidemiology and molecular epidemiology to be applied in identifying the host species for bacterial strains that are isolated from environmental samples in the same manner that outbreaks of infectious diseases are detected (Samadpour, 1995). There are several foundations on which the MST method is based:

- First, in any given pollution scenario, there are multiple contributing animal sources of microbial pollution, each of which has its own unique clones of bacteria that constitute their normal flora.

- Second, collections of isolates from an appropriate bacterial species can be compiled from the polluted sites and the suspected animal sources of pollution, which are identified through a sanitary survey of the region surrounding the polluted site.

- Third, using an appropriate molecular subtyping method, all the bacteria in the collection can be subtyped.

- Finally, the genetic fingerprints of the bacterial isolates from the polluted site can be compared to those of the bacteria from the suspected animal sources.

- When a strain of bacteria with an identical genetic fingerprint is isolated from both a polluted site and a suspected animal source, the animal is implicated as a contributor of that specific clone of the bacteria to the polluted site.

Correspondence should be addressed to:

Mansour Samadpour, Ph.D.
Assistant Professor, Department of Environmental Health
School of Public Health and Community Medicine
University of Washington, Seattle, WA 98195-7234
Tel: 206-543-5120 Fax: 206-543-8123 E-mail: mansour@u.washington.edu

MST Underlying Assumptions

The MST method is based upon two principles. The first principle is that the bacterial population genetic structure is clonal. This is a well-established element of microbial genetics. Bacteria divide by binary fission. The two daughter cells that are generated as a result of this cell division are virtually identical in all aspects. All descendents of a common ancestral cell are genetically related to each other. Over time, members of a given clone may accumulate genetic changes, which will cause them to diverge from the main lineage and to form one or several new clonal groups. MST makes use of the clonal population structure of bacteria to classify organisms based on their genetic fingerprints into groups of clonal descent.

The second principle behind the MST methodology is the assumption that within a given species of bacteria, various members have adapted to living/environmental conditions in specific hosts/environments. As a result, there is a high degree of host specificity among bacterial strains that are seen in the environment. A bacterial strain that has adapted to a particular environment or host (e.g., animal intestinal tract) is capable of colonizing that environment and competing favorably with members of its indigenous flora. Such a bacterial strain is called a resident strain. Resident strains are usually shed from their host over a long period of time, thus providing a characteristic signature of their source. A transient strain is a bacterial strain that is introduced into a new environment or host, but cannot colonize and persist in that environment. If a host is sampled over time for a given species of bacteria, a few resident strains are consistently being shed while a large number of transient strains are shed for brief lengths of time. A study conducted by Hartl and Dykhuizen (1984) illustrates this point. Over a period of 11 months, 22 fecal samples were taken from a single individual. A total of 550 *E. coli* isolates were characterized, of which two were considered to be resident strains, appearing 252 times. We have accumulated considerable evidence to support this assertion for *E. coli*. Our data show that using our subtyping method (ribosomal RNA typing using two restriction enzyme reactions), more than 96 percent of *E. coli* strains are seen in only one host species (or group of related species) (Buck, 1998).

Subtyping Methods Used in Source Tracking Studies

Another important factor in determining clonality is methodology. Our laboratory's ability to subtype microbes and divide them into groups of clonal origin largely rests upon the sensitivity of the methods that are used to subtype the organisms. For instance, consider a hypothetical collection of 100 *E. coli* strains isolated from 100 different source samples at 100 different times from 100 different sites, which is to be analyzed with three methods representing low, medium, and high degrees of sensitivity. The first method, which has low sensitivity, may divide the 100 strains into eight groups, while the second method divides them into 40 groups and the third method, with a high degree of sensitivity, divides them into 95 groups. A researcher using either of the first two methods may erroneously cluster unrelated strains of *E. coli* as members of the same clone. If this was a source tracking study, the practical implication is that a water isolate that is different from a bovine strain, but is seen by the subtyping method as being identical, will be labeled as *E. coli* of bovine origin; however, this isolate may, in reality, have come from a source other than bovine. While insensitive subtyping methods are not suitable for use in MST studies, we have also found that very sensitive subtyping methods may not be as useful in source tracking studies as one would predict. The main reason is that highly sensitive subtyping methods can detect minute genetic changes that have occurred very recently, on the order of weeks to months. The practical implication of this is that the level of diversity seen by these methods is so high that the number of samples needed to achieve a field-laboratory survey of the study area that

is representative of the population of a given species in a watershed would require the analysis of thousands of bacterial isolates, which would make the venture prohibitively expensive.

Ribotyping

The key methodological problem in tracing sources of bacterial contamination in the environment was the lack of a universal single-reagent typing scheme for bacteria. This has been overcome by the work of several investigators in the fields of population genetics, molecular systematics, and molecular epidemiology. Grimont and Grimont (1986) showed that DNA probes corresponding to specific regions of the rRNA operon can be used to speciate bacteria. Stull et al. (1988) used the rRNA operon to study the molecular epidemiology of several species of bacteria. To trace the indicator bacterium, *E. coli*, from water to its specific source, the bacterial strain must first be uniquely identified. Populations of *E. coli*, like other bacteria, are essentially composed of a mixture of strains of clonal descent. Due to the relatively low rates of recombination, these clones remain more or less independent (Selander et al., 1987). These clones, or strains of bacteria, are uniquely adapted to their own specific environments. As a result, the *E. coli* strain that inhabits the intestines of one species is genetically different from the strain that might inhabit another species.

Ribosomal ribonucleic acids (rRNA), which are integral to the machinery of all living cells and tend to be very highly conserved, make an ideal choice of target in interstrain differentiation. Since the *E. coli* chromosome contains seven copies of the rRNA operon, a rDNA probe can be used as a definitive taxonomic tool (Grimont and Grimont, 1986). That is, when digested with restriction enzymes, resolved by agarose gel electrophoresis, and transferred to a membrane and hybridized with an rRNA probe, an *E. coli* chromosome will produce several bands to create a specific restriction fragment length polymorphism (RFLP) pattern that can be used to uniquely identify the bacterial strain.

Existence of Source-specific Lineages of Bacteria

To investigate the presence of *Eco*types (resident clones) in the natural population of *E. coli*, we conducted a study using a collection of 2,142 *E. coli* strains isolated from 402 human and animal source samples. The collection was subtyped by the ribosomal RNA typing method (Grimont and Grimont, 1986) using two restriction enzymes (*Eco*RI and *Pvu*II).

Table 1 shows the total number of ribotypes for single and double enzyme analysis and includes the percentage of resident, source related, and transient ribotypes. Roughly half of the total ribotypes for *Pvu*II (43 percent) and *Eco*R1 (51 percent) analyzed separately are resident strains, while 94 percent of the total ribotypes are residents when the enzymes are combined. Total ribotypes also increases from 514 to 723 to 844 for *Pvu*II, *Eco*R1, and *Pvu*II/*Eco*R1, respectively.

Table 1. Summary of Ribotype Totals for Single and Double Enzyme Analysis

Enzyme	Total Ribotypes	Source Specific Ribotypes	Source Related Ribotypes	Transient Ribotypes
*Pvu*II	514	221 (43%)	31 (6%)	262 (51%)
*Eco*R1	723	368 (51%)	38 (5%)	317 (44%)
*Pvu*II and *Eco*R1	873	823 (94%)	18 (2%)	32 (4%)

The data clearly shows that while ribotyping using a single enzyme has insufficient sensitivity to identify resident clones, ribotyping using two enzymes shows the existence of host-specific lineages in the natural population of *E. coli*.

Comparison of Ribotyping and Antibiotic Resistance Analysis

The experiment was designed as a blind study in which a set of 120 stab cultures was prepared from 40 *E. coli* strains (each in triplicate) isolated from seven different sources. The collection included *E. coli* strains isolated from a cat (1), cows (10), a harbor seal (1), horses (2), humans (6), sea gulls (8), and sea lions (12). The stab cultures were randomly numbered (from 1 to 120), and sent to our laboratory with no additional information. Upon arrival, each isolate was plated on MacConkey media to check its purity. After the compilation of ribotyping, the results were submitted to our study partners, who in turn sent back the isolate log.

Antibiotic resistance analysis for a set of 10 antibiotics (Amikacin, Ampicillin, Chloramphenicol, Ciprofloxacin, Gentamicin, Kanamycin, Nalidixic Acid, Streptomycin, Tetracycline, and Trimethoprin) was determined using the standard disk diffusion assay (National Committee for Clinical Laboratory Standards, 1999). All of the 120 isolates were susceptible to six of the 10 antibiotics (Amikacin, Chloramphenicol, Ciprofloxacin, Gentamycin, Kanamycin, and Nalidixic Acid). Fifty-three of the 120 isolates were susceptible to all 10 the antibiotics used in the study. These included 22 of the 30 cow isolates, one of the three cat, five of six horse, seven of the 18 human, four of the 24 sea gull, 13 of the 36 sea lion, and one of the three seal isolates.

Sensitivity. Sensitivity refers to the inter-strain differentiative ability of a given method. The collection of 120 isolates, 40 strains in triplicates, was divided into 27 different ribogroups by the two-enzyme ribotyping method. The same collection was then divided into six groups by antibiotic resistance analysis. The index of diversity (defined as the number of groups divided by the number of isolates) for the two methods was 0.67, and 0.15, for ribotyping and antibiotic resistance analysis, respectively.

Reproducibility. The study was designed to measure reproducibility by the ability of the methods to group members of the triplicate sets correctly. The rates of correct classification for ribotyping and antibiotic resistance analysis was 100 and 90 percent, respectively.

Host Specificity. This refers to the ability of the method to group the isolates in the collection on the basis of their host origin. Ribotyping correctly grouped the members of each of the 40

triplicate sets. None of the seven animal groups shared any of the ribotypes. The ribotypes for this limited set of isolates showed 100-percent host specificity. With a larger set of isolates (2,143 strains), our results indicate a host specificity of 96 percent (Buck, 1998). For the antibiotic resistance analysis method, while four of the six (66 percent) patterns were only seen in one host, the four patterns represented only 10 percent of the isolates. The method had no host specificity for the 90 percent of the remaining isolates in the study.

Universal Presence of the Subtyping Markers in the Target Organism. One of the main advantages of the ribotyping method is that the ribosomal RNA genes are universally present in all bacterial cells. In comparison, 120 of the 120 isolates in the study were susceptible to six of the 10 antibiotics used in the study. With the antibiotic resistance analysis method, 81 percent of the isolates were susceptible to all 10 of the antibiotics.

Conclusions Regarding Antibiotic Resistance Analysis. The data indicates that the antibiotic resistance analysis method lacks the sensitivity, reproducibility, and host specificity required of a method to be used in source tracking and source identification. This is hardly surprising since the same conclusions have been reached regarding the use of antibiotic resistance analysis in epidemiological investigations; however, we believe that the antibiotic resistance analysis could be a useful tool to supplement the data generated by a more sensitive method, such as ribotyping. The antibiotic resistance analysis will result in enhancing our understanding of the movement and transfer of antibiotic resistance genes in the environment.

MST at the Institute for Environmental Health in Seattle, Washington

We have conducted more than 80 MST studies across the United States and in Canada. The studies have included:

- Studies of water sources for municipal drinking water utilities.
- Identifying the sources of microbial pollution in swimming beaches, lakes, rivers, shellfish growing areas, groundwater, and storm drains.
- Assessing the impact of point sources.
- Assessing the impact of tourism on water quality in national parks.
- Identifying reservoirs of human pathogens in watersheds.
- Total maximum daily load studies.

Currently, we have the ability to conduct 40,000 molecular subtyping reactions a year. The methods that are in use at the Institute for Environmental Health include:

- Ribotyping.
- Pulsed-field gel electrophoresis.
- MRF.
- PCR-based subtyping methods.
- DNA sequencing.
- Phage typing.

- Antimicrobial resistance analysis.
- Biotyping.
- Biomarkers of human and animal impact (detection of viruses, bacteria, and microbial genes that are host species specific).
- Detection of microbial pathogens (bacterial, viral, parasitic, and fungal).
- Characterization of virulence determinants,
- Host specificity.

References

Buck, F. (1998). *Comparison of one vs. two enzyme ribotyping in microbial source tracking studies*. Master's Thesis, Department of Environmental Health, University of Washington.

Feachem, R. (1975). "An improved role for fecal coliform to fecal streptococci ratios in the differentiation between human and nonhuman pollution sources." *Water Research*, 9: 689-690.

Grimont, F., and P.A.D. Grimont (1986). "Ribosomal ribonucleic acid gene restriction patterns as potential taxonomic tools." *Annules de L'Institut Pasteur/Microbiology*, 137 B(2): 165-175.

Hartl, D.L., and D.E. Dykhuizen (1984). "The population genetics of *Escherichia coli*." *Annual Review of Genetics*, 18: 31-68.

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." *Canadian Journal of Microbiology*, 36: 891-894.

Maynard Smith, J. (1996). "Population genetics: An introduction." *Escherichia coli* and Salmonella: *Cellular and molecular biology*, Second edition, F. C. Neidhardt (ed.). ASM Press, Washington D.C.

National Committee for Clinical Laboratory Standards (1999). *Methods for dilution of Antimicrobial susceptibility tests for bacteria that grow aerobically*. Third Edition; Approved Standard M7-A3. National Committee for Clinical Laboratory Standards, Villanova, Pennsylvania.

Samadpour, M. (1995). "Molecular epidemiology of *Eshcherichia coli* O157:H7 by restriction fragment length polymorphism using shiga-like toxin genes." *Journal of Clinical Microbiology*, 33: 2,150-2,154.

Selander, R.K., D.A. Caugant, and T.S. Whittam (1987). ÒGenetic structure and variation in natural populations of *Escherichia coli*." *Cellular and molecular biology*, F.C. Neidhardt et al. (eds.). American Society for Microbiology, Washington D.C.

Stull, T.L., J.J. LiPuma, and T.D. Edlind (1988). ÒA broad-spectrum probe for molecular epidemiology of bacteria: Ribosomal RNA." *Journal of Infectious Diseases*, 157(2): 280-286.

MANSOUR SAMADPOUR has taught at the University of Washington, since he received his doctorate in 1990. Currently, he is Assistant Professor in the Department of Environmental Health in the School of Public Health and Community Medicine. Research interests include the microbiology of water, wastewater, food and air; fate and transfer of microbial contaminants in the environment; and early detection of food and waterborne outbreaks. Samadpour has worked on developing the MST method for the last 12 years. His work in the area included more than 80 MST studies in the U.S. and Canada, including total maximum daily load studies; MST studies of swimming beaches, shellfish beds, drinking water sources, drinking water distribution systems, groundwater, impact assessment for point sources (industrial wastewater and municipal wastewater effluent); hazard analysis critical control point analysis of watersheds; identification and tracking of microbial pathogens in water; and investigation of waterborne outbreaks. Samadpour received both a B.S. and M.S. in Microbiology and a Ph.D. in Food Science and Technology (minor in Molecular Biology) from the University of Washington.

Ribotyping Enterococci

Peter G. Hartel
Associate Professor, Department of Crop & Soil Sciences
University of Georgia

Author's Note: Because of the interest of regulatory personnel in microbial source tracking, I have written this extended abstract more informally and tried to link our work with others presenting their research at this workshop. Also, in the spirit of the workshop, I have added more speculation. If it has not already been explained in a previous presentation, our research focuses on ribotyping, which looks at "fingerprints" of the portion of bacterial DNA encoding for ribosomal RNA (rRNA). By matching fingerprints of a specific bacterial species isolated from various warm-blooded animal species against fingerprints of the same bacterial species isolated from environmental samples, the source of the non-point fecal contamination can be identified. Although ribotyping is expensive and time-consuming, it is extremely reproducible. Our most recent work is with the enterococci, a subgroup of the fecal streptococci. Enterococci share all the characteristics of fecal streptococci except that they can also grow in a medium with 6.5-percent sodium chloride (NaCl) and can grow at 10°C. I've recapped all of our earlier Escherichia coli work because it sets the stage for why we are working with the enterococci. At the end of the abstract, I've listed the uniform resource locator (URL) for our ribotyping protocol and the six manuscripts associated with this work. Only two of the six manuscripts are published. Anyone is welcome to download the protocol and e-mail me (pghartel@arches.uga.edu) if they would like reprints or preprints of any of the manuscripts.

We began microbial source tracking research shortly after we first heard Mansour Samadpour (University of Washington) and George Simmons (VPI) talk about microbial source tracking at the 1998 American Society of Agronomy meetings in Baltimore. We were most interested in ribotyping because of its reproducibility and because we had access to a Qualicon RiboPrinter, an instrument that could do ribotyping. In our first manuscript (Hartel et al., 1999), we obtained 35 isolates of *E. coli* from cow manure in a pasture and compared these isolates to:

- 40 isolates of *E. coli* from a stream draining that pasture.
- 44 isolates from a geographically separated wooded stream remote from domesticated animals and human habitation.

The isolates were ribotyped with the restriction enzyme EcoRI. Among the 119 isolates, the RiboPrinter system identified eight different ribotypes. The wooded stream and the cow manure each contained six ribotypes, while the pasture stream contained all eight. In the wooded stream, 29 of the 44 isolates (66 percent) were Ribotypes #6 and #8, while in the cow manure, 19 isolates (54 percent) were Ribotype #4. The results suggested that either a limited variation in ribotypes of

Correspondence should be addressed to:

Peter G. Hartel, Ph.D.
Associate Professor, Department of Crop & Soil Sciences
University of George
3111 Plant Sciences Bldg., Athens, GA 30602
Tel: 706-542-0898 Fax: 706-542-0914 E-mail: pghartel@imap.arches.uga.edu

E. coli exists among the two streams and the cow manure, or that a more discriminatory ribotyping system was needed.We decided to try a more discriminatory ribotyping system. We called Mansour Samadpour and he kindly gave us a brief outline of his ribotyping method. This method uses two restriction enzymes, EcoRI and PvuII, and the additional bands give more discrimination. This method made the RiboPrinter too expensive to use, so all of this work was done manually. This is the method we have used with all our remaining *E. coli* work and is the protocol listed under the URL. Our next manuscript (Hill et al., 2001) was a small study determining the source of fecal contamination in three household wells near Cochran, Georgia. All the wells had tested positive for fecal coliforms. We quantified the number of fecal coliforms in water samples from one pond, two streams, a sinkhole, and three household wells and their accompanying septic systems. The pond, streams, and sinkhole were all connected. We isolated *E. coli* from the fecal coliforms and ribotyped 51 isolates. Twelve different ribotypes were observed among the water sources and the household wells with their accompanying septic systems. Two ribotype patterns were observed from the septic systems, 10 patterns among the pond, sinkhole, and two streams, and six patterns among the three household wells. At 100 percent similarity, the ribotype patterns of *E. coli* from the household wells were associated with six of the 10 ribotype patterns from the pond, sinkhole, and two streams. The similarity of *E. coli* ribotype patterns between the household wells and the septic systems was only 80 to 86 percent. The results suggest that the point source of the *E. coli* contamination was the pond, sinkhole, or two streams, and not the septic systems.

These results were sufficiently encouraging that the Georgia Environmental Protection Division wanted us to establish a host origin database for *E. coli* in Georgia. The reason for this was the potential integration of microbial source tracking with the State's total maximum daily load work. If this has not already been explained, a total maximum daily load is a calculation of the maximum amount of non-point contamination that a water body can receive and still meet water-quality standards. Although "contamination" can mean nitrogen, phosphorus, pesticides, and heavy metals among other things, about 70 percent of the Georgian waters that do not meet total maximum daily load standards do so because they exceed the limit for fecal coliforms (for recreational waters in Georgia, the standard is 200 fecal coliforms per 100 milliliter [mL]). The current method for identifying non-point fecal contamination, land use, does not work very well. The problem with creating the host origin database for Georgia was that we were worried that a geographically or temporally limited host origin database might not be universally applicable. Determining the geographic and temporal variability of *E. coli* ribotypes was too much work for us, so we decided to split the geographic work with Jim Entry (U.S. Department of Agriculture-ARS, Kimberly, Idaho). Mike Jenkins (U.S. Department of Agriculture-ARS, Watkinsville, Georgia) kindly volunteered to do the work on temporal variability.

Both manuscripts for geographic and temporal variability are finished. The temporal variability manuscript (Jenkins, M.B., P.G. Hartel, T.J. Olexa, and J.A. Stuedemann [2002]. "Temporal variability of Escherichia coli ribotypes from an agricultural source of fecal pollution.") is currently under U.S. Department of Agriculture review, and readers interested in this manuscript should contact Mike Jenkins directly at <mjenkins@arches.uga.edu>. In the geographic variability manuscript (Hartel et al., 2002a), we used ribotyping to determine the geographic variability of *E. coli* from one location in Idaho and three locations in Georgia for cattle, horse, swine, and chicken. A total of 568 fecal *E. coli* isolates from Kimberly, Idaho (125 isolates) and from Athens (210 isolates), Brunswick (102 isolates), and Tifton, Georgia (131 isolates) yielded 213 ribotypes. The percentage of ribotype sharing within an animal species increased with decreased distance between geographic locations for cattle and horses, but not for swine and chicken. When the *E. coli* ribotypes among the four host species were compared at one location, the percent of unshared ribotypes was 86, 89, 81, and 79 percent for Kimberly, Athens,

Brunswick, and Tifton, respectively. These data suggest that there is good ribotype separation among host animal species at each location. We concluded that ability to match environmental isolates to a host origin database might depend on a large number of environmental and host origin isolates that, ideally, are not geographically separated.

In the meantime, we were also concerned about other factors that affect ribotype variability within a specific host species. The main factors we thought important were age, diet, physiology (e.g., ruminants versus non-ruminants), and intimacy (e.g., humans and dogs sharing the same house). We decided to try diet first and determined the effect of diet on ribotype diversity for *E. coli* in penned and wild deer in a 13-hectare forested watershed (Hartel et al., 2002b). A total of 298 *E. coli* isolates was obtained, 100 from penned deer, 100 from wild deer, and 98 from the stream in the watershed. The penned deer had 11 ribotypes and wild deer had 35 ribotypes, and this difference was significant ($p = 0.05$). This suggests that diet affected ribotype diversity, and that a host origin database for microbial source tracking should contain bacterial isolates from wild rather than from captive animals. Little differences were observed between the ribotypes of *E. coli* from penned bucks and does. Also, 42 of 98 (42.9 percent) environmental isolates matched deer ribotypes. If microbial source tracking determines that fecal contamination is predominantly from wildlife, then it may be unnecessary to monitor these watersheds because control over wildlife is difficult, if not impossible.

If there was geographic and temporal variabilities for ribotypes of host origin isolates, then we should see a reflection of these variabilities in the ribotypes of environmental isolates, particularly during runoff conditions. This was also important because the more *E. coli* ribotype variability, the more sampling would be required to adequately describe the watershed. This next study was done in collaboration with Adrienne Funk, Betsy Frick, and Brian Gregory (U.S. Geological Survey, Atlanta, Georgia) (Hartel et al., 2002c). The study area was a 77-kilometer reach of the Chattahoochee River and its tributaries in metropolitan Atlanta. A total of 659 *E. coli* isolates was obtained from eight tributaries and four main stem sites during base flow and wet weather conditions, and 346 isolates were ribotyped and assessed for their similarity. During base flow conditions, 92 of 162 ribotypes were unique; during wet weather conditions, 57 of 86 ribotypes were unique. To be unique, a ribotype could only be observed once at one location during one flow condition. This suggests that geographic variability exists. When uniqueness was redefined as being observed in only one location during both base flow and wet weather conditions, 110 of the combined 149 unique ribotypes (73.8 percent) remained unique. This suggests that at least some of these unique ribotypes may have been introduced during wet weather conditions and that temporal variability may exist. The large number of unique ribotypes again suggests that considerable variability exists among *E. coli* ribotypes, and a large sampling of *E. coli* isolates is needed from watersheds with complex land use patterns and varied flow conditions for adequate microbial source tracking.

In the meantime, we were dissatisfied with using fecal coliforms as indicator organisms for all environmental conditions. Our research with bacterial survival in broiler litter suggested that fecal streptococci were better fecal indicators than fecal coliforms. This is not a trivial problem for Georgia because the State is ranked No. 1 in broiler production in the United States and, last year, produced over 12-million Mg of litter. Chuck Hagedorn (VPI) also observed that fecal streptococci were better indicators of fecal contamination than fecal coliforms for composted biosolids. At the same time, the Georgia Environmental Protection Division wanted us to prioritize the total maximum daily load implementation plans for fecal contamination. Ideally, those stream segments with human fecal contamination should be given first priority for cleanup because human fecal contamination represents the greatest hazard to humans, yet is a relatively easy non-point source to remedy. The Georgia Environmental Protection Division was,

ultimately, interested in a simple, inexpensive test to identify human fecal contamination that could be done in any microbiological laboratory. After combining this decades-old problem with our dissatisfaction using fecal coliforms as indicator organisms, we decided to look more closely at the fecal streptococci. This was the beginning of our interest in Enterococcus faecalis. The vast majority of literature suggests that this organism does not have a limited host range, but one paper by Pourcher et al. (1991) suggested that it did. On this basis, we attempted to repeat Pourcher's work (Wheeler et al., 2002). Of 583 fecal streptococcal isolates obtained on Enterococcosel agar from Canada goose, cattle, deer, dog, human, chicken, and swine, 392 were considered presumptive enterococci and were subsequently speciated with the API 20 Strep system. Of these isolates:

- 22 were *Ent. durans* (5.6 percent).
- 61 were *Ent. faecalis* (15.6 percent).
- 98 were *Ent. faecium* (25.0 percent).
- 86 were *Ent. gallinarum* (21.9 percent).
- 125 were unidentified (31.9 percent).

The host range of the *Ent. faecalis* isolates was limited to dogs, humans, and chickens, and our results were similar to those of Pourcher et al. (Table 1). We then developed new media to isolate and identify *Ent. faecalis* quickly from fecal samples; this scheme eliminated *Ent. faecalis* isolates from dogs. When the remaining *Ent. faecalis* isolates were ribotyped, it was possible to differentiate clearly among the isolates from human and chicken. It may be that combining the potentially limited host range of *Ent. faecalis* with ribotyping is useful for prioritizing watersheds with fecal contamination.

Table 1. Number of Ent. faecalis Isolates from Presumptive Streptococcal or Enterococcal Isolates Obtained from the Feces of Humans and Other Warm-blooded Animals

Source	Pourcher et al. (1991)		Wheeler et al. (2002)	
	Total isolates	*Ent. faecalis*	Total isolates	*Ent. faecalis*
Cattle	56	0	69	0
Chicken	42	8	35	12
Deer	Not Done	Not Done	131	0
Dog	Not Done	Not Done	56	30
Human	72	27	47	19
Horses	28	1	Not Done	Not Done
Rabbits	5	0	Not Done	Not Done
Sheep	12	0	Not Done	Not Done
Swine	45	3	48	0
Wild birds*	10	7	6	0

*Wild birds for Pourcher et al. (1991) were seagulls; wild birds for Wheeler et al. (2002) were Canada goose.
The identification of the host isolates was determined with the API 20 Strep system.

In a conversation several weeks ago with Fred Genthner (U.S. Environmental Protection Agency), we discussed the possibility that the initial isolation medium may be responsible for some of the limited host range of *Ent. faecalis*. We are currently testing this idea. Of 92 isolates from chicken excreta (each obtained on Enterococcosel, KF, mE [with esculin iron agar], and m*Enterococcus* media), 42, 36, 60, and 49 isolates were *Ent. faecalis*, respectively. Of the combined 368 isolates from the four media obtained from deer feces, none was *Ent. faecalis*. These results are what we would have predicted and suggest that the initial isolation medium is not responsible for the limited host range; however, because fecal contamination from humans is typically from sewage and broilers is from broiler litter, we are testing these sources to confirm that they contain *Ent. faecalis*. Of raw sewage samples from four different wastewater treatment plants, 110 of 637 isolates (17.3 percent) obtained on Enterococcosel agar were *Ent. faecalis*. Surprisingly, of 728 isolates obtained on Enterococcosel agar from four broiler litter samples, a majority (73.8 percent) were not enterococci, and of those that were enterococci, none was *Ent. faecalis*. This continues to suggest that *Ent. faecalis* may be a good indicator of human fecal contamination. Obviously, more animals need to be tested, not only of the same animal species already tested, but also new animal species (particularly wild animals). We have tested more dogs with the Wheeler et al. (2002) medium, and of 182 isolates obtained on Enterococcosel agar, 9 (4.9 percent) were *Ent. faecalis*. We are currently testing more sewage and broiler litter samples, and are preparing to field-test two stream segments, one urban and one rural, both of which exceed their total maximum daily loads for fecal contamination. If *Ent. faecalis* is found not at all or in reduced numbers in other animal species besides humans and wild birds, then this will seriously reduce the amount of host origin sampling needed to establish human fecal contamination.

Besides the proposed research already mentioned, we will also concentrate on two other research areas. First, because of the apparent variability of *E. coli* ribotypes, we badly need to automate ribotyping to handle an increased numbers of isolates. We don't know the resolution to this yet. Cost is a real problem. It currently costs us $22 to run each isolate manually, and the cost to run each isolate on something like a RiboPrinter (with no discount for volume) is more than double that. Second, regardless of the situation with *Ent. faecalis*, it seems reasonable to us that not all bacterial species are found in all animal hosts. We will continue to investigate the possibility of using other enterococcal species or ratios of various enterococcal species to determine sources of fecal contamination from specific animal species. There may be other bacteria with limited host ranges that are also useful in this regard.

URL for ribotyping protocol: <http://dmsylvia.ifas.ufl.edu/msp/other.htm>

References

Hartel, P.G., W.I. Segars, N. Stern, J. Steiner, and A. Buchan (1999). "Ribotyping to determine the host origin of *Escherichia coli* isolates in different water samples." Wildland hydrology, D. S. Olsen and J. P. Potyondy (eds.). American Water Resources Association Technical Publications Series TPS-99-3, Herndon, Virginia.

Hartel, P.G. J.D. Summer, J.L. Hill, J.V. Collins, J.A. Entry, and W.I. Segars (2002a). "Geographic variability of *Escherichia coli* ribotypes from animals in Idaho and Georgia." *Journal of Environmental Quality*, favorably reviewed.

Hartel, P.G., J.D. Summer, and W.I. Segars (2002b). "Effect of diet on ribotype diversity for microbial source tracking." *Water Research*, submitted.

Hartel, P.G., J.L. Hill, J.D. Summer, A.L. Funk, E.A. Frick, and M.B. Gregory (2002c). "Variability of *Escherichia coli* ribotypes in a complex watershed during varied flow conditions." Journal of Environmental Quality, under U.S. Geological Survey review.

Hill, J.L., P.G. Hartel, W.I. Segars, and P. Bush (2001). "Ribotyping to determine the source of fecal coliform contamination in three household wells near Cochran, Georgia." *Proceedings, 2001 Georgia Water Resources Conference*, K. J. Hatcher (ed.). University of Georgia, Athens, Georgia.

Pourcher, A.M., L.A. Devriese, J.F. Hernandez, and J.M. Delattre (1991). "Enumeration by a miniaturized method of *Escherichia coli*, *Streptococcus bovis*, and enterococci as indicators of the origin of fecal pollution of waters." *Journal of Applied Bacteriology*, 70: 525-530.

Wheeler, A.L., P.G. Hartel, D.G. Godfrey, J.L. Hill, and W.I. Segars (2002). "Potential of *Enterococcus faecalis* as a human fecal indicator for microbial source tracking." *Journal of Environmental Quality*, favorably reviewed.

PETER HARTEL is an Associate Professor in the Department of Crop & Soil Sciences at the University of Georgia, where he has taught since 1985. A descendant from a long line of farmers, his primary research interest is in soil microbiology. He is co-editor of a leading textbook on soil microbiology as well as books on environmental ethics and agricultural ethics. He currently administers the University of Georgia's internationally recognized Environmental Ethics Certificate Program. Hartel received a B.S. in Biology from the University of Alaska, Fairbanks, a B.S. in Microbiology from the University of Washington, and a Ph.D. in Microbiology from Oregon State University. He conducted post-doctoral work in the Laboratory for Soil Microbiology in the Department of Agronomy at Cornell University.

Fecal Source Tracking with *Bacteroides*

Katherine G. Field
Assistant Professor, Department of Microbiology
Oregon State University

Introduction

Fecal contamination of aquatic environments afflicts many regions of the United States, with associated human health risks and environmental damage. Often, the source of the fecal contamination cannot be determined. For example, failing septic systems, overloads at sewage treatment facilities, wildlife, domestic pets, and runoff from non-point sources (such as farm manure) may all be candidates. The standard methods of measuring fecal pollution do not distinguish between human and animal sources. A method of fecal source discrimination is critical for control. Ideally, such a method would:

- Be inexpensive, quick, and reliable.
- Not require culturing isolates.
- Not require a large library of reference strains.
- Be flexible for easy field handling and storage of samples.
- Require a minimum of specialized equipment.

Taking advantage of recent methodological advances in the field of molecular ecology of microorganisms, we have developed a method of fecal source discrimination based on host-specific genetic markers from the *Bacteroides-Prevotella* group of fecal anaerobic bacteria. This method appears to fit the list of desired characteristics. In particular, because the method does not require culturing bacterial isolates, it is less labor intensive and more rapid than many other approaches. Also, as the *Bacteroides-Prevotella* markers are widespread; they appear to be common to all members of a particular host group. This means that libraries of reference strains are not needed.

The shortcoming of the method is that, thus far, markers for only a few species have been developed, although markers for more species are currently being tested. In addition, like any indicator, it is important to understand how the survival of *Bacteroides-Prevotella* bacteria in water compares to the survival of human pathogens.

Bacteroides-Prevotella

Bacteroides is a group of non-spore-forming, obligate anaerobes that make up one-third of the

Correspondence should be addressed to:

Katharine G. Field, Ph.D.
Assistant Professor, Department of Microbiology
Oregon State University
220 Nash Hall, Corvallis, OR 97331
Tel: 541-737-1837 Fax: 541-737-0496 E-mail: fieldk@bcc.orst.edu

human fecal flora, far outnumbering coliforms (Holdeman et al., 1976; Moore and Holdeman, 1974; Salyers, 1984). *Bacteroides* and its close relatives in the genus Prevotella are found exclusively in feces, animal rumens, and other cavities within animals and humans (Paster et al., 1994). When found in water, they are invariably diagnostic of pollution. For several days after dispersal in water, *Bacteroides* has been identified from environmental water samples (Kreader, 1995; Straub and Dixon, 1997). It has been detected by polymerase chain reaction (PCR) for up to 14 days at 4°C, even in the presence of predators. At higher temperatures more typical of natural waters (14°C), *Bacteroides* has been detected for 4 to 5 days (Kreader, 1998).

The most abundant *Bacteroides* species in human feces are either human-specific or present only at very low levels in other species (Allsop and Stickler, 1985; Straub and Dixon, 1997). This group of bacteria has not been used as an indicator because of the difficulty of growing anaerobic bacteria. We chose *Bacteroides-Prevotella* to develop genetic markers because the group contains a great deal of sequence-level diversity; species of *Bacteroides* are as genetically diverse as genera of other groups, including coliforms (Abigail Salyers, personal communication). In addition, because they make up such a high proportion of the fecal flora, they are easier to detect by PCR directly from water without a growth step than the more rare coliforms.

PCR, LH-PCR, and T-RFLP

Any test for fecal pollution must first increase the indicators to a level sufficient for detection. Coliforms tests do this by growing the bacteria. To amplify our genetic markers to a detectable level, we use PCR, a method of making many copies of a target gene sequence in a test tube. In PCR, the desired gene sequences are "selected" for amplification by the use of "primers" specific for these sequences. Primers are short lengths of single-stranded DNA that can be synthesized in any sequence required. We chose to use ribosomal RNA genes because they are widely used in studies of diversity in bacteria, and sequences are readily available. We used *Bacteroides-Prevotella* sequences from sequence databases to design PCR primers that specifically amplify 16S ribosomal RNA genes from *Bacteroides-Prevotella* bacteria (Bernhard and Field, 2000a).

To establish diagnostic markers, we used recently developed technologies that discriminate among mixtures of bacterial gene sequences by detecting differences in the length of gene fragments (Avaniss-Aghajani et al., 1994; Brunk et al., 1996). Length heterogeneity polymerase chain reaction (LH-PCR) (Suzuki et al., 1998) and terminal restriction fragment length polymorphism (TRFLP) (Liu et al., 1997) analyze differences in lengths of gene fragments due to insertions and deletions and estimate the relative abundance of each fragment.

Identification of Bacteroides Genetic Markers

We used cow and human feces for our initial experiments (Bernhard and Field, 2000a) and looked for fragments that were found in one type of feces, but not the other. Both LH-PCR and TRFLP identified several fragments unique to either cow or human feces (Figure 1). Next, we asked whether we could recover these PCR fragments from contaminated water. We filtered sewage, river, and bay water samples, extracted DNAs from the filters, and amplified DNAs using the *Bacteroides* PCR primers. We observed the same fragments as those found in feces (data not shown). To test whether the fragments we found in the water were really fecal DNA markers, we cloned and sequenced the fecal markers and the fragments from the water samples. The fecal and water sequences fell into four groups of related sequences: two unique to human feces and two unique to cattle feces (data not shown). Using these sequences, we designed PCR primers specific to each group (Bernhard and Field, 2000b). We tested the PCR primers and found that the two human primers only amplified DNA from human feces (Figure 2). The cow

primers did not amplify human fecal DNA (Figure 3), but did amplify fecal DNAs from all ruminants (Bernhard and Field, 2000b). The sensitivity of detection of the *Bacteroides* markers

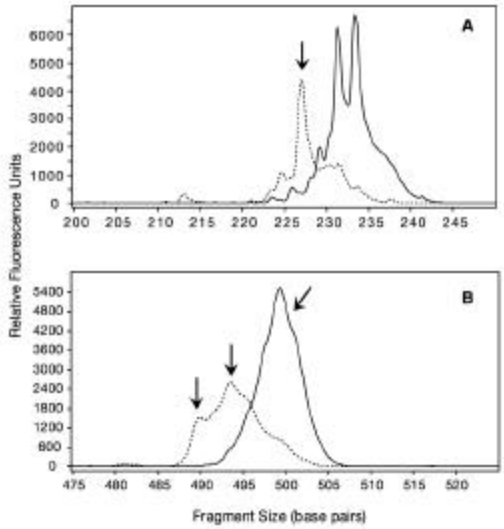

Figure 1. Example of TRFLP analysis of Bacteroides 16S rDNA fragments. Arrows indicate ruminant and human-specific markers. Fecal DNAs were amplified with PCR primers Bac32F-FAM and BAC708R and cut with restriction enzyme AciI. Solid lines are from human fecal DNA; dotted lines are from cow fecal DNA.

was comparable to the sensitivity of detection of fecal coliforms (Figure 4). We have since designed and tested primers specific for pigs, elks, and dogs/cats, and are developing primers for other groups, including seals, gulls, chickens, ducks/geese, and gulls.

Primers developed for Oregon fecal samples work for fecal samples from New York, Canada, and New Zealand. Sequence evidence suggests that the PCR primers sample entire species of bacteria, which are distributed throughout entire host species and groups of species, not just local populations. There is some evidence that the amount of markers present in an individual may vary; thus, this approach may be more reliable for detecting groups of individuals (for example, sewage or farm runoff) than single individuals.

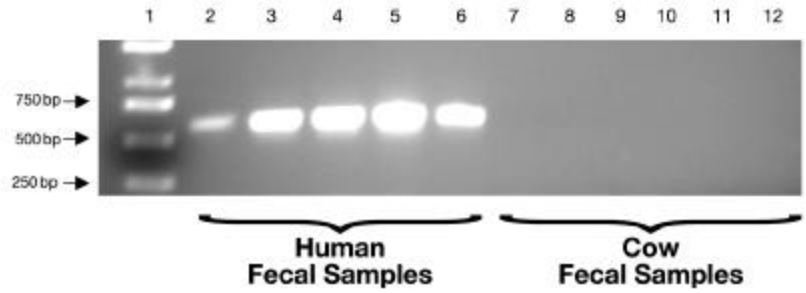

Figure 2. Specific amplification of a human fecal DNA marker with Bacteroides *human-specific 16S rDNA PCR primer HF10 and Bac32F. Bright bars are amplified DNAs separated on an agarose gel and stained with ethidium bromide. Ladder on left is a size standard.*

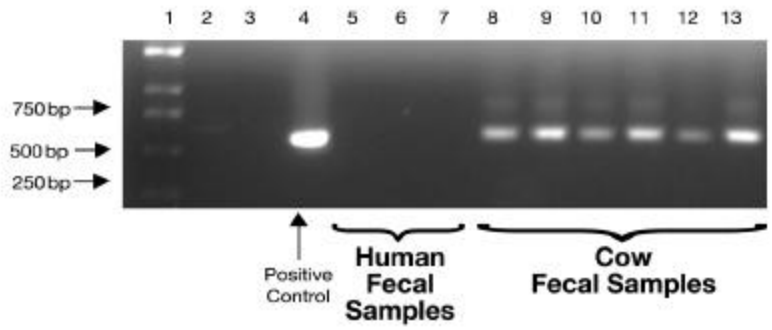

Figure 3. Specific amplification of a ruminant fecal DNA marker with Bacteroides *ruminant-specific 16S rDNA PCR primer CF151 and Bac32F. Bright bars are amplified DNAs separated on an agarose gel and stained with ethidium bromide. Ladder on left is a size standard.*

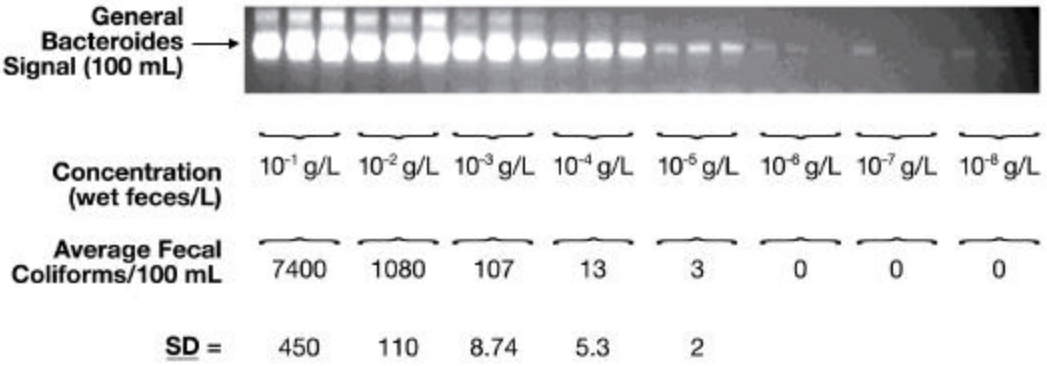

Figure 4. Detection of fecal contamination with PCR/gel analysis using general Bacteroides *primers compared to fecal coliform counts (CFU) by membrane filtration. Cow feces were mixed with water and diluted. Triplicate subsamples of each dilution were used for membrane filtration of fecal coliforms and DNA extraction followed by PCR detection. 1g wet fecal mass = 1.4 x 10^{-1} g dry fecal mass.*

The Procedure

Figure 5 shows the protocol for fecal source discrimination using the *Bacteroides-Prevotella* PCR primers. Specialized equipment includes PCR primers and reagents, a thermal cycler for PCR, and the gel apparatus and documentation device used to reveal the results of the PCRs. Also needed is equipment for water collection and filtration, a refrigerator and freezer, buffers, and DNA extraction kits. The entire procedure lends itself to automation and high-throughput. Because the analysis does not require living cells, sample handling and processing times are simplified. The analysis can be finished in less than 24 hours. After water samples are filtered, the filters can be stored in a preservative buffer for an indefinite length of time before analysis. Quality assurance requires including multiple negative and positive controls in every set of analyses. Cost, if done in bulk, could be as low as $25 to $35 per sample analyzed with two species-specific primers.

Conclusions

Amplification of genetic markers from *Bacteroides* provides a sensitive and accurate method of fecal source identification without growing bacteria. In most cases, markers appear to have co-evolved with host animals, so they are widely distributed. The ability to accurately and sensitively

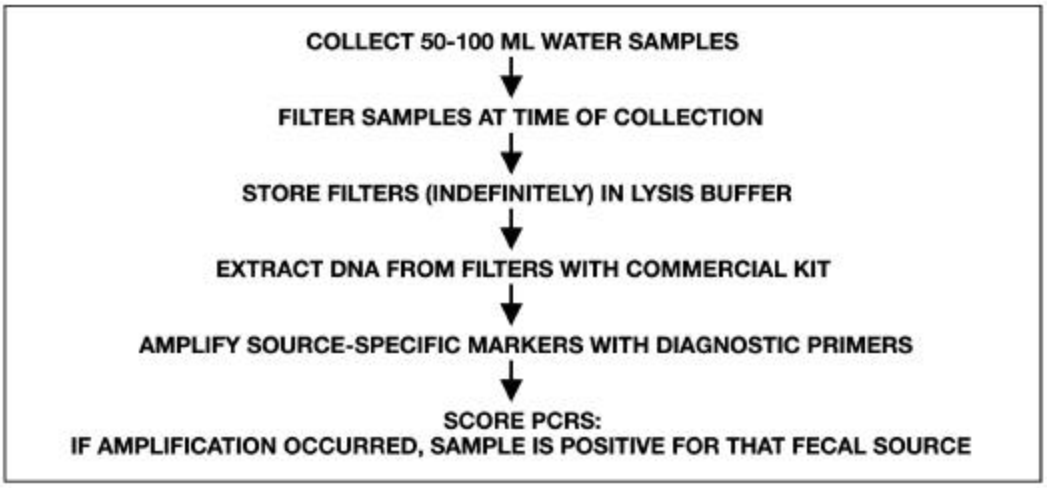

Figure 5. Protocol for source detection with fecal-source-specific Bacteroides-Prevotella *primers.*

track a single marker sequence or indicator in a complex environment has widespread applications in molecular ecology.

References

Allsop, K., and J. D. Stickler (1985). "An assessment of *Bacteroides* fragilis group organisms as indicators of human faecal pollution." *Journal of Applied Bacteriology*, 58: 95-99.

Avaniss-Aghajani, E.A., K. Jones, D. Chapman, and C. Brunk (1994). "A molecular technique for identification of bacteria using small subunit ribosomal RNA sequences." *BioTechniques*, 17: 144-149.

Bernhard, A.E., and K.G. Field (2000a). "Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from fecal anaerobes." *Applied and Environmental Microbiology*, 66: 1,587-1,594.

Bernhard, A.E., and K.G. Field (2000b). "A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA." *Applied and Environmental Microbiology*, 66: 4,571-4,574.

Brunk, C.F., E. Avaniss-Aghajani, and C.A. Brunk (1996). "A computer analysis of primer and probe hybridization potential with bacterial small-subunit rRNA sequences." *Applied and Environmental Microbiology*, 62: 872-879.

Holdeman, V., E.P. Cato, and W.E.C. Moore (1976). "Human fecal flora: Variation in bacterial composition within individuals and a possible effect of emotional stress." *Applied and Environmental Microbiology*, 31: 359-375.

Kreader, C.A. (1995). "Design and evaluation of *Bacteroides* DNA probes for the specific detection of human fecal pollution." *Applied and Environmental Microbiology*, 61: 1,171-1,179.

Kreader, C.A. (1998). "Persistence of PCR-detectable *Bacteroides* distasonis from human feces in river water." *Applied and Environmental Microbiology*, 64: 4,103-4,105.

Liu, W.T., T.L. Marsh, H. Cheng, and L.J. Forney (1997). "Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA." *Applied and Environmental Microbiology*, 63: 4,516-4,522.

Moore, W.E.C., and L.V. Holdeman (1974). "Human fecal flora: The normal flora of 20 Japanese-Hawaiians." Applied Microbiology, 27: 961-979.

Paster, B.J., F.E. Dewhirst, I. Olsen, and G.J. Fraser (1994). "Phylogeny of *Bacteroides*, Prevotella, and Porphyromonas spp. and related bacteria." *Journal of Bacteriology*, 176: 725-732.

Salyers, A.A. (1984). "*Bacteroides* of the human lower intestinal tract." *Annual Reviews of Microbiology*, 38: 293-313.

Straub, D.V., and B. Dixon (1997). "*Bacteroides* vulgatus, an alternative indicator for the assessment of fecal contamination of shellfish and estuarine waters." *Proceedings, 38th Annual Western Fish Disease Workshop*, Bodega Marine Laboratory, Bodega Bay, California.

Suzuki, M., M.S. Rappe, and S.J. Giovannoni (1998). "Kinetic bias in estimation of coastal picoplankton community structure obtained by measurements of small subunit rRNA gene PCR amplicon length heterogeneity." *Applied and Environmental Microbiology*, 64: 4,522-4,529.

KATHERINE FIELD has taught at Oregon State University since 1988. Currently, she is Assistant Professor in the Department of Microbiology as well as Co-Director of the Bioresource Research Program at Oregon State University. In her research, Field applies molecular methods to environmental problems, and studies microbial ecology and evolution. She is interested in bacterial source discrimination and in the spread of antibiotic resistance genes from fecal pollution. She also holds a patent in "Source Detection of Fecal Pollution Using Host-specific 16S rDNA Genetic Markers from Fecal Anaerobic Bacteria." Field received a B.A. in Biology from Yale University, an M.A. in Biology from Boston University, and a Ph.D. in Biology from the University of Oregon. She also obtained post-doctoral training in Molecular Biology at Indiana University.

Urbanization and Coastal Water Quality: What Can Molecular Fingerprinting Tell Us?

Michael G. LaMontagne and Patricia A. Holden
The Donald Bren School of Environmental Science & Management
University of California, Santa Barbara

Introduction

Human development of coastlines changes the quantity and quality of freshwater input to coastal waters. Overland flow of stormwater runoff and subsurface groundwater transport can alter watershed waters chemically and biologically, and both types of changes may affect the phylogenetic diversity and composition of coastal creek aquatic bacterial communities. We ask:

- What is the relationship between human development and bacterial communities for two coastal creeks in Santa Barbara, California?
- Is the coastal lagoon for one creek harboring a bacterial community distinct from the creek upstream or downstream?

We characterized the aquatic bacterial community fingerprints for two watersheds in Santa Barbara, California (Mission Creek and Arroyo Burro Creek), and related these fingerprints to other indicators of urban water pollution. The goal is to understand the use of DNA fingerprinting as a potential source tracking tool for urban pollution in watersheds.

Background

Detecting small changes in aquatic communities using clone libraries (Acinas et al., 1999; Acinas et al., 1997; Benlloch et al., 1995; Moyer et al., 1994; Murray et al., 1998) is laborious and introduces a potential bias of cloning efficiency (Lueders and Friedrich, 2000; Wintzingerode et al., 1997). Alternatively, polymerase chain reaction (PCR) amplification products are sorted to detect changes by denaturing gradient gel electrophoresis (Muyzer et al., 1996) or terminal restriction fragment length polymorphism (TRFLP) (Liu et al., 1997). TRFLP is suitable for detecting small differences between similar microbial communities because of the reproducibility of profiles (Osborn et al., 2000) and appears more sensitive than denaturing gradient gel electrophoresis (Moeseneder et al., 1999). For these technical reasons, we have chosen to use TRFLP as the fingerprinting method of choice.

Methods

TRFLPs of PCR-amplified 16S ribosomal genes (PCR-TRFLP) (Liu et al., 1997) were analyzed for bacterial communities in water samples. PCR-TRFLP utilizes a fluorescently labeled primer

Correspondence should be addressed to:

Patricia Holden, Ph.D.
Assistant Professor, Donald Bren School of Environmental Science & Management
University of California, Santa Barbara
4670 Physical Sciences North
Santa Barbara, CA 93106-5131
Tel: 805-893-3195 Fax: 805-893-7612 E-mail: holden@bren.ucsb.edu

in the amplification of genes from a mixed template pool (in this case, environmental DNA). The resulting end-labeled PCR products are digested with suitable restriction enzymes to generate labeled fragments. The fragments can be separated and detected with a DNA sequencer. Based on the assumption that a unique fragment represents a unique sequence, patterns can be compared to detect changes in communities or to estimate the diversity of targeted genes, as well as be compared to predicted fragment lengths to infer community composition.

Water samples (1 liter) were taken in sterile sampling bottles and placed on ice during transport to the laboratory. Sub-samples (100 milliliters [mL]) were transferred onsite into a standard vessel containing thiosulfate preservative for indicator organisms. The sub-samples were delivered to Santa Barbara County and City laboratories for indicator organism analysis. The remaining 900 mL was pre-filtered through Nuclepore¨ 3.0-micrometer (µm) polycarbonate membranes. Bacterioplankton were collected from the pre-filtrate by vacuum-filtering through 0.2-µm SterivexTM GV filter cartridges (Millipore, Bedford, Massachusetts). Bacteria in the filter housing were lysed with lysozyme/sodium dodecyl sulfate (Murray et al., 1998). To minimize humic acid contamination, nucleic acids were purified by size exclusion spin columns (Sepharose G75 overlayed with polyvinylpolypryrrolidone) and concentrated by polyethylene glycol (molecular weight 8,000) precipitation.

16S rDNA from purified DNA samples was PCR-amplified using universal eubacterial primers 8F hex (flourescently labeled forward primer; 5'AGAGTTTGATCCTGGCTCAG; [Liu et al., 1997]) and 1389R (5'ACGGGCGGTGTGTACAAG; [Osborn et al., 2000]). PCR products were purified with the High PureTM kit (Boehringer Mannheim, Indianapolis, Indiana), and digested with either *Hha*I or *Rsa*I. TRFLP patterns were used to estimate bacterial diversity and community structure. Shannon-Weaver diversity indices (Atlas and Bartha, 1993) were calculated using the number of peaks as an estimate of the number of species. TRFLP patterns were compared between communities using the Ribosomal Database Project TRFLP Profile Matrix program (Maidak et al., 1997) to calculate percent similarities. Principle components analyses were performed to assess the significance of overall similarity between bacterial communities at different sites and at different sampling times.

Results

We applied our methodologies as described above to a preliminary study of Arroyo Burro Creek and Mission Creek during the spring of 2001. We sampled Arroyo Burro Creek and Mission Creek waters at several locations extending from the ocean, then upstream into the more pristine reaches (Mission Creek only). With assistance from the City of Santa Barbara, we also sampled sanitary sewage and El Estero wastewater treatment plant influent. The goal of this study was to determine, for this limited sample set, the degree to which microbial communities from sewage and creeks/ocean were similar. Similarity would directly imply influence of human waste on water quality. Both the City and County of Santa Barbara assisted by performing all of the indicator organism assays. Principle components analyses suggested a gradient in bacterial community composition from upstream to downstream in the creeks (Figure 1). When plotted against the water-quality metric of an enterococcus concentration, it would appear that the whole bacterial community is a reasonable tracer for this indicator (Figure 2) and, thus, is potentially relatable to anthropogenic effects, including pathogenic bacteria and, possibly, nutrients.

Bacterial diversity appeared highest in the "human impacted" stretches of the watersheds, which were downstream of a relatively pristine site (Tunnel Road on Mission Creek) and upstream from the coastal receiving waters (Figure 3). Coastal waters had relatively low diversity, both in terms of richness (number of terminal restriction fragments) and Shannon-Weaver indices. Different

restriction enzymes (*Hha*I and *Rsa*I) showed similar trends, but estimates obtained with RsaI were consistently lower than those obtained with *Hha*I.

We also analyzed microbial diversity in water collected in the human-impacted coastal Arroyo Burro lagoon by TRFLPs of PCR-amplified 16S rDNA. Water samples were fractionated by pre-filtration to differentiate particle-associated and free-living microbes. Pre-filtration removed 23 to 44 percent of bacteria, as assessed by direct counts and Most Probable Number, and 99 percent of phytoplankton, as assessed by chlorophyll a. Analysis of TRFLPs obtained by restriction with HhaI revealed that species richness and evenness were higher in the particle-associated fractions than in the free-living fractions. Restriction with *Rsa*I yielded fewer peaks than *Hha*I and did not reveal consistent differences in diversity between water fractions. Diversity, estimated with either enzyme, was higher in the lagoon than adjoining coastal waters. The variability between water fractions was greater in samples collected during the rainy season.

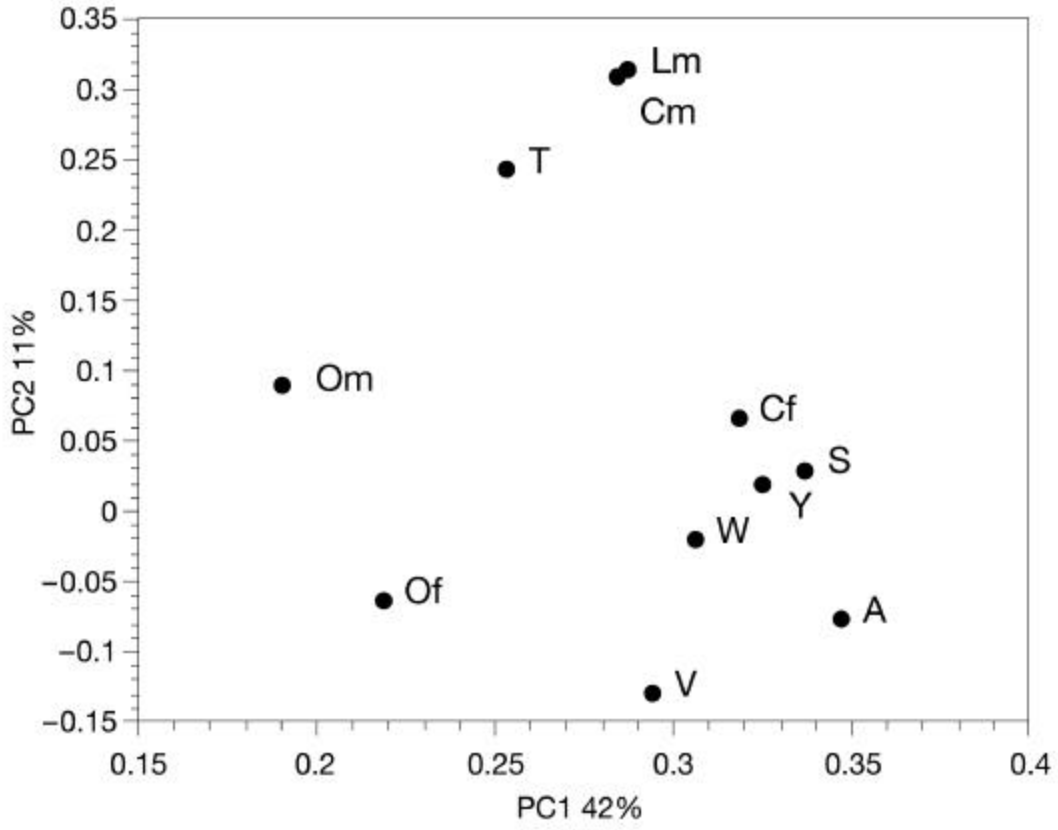

Figure 1. Plot of two principle components explaining 11 percent (y axis) and 42 percent (x axis) of the variation between bacterial communities (based on HhaI-generated TRFs) at sampling sites on Arroyo Burro Creek and Mission Creek watersheds and the ocean in May 2001. A development gradient is apparent in the creek samples, suggesting that urbanization affects bacterial communities. Legend: O(ocean); C (Arroyo Burro Creek just upstream of Lagoon); T (Tunnel Road – Headwaters of Mission Creek); L (Arroyo Burro Lagoon); W (West Mission); V (Victoria); Y (YMCA on Arroyo Burro Creek); A (Amtrack Station at Mission Creek). Subscripts "m" and "f" are May 2001 and February 2001, respectively. The ocean and lagoon at Mission Creek were not sampled. Tunnel (Mission Creek) is the only pristine site for the May 2001 study.

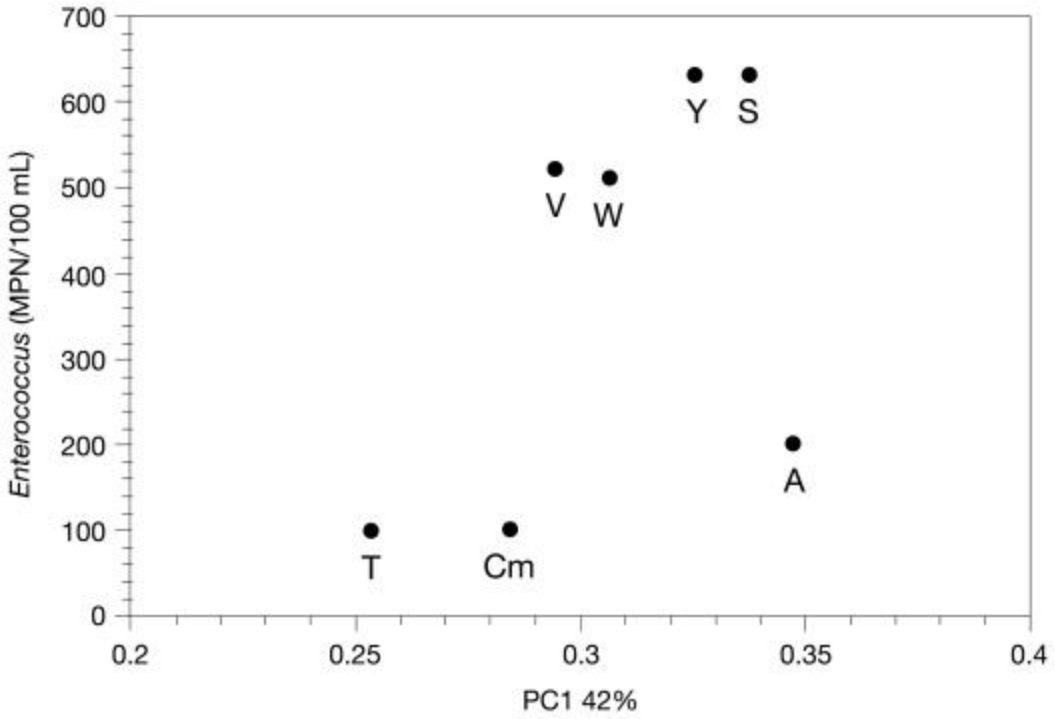

Figure 2. Entercoccus concentration versus the principle component (PC) explaining most (42 percent of the variability in bacterial community composition (based on Hhal-generated terminal restriction fragments) for Arroyo Burro Creek and Mission Creek in May 2001. By plotting a statistical synopsis of the bacterial community composition (i.e., PC) against an ecological variable, we can evaluate the degree to which bacterial communities are tracing either bacterial indicators of human waste or other pollutants. Legend as per Figure 1.

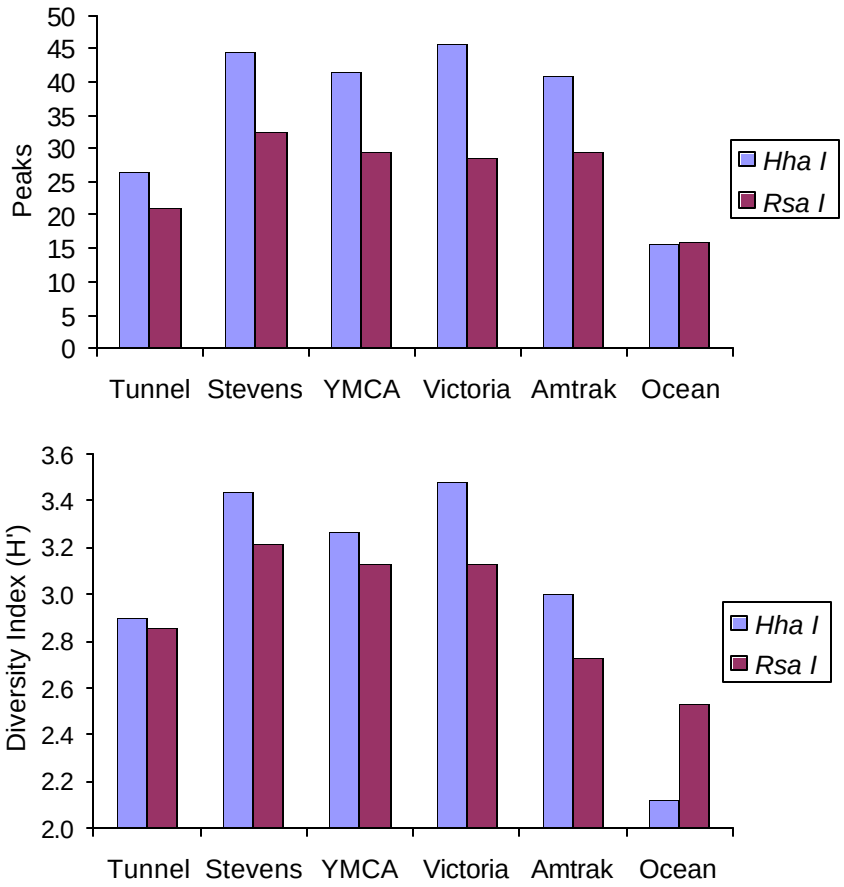

Figure 3. Aquatic bacterial diversity, as assessed by analysis of TRFLPs generated with either HhaI or RsaI, in Mission Creek and Arroyo Burro Creek watersheds during spring 2001 and a specified sampling sites. Top graph is species richness (number of terminal restriction fragments) versus location. Bottom graph is of the diversity index, where richness is weighted by evenness versus location.

Conclusions

By TRFLP analysis of 16S rDNA amplified from creek, lagoon, and ocean samples, we determined that:

- A development gradient in bacterial community composition appears in creek water.
- The lagoon and creek bacterial communities were similar in wet weather.

Our work suggests that fingerprinting whole bacterial communities is a sensitive, accurate, and reproducible approach for relating urbanization to watershed and coastal water quality.

References

Acinas, S.G., J. Anton, and F. Rodriguez-Valera (1999). "Diversity of free-living and attached bacteria in offshore western Mediterranean waters as depicted by analysis of genes encoding 16S rRNA." *Applied and Environmental Microbiology*, 65(2): 514-522.

Acinas, S.G., F. Rodriguez-Valera, and C. Pedros-Alio (1997). "Spatial and temporal variation in marine bacterioplankton diversity as shown by RFLP fingerprinting of PCR amplified 16S rDNA." *FEMS Microbiology Ecology*, 24(1): 27-40.

Atlas, R.M., and R. Bartha (1993). *Microbial ecology: Fundamentals and applications*, 3 ed. Benjamin Cummings, Redwood City, California.

Benlloch, S., F. Rodriguez-Valera, and A.J. Martinez-Murcia (1995). "Bacterial diversity in two coastal lagoons deduced from 16S rDNA PCR amplification and partial sequencing." *FEMS Microbiology Ecology*, 18(4): 267-279.

Liu, W.T., T.L. Marsh, H. Cheng, and L.J. Forney (1997). "Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA." *Applied and Environmental Microbiology*, 63(11): 4,516-4,522.

Lueders, T., and M. Friedrich (2000). "Archaeal population dynamics during sequential reduction processes in rice field soil." *Applied and Environmental Microbiology*, 66(7): 2,732-2,742.

Maidak, B.L., G.J. Olsen, N. Larsen, R. Overbeek, M.J. McCaughey, and C.R. Woese (1997). "The RDP (Ribosomal Database Project)." *Nucleic Acids Research*, 25(1): 109-110.

Moeseneder, M.M., J.M. Arrieta, G. Muyzer, C. Winter, and G.J. Herndl (1999). "Optimization of terminal-restriction fragment length polymorphism analysis for complex marine bacterioplankton communities and comparison with denaturing gradient gel electrophoresis." *Applied and Environmental Microbiology*, 65(8): 3,518-3,525.

Moyer, C.L., F.C. Dobbs, and D.M. Karl (1994). "Estimation of diversity and community structure through restriction fragment length polymorphism distribution analysis of bacterial 16S rRNA genes from a microbial mat at an active, hydrothermal vent system, Loihi Seamount, Hawaii." *Applied and Environmental Microbiology*, 60(3): 871-879.

Murray, A.E., C.M. Preston, R. Massana, L.T. Taylor, A. Blakis, K. Wu, and E.F. DeLong

(1998). "Seasonal and spatial variability of bacterial and archaeal assemblages in the coastal waters near Anvers Island, Antarctica." *Applied and Environmental Microbiology*, 64: 2,585-2,595.

Muyzer, G., T. Brinkhoff, U. Nubel, C. Santegoeds, H. Schafer, and C. Wawer (1996). "Denaturing gradient gel electrophoresis (DGGE) in microbial ecology." *Molecular microbial ecology manual*, A. D. L. Akkermans, J. D. Van Elsas, and F. J. de Bruijn (eds.). Kluwer Academic Publishers, Dordrecht; Boston, Massachusetts.

Osborn, A.M., E.R.B. Moore, and K.N. Timmis (2000). "An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics." *Environmental Microbiology*, 2(1): 39-50.

Wintzingerode, F.V., U.B. Goebel, and E. Stackebrandt (1997). "Determination of microbial diversity in environmental samples: Pitfalls of PCR-based rRNA analysis." *FEMS Microbiology Reviews*, 21(3): 213-229.

PATRICIA HOLDEN is a licensed Professional Engineer with 8 years professional design experience. Currently, she is Assistant Professor of Environmental Microbiology in The Donald Bren School of Environmental Science & Management at the University of California, Santa Barbara, where she has taught since 1997. Holden's research program focuses on environmental bacteriology, addressing: 1) whole community, phylogenetically based fingerprinting as a metric for water quality, and understanding the fate of indicators and pathogens in the coastal environment; and 2) biodegradation of petroleum hydrocarbons by vadose zone bacteria with emphasis on biofilm and its general presence and ecology in unsaturated systems. Holden received a B.S. in Civil Engineering from the University of Tennessee in Knoxville, an M.S. in Civil Engineering from Purdue University, and both a Master's of Engineering in Civil Engineering and Ph.D. in Soil Microbiology from the University of California, Berkeley.

Comparison of Environmental and Clinical Isolates of *Escherichia Coli* Using Various Genetic Fingerprinting Methods

Cindy H. Nakatsu, Byoung-Kwon Hahm, Jennifer Fiser,
Steven Yeary, and Arun Bhunia
Purdue University

Among the many microorganisms introduced into the environment from sources like waste treatment systems and agricultural runoff, coliform bacteria and, in particular, *Escherichia coli* have been used extensively by authorities as an indicator of water quality. The accumulation of materials down a watershed often makes it difficult to accurately determine the main source of a microorganism. The challenge ahead of us now lies in improving technologies capable of objectively identifying microbial sources to effectively control their introduction into our water supply. Molecular (DNA) fingerprinting methods are powerful tools for identifying and differentiating microbial strains. There are a number of genetic fingerprinting methods; our present objective is to determine the method that will most accurately identify sources of *E. coli* in the environment. Additionally, we wish to use a method that minimizes cost, labor, and time. We are comparing methods that examine the *E. coli* genome indirectly using polymerase chain reaction (PCR) amplification (amplified fragment length polymorphism, rep-PCR) and, to a lesser extent, direct genome analysis (pulsed-field gel electrophoresis and ribotyping). Isolates of *E. coli* tested were collected from:

- Animal fecal samples.
- Water from the Hoagland Ditch and Honey Creek watersheds in Indiana.
- Known-pathogenic strains (mainly serotype O157:H7) from food and clinical samples.

This variety of *E. coli* isolates is being used to determine the level of discrimination that can be made through the various genetic fingerprinting methods.

We have used rep-PCR to examine the majority of our isolates thus far because of the relative ease of the procedure. The rep-PCR DNA fingerprinting methods amplify portions of the genome using primers based on highly repeated sequences found throughout the *E. coli* genome. There are a number of different repeated sequences found in *E. coli*, and we have chosen primers from two groups to test in our analyses: Repetitive Extragenic Palindromes (REP) and BOX elements. We are also testing amplified fragment length polymorphism because of the potential for greater resolution and automation of fingerprint analysis. This technique combines restriction enzyme digestion and PCR. Briefly, after restriction enzymes digest the genomic DNA, linkers are ligated to the fragments, amplified by PCR using primers that complement the linker, and then a subset of the fragments are amplified using selective primers that have one to three additional bases. In theory, by using all the combinations of selective primers, the entire genome can be examined;

Correspondence should be addressed to:

Cindy H. Nakatsu, Ph.D.
Associate Professor, Department of Agronomy
Lilly Hall of Life Sciences, Purdue University
West Lafayette, IN 47907-1150
Tel: 765-496-2997 Fax:: 765-496-2926 E-mail: cnakatsu@purdue.edu

however, under most circumstances, the objective is to determine the primer combination that discriminates isolates at the desired level of resolution.

Results from all three of these PCR-based methods showed that complex fingerprint patternswere generated, which facilitated discrimination between isolates. In general, there were more bands in fingerprints generated using BOX A1-R versus REP-IR-1/2-1 primers. The number of fragments analyzed using amplified fragment length polymorphism could be changed by the combination of selective primer sets chosen.

We are presently using the following six selective primer set combinations:

- *EcoRI*-A+M*se*I-CA.
- *EcoRI* I-A+M*se*I-CT.
- *EcoRI* -AG+M*se*I-T.
- *EcoRI* -AG+M*se*I-C.
- *EcoRI* -C+M*se*I-CC.
- *EcoRI* -AT+M*se*I-T.

An initial study comparing 54 fingerprint patterns showed that all three methods were able to differentiate between the O157:H7 isolates, other pathogenic serovars, and some environmental isolates. The fingerprint profiles of O157:H7 isolates were almost identical, suggesting that this pathogen could be identified using any of these methods. In some cases, fingerprint patterns from environmental isolates appeared to group according to the animal sources, but the number of profiles was limited; therefore, we have begun to examine animal sources more extensively.

Isolates from feces collected from individual animals are being used to determine the range of genotypic variation in the *E. coli* population between and within an animal host. Our initial collections have been from pigs raised in three different facilities in Indiana to limit potential geographic variation. Between 10 and 15 *E. coli* isolates were randomly chosen from each pig, and a total of five pigs were tested from each farm. Our preliminary results from genetic fingerprint analyses showed at least five different groups using REP primers; however, one fingerprint group dominated the collection, representing about 72 percent of the isolates. The number of REP-fingerprint groups observed from isolates from a single animal ranged from one to four. The results indicate that a better understanding of genetic variation between *E. coli* isolates will improve our efforts to identify their sources into the environment.

This set of 54 isolates was also examined using pulsed-field gel electrophoresis. Pulsed-field gel electrophoresis separates large DNA molecules (megabase range) in an agarose gel by changing the orientation of the electric field at different intervals. The Center for Disease Control has been using this method to identify pathogenic *E. coli* (including O157:H7) from disease outbreaks associated with food and water. In our analysis, pulsed-field gel electrophoresis grouped the O157:H7 and, additionally, the fingerprint patterns of isolates from different outbreaks could be distinguished. Although the pulsed-field gel electrophoresis technique has been useful in distinguishing sources of disease outbreaks, it is time consuming and may not be suitable for rapidly identifying a large number of isolates.

We also analyzed ribotyping, another approach used by a number of research groups. The genetic

fingerprint for ribotyping is generated after hybridizing an *E. coli* rRNA gene probe to genomic DNA that has been digested with a restriction enzyme that cuts at a moderate frequency. We have been using the RiboPrinter¨ (Qualicon, Inc.), an automated system that can perform all the required steps overnight from cultured cells. The method minimizes the amount of labor needed, but the number of samples is limited to six per run and the reagents are costly. Ribotyping has been beneficial in our studies for differentiating *E. coli* strains from non-*E. coli* strains when the strains were misidentified by the cultivation/biochemistry tests used. This method could also differentiate the pathogenic *E. coli* from the environmental isolates. Because of the limited number of samples we tested, we presently cannot conclude whether ribotyping is able to differentiate animal sources.

The overall goal of our research is to determine whether molecular fingerprinting can be used to determine *E. coli* sources in the environment. To achieve this objective, we have been testing watershed isolates collected as part of the Purdue Lake Shafer Research Project, which is in the progress of identifying possible types of land use leading to increased *E. coli* levels in surface waters. Twenty-one sampling sites are being monitored along the Hoagland Ditch and Honey Creek watersheds that feed into Lake Shafer. Both agricultural and urban landscapes are found in the watersheds. For most of the year, *E. coli* concentrations in these waters are above the State of Indiana limit for full body contact of 235 colony-forming units (CFU) per 100 milliliter (mL) of water. Compared to upstream locations, higher numbers of *E. coli* have been found in areas downstream of some of the livestock and wastewater treatment facilities at various times throughout the year. Comparisons of rep-PCR fingerprint patterns of isolates collected over a year in these watersheds show patterns that can be found throughout the watershed. There were three to seven REP-fingerprint groups that dominated the samples. It is difficult to conclude possible sources of *E. coli* at various locations within the watershed because of the limited number of isolates that were purified from each site on each sampling date. This is an important factor based on the genotypic variations that we have observed from a single animal source.

To improve water quality where *E. coli* has been detected, the sources of the *E. coli* must be identified. Each of the methods that we have tested can be used to determine *E. coli* sources, although the methods vary in level of discrimination between isolates. Important factors that should be considered when monitoring point and non-point sources contamination into water bodies are the frequency and number of samples collected. Other valid factors to consider are the cost of supplies, availability of equipment, and availability and expertise of personnel to perform these approaches.

CINDY NAKATSU joined the Agronomy Department at Purdue University in 1995 after completing post-doctoral research in the Center for Microbial Ecology at Michigan State University. Currently, she is Associate Professor at Purdue University, teaching courses in soil ecology and molecular microbial ecology. Her recent research focus is on both basic microbiology and molecular biology techniques that are used to gain a better understanding of the adaptation of environmental microorganisms to perturbed ecosystems. Ultimately, she hopes to determine the different mechanisms involved in bacterial adaptation and survival in the environment. Nakatsu received both a B.Sc. in Biology and an M.Sc. in Botany from the University of Toronto, and a Ph.D in Environmental Microbiology from Carleton University in Ottawa, Canada.

Bacterial Endemism and Co-speciation

James T. Staley
Professor, Department of Microbiology
University of Washington

Endemic organisms are defined as those that live in only one area on Earth. Endemism in plants and animals is commonplace. These are the plants and animals that are most likely threatened with extinction because:

- If they are over-harvested or over-hunted, they cannot be replenished from another source.
- Their habitats can be threatened by increased human activities, such as farming and natural activities.
- They are susceptible to the introduction of invasive species that may displace them from their habitats.

Although some have made claims of bacterial endemism (Fulthorpe et al., 1998), no convincing evidence supports endemism in free-living bacteria; hence, there are no free-living bacterial species that are known to be endangered. In contrast, ample evidence supports the view that at least some bacterial species of both the Archaea and Bacteria have a worldwide or cosmopolitan distribution. For example, the archaeon, Archaeoglobus fulgidus, which was originally isolated from Italian thermal sources, has also been found in Alaskan oil wells as well as at other locations (Beeder et al., 1994).

One of the reasons that bacterial species may be more cosmopolitan than endemic is their broad species definition (Staley, 1997, 1999). When compared on a molecular basis (that is, mol% G + C content, range in 16S and 18S rDNA sequence, and DNA-DNA reassociation), the definition of *E. coli* is more like that of its human host family or order rather than its host species, *Homo sapiens* (Table 1). Since the DNA-DNA reassociation value of >70 percent is used as the definition of all bacterial species, what applies to *E. coli* applies to all bacteria.

Table 1. Comparison of E. coli and its Primate Host Species[1]

Property	*E. coli*	*Homo sapiens*	Primates
Mol% G + C	48 to 52	42	42[2]
16S-18S rRNA variability	>15 bases	?	<16[3]
DNA/DNA reassociation	>70%	98.6%[4]	70%[5]

[1]Adapted from J.T. Staley (1999).
[2]Value for all primates.
[3]Mouse 18S rRNA differs from humans by 16 bases.
[4]Comparison between *Homo sapiens* and chimpanzee.
[5]Comparison between *Homo sapiens* and lemur.

Correspondence should be addressed to:

James T. Staley, Ph.D.
Professor, Department of Microbiology and Immunology, SC-42
University of Washington, Seattle, WA 98195
Tel: 206-543-0461 Fax: 206-543-8297 E-mail: jtstaley@u.washington.edu

We have proposed a set of guidelines to determine whether a bacterium is endemic (Table 2) (Staley and Gosink, 1999). The basic assumption of these postulates is that if an organism is endemic to one area on Earth, the organism should have evolved independently from other related organisms for a sufficient length of time that the divergence between the organisms from two different sites is recorded in the sequence of some of the genes. These guidelines state that four strains of a purported "species" should be isolated from two different geographic areas. Then the sequence of at least two of the genes should indicate that separate clades of the organism exist at the two different locations.

Recently, we have tested the postulates using Cycloclasticus, a genus of marine bacteria that carries out the degradation of polycyclic aromatic hydrocarbon compounds (Hedlund and Staley, submitted). Strains of Cycloclasticus, which are members of the gamma-Proteobacteria, were isolated from the Pacific Ocean at Puget Sound (Washington) and in the Atlantic Ocean at the Gulf of Mexico (off Texas). The sequence of two genes, 16S rDNA and DNA gyrase B, were determined in a collection of eight strains from each location. Results indicate that the strains from Washington State cannot be distinguished from the strains from Texas, providing yet another example of a purported cosmopolitan bacterium.

Although examples of endemism of free-living bacteria have not been confirmed using these guidelines, evidence does exist that some symbiotic bacteria have co-speciated with their hosts. In this regard, we have studied the genus Simonsiella that lives in the oral cavity of mammals. These filamentous bacteria exhibit dorsal-ventral asymmetry. Sixteen strains of Simonsiella, which are members of the beta-Proteobacteria, were included in a study in which the 16S rDNA was sequenced (Hedlund and Staley, in press). Four strains were from human hosts, four from sheep, four from dogs, and four from cats. The phylogenetic tree indicates that the four strains from each of the animal hosts form a separate clade. This is exactly the result one would expect for co-speciation; however, it would be desirable to know if this speciation process is identifiable between strains of Simonsiella from closely related host species, say those from humans and their closest relative, the chimpanzee.

The suggestion that co-speciation may occur in some bacterial species that are commensals of animals suggests that these bacteria might be useful for tracking. Thus, strains of some commensal symbionts that have been found in non-host environments might be traced back to their host. If so, appropriate molecular procedures, such as sequence analyses of genes that reflect the co-speciation process, could be useful in this tracing approach.

Table 2. Bacterial Biogeography and Co-speciation Postulates1

No. 1: At least four bacterial strains of a purported taxon must be isolated from different samples taken from one ecosystem (or host). Ideally, bacteria should be isolated using extinction dilution as well as gross enrichment methods to ensure that the greatest variety of organisms of a particular type can be selected and obtained in a pure culture.
No. 2: These strains must be shown to be indigenous to the habitat from which they were obtained by showing their growth in the habitat at some time during the annual or other periodic cycle of the environment. For some organisms, such as mat-forming cyanobacteria, observation alone may be sufficient. For most bacteria, further research will be needed.
No. 3: At least four strains, potentially of the same taxon as in Number 1, must be obtained from one or more other geographically separate ecosystems (or hosts) that are similar to the ecosystem from which the first strains were obtained using the same procedures used to isolate the first strains. These new strains must be shown to be indigenous to these other ecosystems (or hosts).
No. 4: The two or more groups of strains from the two or more geographically separate ecosystems (or hosts) must be analyzed phylogenetically by sequencing two or more appropriate genes. The choice of genes may differ from one phylogenetic group to another. For example, 16S rDNA sequences do not appear to provide sufficient resolution for most prokaryotic groups. If the strains show no evidence of geographic clustering (clades) in either phylogenetic analysis, then they are considered cosmopolitan. In contrast, if the phylogenetic analysis indicates that geographic clustering occurs, then they are provisionally considered to be endemic to those areas. Such strains would be called geovars. Geovars may or may not be the same species. This would need to be determined by separate analyses in Number 5 (below). Cultures of the geovars should be maintained in national culture collections for other researchers to study.
No. 5: (Optional) If one wishes to determine whether the two or more groups of strains from the two or more geographically separate ecosystems comprise the same species, then they must be compared with one another directly. Tests to be conducted include, but are not limited to, DNA/DNA reassociation, sequencing of 16S rDNA, and phenotypic analyses that are appropriate to the phylogenetic group including, but not limited to, nutrition, physiology, fatty acid composition, and cell and colony morphology.

[1] Adapted from J. T. Staley (1999). See Staley and Gosink (1999) for more detail.

References

Beeder, J.R., K. Nielsen, J.T. Rosnes, T. T¿rsvik, and T. Lien (1994). "Archaeoglobus fulgidus isolated from hot North Sea oil field waters." *Applied and Environmental Microbiology*, 60: 1,227-1,234.

Fulthorpe, R.R., A.N. Rhodes, and J.M. Tiedje (1998). "High levels of endemicity of 3-benzoate-degrading soil bacteria." *Applied and Environmental Microbiology*, 64: 1,620-1,627.

Hedlund, B.P., and J.T. Staley (2002). "Phylogeney of the genus Simonsiella, including S. kuhniae sp. nov. and other members of the Neisseriaceae." *Journal of Systematic and Evolutionary Microbiology*, in press.

Hedlund, B. P., and J. T. Staley. "Biogeography of Cycloclasticus strains: A widely dispersed genus of polycyclic aromatic hydrocarbon degrading marine bacteria." *Archives of Microbiology*, submitted.

Staley, J.T. (1997). "Biodiversity: Are microbial species threatened?" *Current Opinion in Biotechnology*, 8: 340-345.

Staley, J.T. (1999). "Bacterial biodiversity: A time for place." *ASM News*, 10: 1-6.

Staley, J.T., and J.J. Gosink (1999). "Poles apart: Biodiversity and biogeography of polar sea ice bacteria." *Annual Reviews of Microbiology*, 53: 189-215.

JIM STALEY is a Professor in the Department of Microbiology at the University of Washington, where he has taught since 1971. Currently, he is a Professor in the Department of Microbiology. A general microbiologist, he is interested in environmental microbiology, microbial ecology, sea ice bacteriology, microbial diversity, and the evolution and taxonomy of Bacteria. Recently, he co-authored a textbook entitled, Microbial Life, which will be published by Sinauer in the spring of 2002. Staley received an A.B. in Mathematics at the University of Minnesota, an M.SC. in Microbiology at Ohio State University, and a Ph.D. in Bacteriology from the University of California, Davis.

Source Tracking Fecal Bacteria in the Environment Using rep-PCR DNA Fingerprinting: Prospects and Problems

LeeAnn K. Johnson, Mary B. Brown, Priscilla E. Dombek,
and Michael J. Sadowsky
Department of Soil, Water, and Climate
University of Minnesota

Introduction

Many of Minnesota's rivers and streams do not achieve the Clean Water Act's "swimmable" goal due to elevated numbers of fecal coliform bacteria. In the 1996 report to Congress on the condition of Minnesota's rivers, lakes, and streams (as reported by the Minnesota Pollution Control Agency), 47 percent of the river miles assessed could not support swimming due to high levels of fecal bacteria. Sources of fecal coliform bacteria include runoff from feedlots and manure-amended agricultural land, wildlife, inadequate septic systems, urban runoff, and sewage discharges. The ability to distinguish between human and animal sources of fecal contamination is an important assessment tool, both for evaluating possible health risks and for developing effective control strategies.

We have been using the rep-polymerase chain reaction (rep-PCR) technique to generate DNA fingerprints of *E. coli*. This PCR-based method does the following:

- Amplifies DNA sequences between naturally occurring, highly conserved, repetitive elements in the *E. coli* genome.

- Generates a unique pattern of PCR products for each strain.

The rep-PCR technique has the necessary sensitivity and resolving power to differentiate between strains of fecal coliform bacteria originating from different human and animal sources (Dombeck et al., 2000). Of the various genetic fingerprinting strategies, rep-PCR is a relatively rapid, simple, and low-cost technique. We have created a "known source" *E. coli* rep-PCR DNA fingerprint database, including 2,466 isolates obtained from 13 sources: cows, pigs, sheep, goats, turkeys, chickens, ducks, geese, deer, horses, dogs, cats, and humans. The known-source library has been applied to fecal bacteria isolated from four impacted Minnesota waterways.

Organisms yielding similar DNA banding patterns can be regarded as being identical or near identical and, as such, may be useful to define sources of fecal contamination.

Correspondence should be addressed to:

LeeAnn K. Johnson
Senior Scientist, Department of Soil, Water, and Climate
University of Minnesota
1991 Upper Buford Circle, 439 Borlaug Hall, St. Paul, MN 55108
Tel: 612 625-4273 Fax: 612-625-6725 Email: ljohnson@soils.umn.edu

Methods

***E. coli* Isolation**. Fresh fecal material was collected from individual animals, streaked onto mFC agar plates, and incubated at 44.5°C for 24 hours. Water samples were collected in sterile Whirl-Pac bags and kept at 4°C until processed. Water samples were analyzed for fecal coliforms on mFC agar by membrane filtration method (Standard Methods 9222 D). Characteristic blue colonies from the mFC plates were streaked to MacConkey agar (Difco), patched onto CHROMagar ECC (CHROMagar Microbiology, Paris, France), and incubated overnight at 37°C. Colonies from the MacConkey plates were used to inoculate citrate agar, EC-MUG broth (Difco), 1-percent tryptone (Difco), and MR-VP broth (Difco). Isolates yielding typical responses for *E. coli* on all media were designated as *E. coli* and used for subsequent studies. Isolates giving atypical responses to any test were further screened using an API 20E test kit.

***E. coli* Preparation and PCR Conditions**. *E. coli* isolates were streaked onto Plate Count Agar (Difco) and grown overnight at 37°C. Colonies were picked with a 1-microliters [µL] sterile inoculating loop and suspended in 100 µL of sterile water. Rep-PCR fingerprints were obtained using the BOX A1-R primer, 5' CTACGGCAAGGCGACGCTGACG 3'. PCR reactions were set up as described previously (Rademaker et al., 1998) using 2 µL of cell suspension as template. PCR was performed using an MJ Research PTC 100 (MJ Research, Waltham, Massachusetts), using the protocol specific for these thermocyclers (Rademaker et al., 1998). Five µL of 6 X loading dye was added to each 25 µL PCR reaction, and 10 µL of each reaction mixture was separated on a 1.5-percent horizontal agarose gel. A 1-kilobase (kb) size ladder (Life Technologies, Gaithersberg, Maryland, 0.5 micrograms per well [µg/well]) was loaded into the two outside and one middle well of the gel. Gels were electrophoresed in 0.5 X TAE (20 millimole [mM] Tris-acetate and 0.5 nM ethylenediaminetetraacetic acid [EDTA]) buffer at 4°C for 17.5 hours at 70 volts with constant buffer recirculation. Gels were stained for 20 minutes in 0.5-µg/mL ethidium bromide. Gel images were captured using a FOTO/analyst Archiver electronic documentation system (Fotodyne Inc., Hartland, Wisconsin).

Computer-assisted rep-PCR Fingerprint Analysis. Gel images were normalized and analyzed using BioNumerics v.2.5 software (Applied Maths, Sint-Martens-Latem, Belgium). Lanes were normalized using a 1-kb ladder from 298 to 5,090 base pairs (bp), as external reference standards. Similarity coefficients were generated by using the curve-based cosine correlation algorithm, with 1-percent optimization. The percentages of isolates assigned to their correct source group were calculated using cosine similarity coefficients and Jackknife analysis with maximum similarity. Unknowns were identified by comparing the known source database using cosine similarity coefficients with 1-percent optimization and maximum similarity.

Results and Discussion

Known Source Library

The rep-PCR DNA fingerprint technique was investigated as a means to differentiate human from animal sources of fecal bacteria. BOX A1-R PCR primers were used to generate 2,466 DNA fingerprints from *Escherichia coli* strains from human and animal sources (humans, dogs, cats, horses, deer, geese, ducks, chickens, turkeys, cows, pigs, goats, and sheep) in Minnesota (Table 1). This constituted our known source DNA fingerprint library.

Table 1. Animal Source Groups and rep-PCR DNA Fingerprints Generated from E. coli Isolates

Source Group	Individuals Sampled	Fingerprints	Unique Fingerprints[a]
Cat	37	108	48
Chicken	86	231	144
Cow	115	299	191
Deer	64	179	96
Dog	71	196	106
Duck	42	122	81
Goose	73	200	135
Goat	36	104	42
Horse	44	114	79
Human	197	307	211
Pig	111	303	215
Sheep	37	101	61
Turkey	69	202	126
Total	982	2466	1535

[a]Identical *E. coli* genotypes from the same animal removed.

The relatedness of the known source isolates to each other was determined by cluster analysis using cosine correlation coefficient, a curve-matching algorithm. Jackknife analyses indicated that 79 to 96 percent of animal and human isolates were assigned into the correct source groups (Table 2); however, when only unique isolates were examined (isolates from a single animal having distinct DNA fingerprints), Jackknife analyses indicated that 44 to 75 percent of the isolates were assigned to the correct source group (Table 2). This indicates that:

- Failure to remove identical fingerprints from the analysis results in an overestimation of the ability of the database to assign isolates to their correct source group.
- The current library size does not adequately capture genetic diversity present in natural *E. coli* populations.

Table 2. Percentage of known source DNA fingerprints assigned to the correct source group by jackknife analysis[a]

Source	All Fingerprints	Unique Fingerprints
Cat	95	44
Chicken	82	58
Cow	83	62
Deer	89	52
Dog	90	65
Waterfowl	80	63
Goat	96	64
Horse	69	47
Human	79	57
Pig	80	62
Sheep	81	54
Turkey	90	75

[a]Done using cosine correlation, maximum similarity, and 1-percent optimization.

However, if broader source groupings are allowed, then our unique DNA fingerprint library, in its current size, can correctly assign a larger percentage of isolates. For example, if the 13 animal sources are grouped as "domestic" (cat, chicken, cow, dog, goat, horse, pig, sheep, turkey), "wildlife" (deer, duck, goose), and "human," correct source group assignments are 88, 66, and 57 percent, respectively.

Moreover, when a limited subset of the unique fingerprint database consisting of humans, cows, pigs, and turkeys was examined by Jackknife analysis, 71 to 79 percent of *E. coli* isolates were assigned to the correct source group. Accordingly, these results indicate that a targeted subset of the DNA fingerprint database should be used to more precisely determine sources of fecal pollutants in watersheds where specific source groups are known to be present or absent.

Identification of Unknown Environmental Isolates

BOX A1-R DNA fingerprints were generated from 300 to 400 *E. coli* isolates obtained from each of the four Minnesota watershed areas (Mississippi River, Prairie Creek, Rush River, and Grindstone River). Source group identification was achieved by comparing DNA fingerprints from the river isolates to the 2,466 *E. coli* DNA fingerprints in our known source library. Based on similarity threshold values of 80 percent or greater, 98 percent of the environmental isolates could be assigned to a source group (Table 3); however, when the environmental *E. coli* isolates are compared to those in our DNA fingerprint library in a more stringent manner (i.e., 85, 90, and 95 percent similarity threshold values), a decreasing number of the isolates could be grouped with those in the library (Table 3). This indicates that:

- The diversity of environmental *E. coli* is much greater than originally anticipated.
- The database library needs to be expanded to include more isolates from each of the animal sources.
- Additional potential animal sources need to be examined.

Table 3. Assignment of Environmental E. coli to Animal Source Groups

Stream	Fingerprints In Database	Percent Assigned at Various Similarity Thresholds			
		>80%	>85%	>90%	>95%
Mississippi River	338	97	84	55	4
Grindstone River	348	96	90	63	8
Prairie Creek	428	99	89	67	9
Rush River	322	98	93	65	20

We are currently expanding the database library to include more beaver, deer, waterfowl, and human *E. coli* isolates to further increase the accuracy of the database and its ability to discriminate among sources of *E. coli* in water.

For our analyses, an environmental organism was assigned to an animal source group if it had >90-percent match, based on cosine correlation and maximum similarity, to a DNA fingerprint pattern in the database library (Table 4). Overall, about 63 percent of the environmental *E. coli* isolates were found to match those in the DNA fingerprint library at a similarity threshold value of >90 percent (Table 3). In general, all tested sites were dominated by *E. coli* bacteria originating from livestock species (especially cows) and other domestic animals. Generally, humans and wildlife species were minor contributors. Taken together, our results indicate that rep-PCR

technique, using the BOX A1-R primer, may be a useful and effective tool to rapidly determine sources of fecal pollution, and that a large scale *E. coli* database is required to accurately distinguish between potential human and animal sources.

Table 4. Environmental (Unknown) E. coli Assigned to Source Groups at >90 Percent Similarity Threshold

	Percent Assigned to Source Group			
Source	Grindstone River	Mississippi River	Prairie Creek	Rush River
Cat	8	1	1	0
Chicken	5	6	16	14
Cow	27	18	25	20
Deer	3	8	8	3
Dog	5	17	4	7
Goat	1	1	2	3
Horse	4	3	4	5
Human	14	8	4	3
Pig	5	12	13	5
Sheep	13	3	6	2
Turkey	5	8	9	9
Waterfowl	10	15	9	28

We are currently examining alternate methods for detecting rep-PCR fragments that will reduce gel-to-gel variability and increase the resolving power of rep-PCR. One method that holds great promise is Fluorophore Enhanced Rep-PCR. Fluorophore Enhanced Rep-PCR uses a fluorescent-labeled primer in the rep-PCR reaction and a molecular weight marker set in each gel lane that is labeled with a second fluorophore. This allows for more precise normalization of the gel. Labeled fragments in agarose gels are then detected by using a fluorescence scanner.

References

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." *Applied and Environmental Microbiology*, 66: 2,572-2,577.

Rademaker, J.L.W., F.J. Louws, and F.J. de Bruijn (1998). "Characterization of the diversity of ecologically important microbes by rep-PCR genomic fingerprinting." *Molecular Microbial Ecology Manual*, 3.4.3: 1-27.

LEEANN JOHNSON is a Senior Scientist in the Department of Soil, Water, and Climate at the University of Minnesota, where she manages a research project utilizing rep-PCR DNA fingerprinting methods to track sources of fecal coliform bacteria in Minnesota watersheds. Prior work experience includes 12 years as a research scientist at a wastewater treatment utility in St. Paul, Minnesota, managing projects that included fecal source tracking, activated sludge bulking control, and the toxicity and treatability of industrial discharges, among others. She also spent 3 years as a pollution control specialist with the Minnesota Pollution Control Agency, the state environmental regulatory agency. Johnson has a B.S. in Microbiology from the University of Minnesota, and is currently completing a Master of Public Health degree in Environmental Microbiology at the University of Minnesota.

Microbial Source Tracking Using Antibiotic Resistance Analysis

Bruce A. Wiggins
Professor, Department of Biology
James Madison University

Fecal contamination in natural waterways can lead to several problems, including an increased incidence of pathogens. Additionally, the increased levels of phosphorous and nitrogen in natural waterways due to fecal pollution can lead to algal blooms that, when degraded, result in the deoxygenation of waterways. This situation is currently leading to the deterioration of many aquatic environments, such as the Chesapeake Bay. Fecal contamination in waterways has consistently been demonstrated by the presence of indicator organisms, such as fecal coliforms or enterococci; however, it is difficult to identify the sources of fecal contamination in waters receiving mixed agricultural and human waste. Once the source is identified, steps can be taken to control the influx of fecal pollution. This is especially important now because of the urgent need for total maximum daily load studies for many polluted watersheds.

Many approaches have been attempted for the source identification of fecal contamination, including chemical methods, species-specific methods, genotypic methods, and phenotypic methods. Chemical methods rely on the detection of a source-specific chemical, such as caffeine or optical brighteners. Species-specific methods are based on the unique occurrences of microorganisms in source animals. Examples of this include measuring the ratio of fecal coliforms to fecal streptococci, looking for source-specific species of fecal streptococci, detecting the presence of unique viruses and bacteriophages, and detecting unique gene sequences using terminal restriction fragment length polymorphism. Genotypic methods rely on detecting differences in genetic sequences among bacteria from different source animals. Examples of this include pulsed-field gel electrophoresis, ribotyping, and rep-polymerase chain reaction (rep-PCR.) Phenotypic methods are based on detecting physical differences in the microorganisms from different source animals. Examples of phenotypic methods include metabolic measurements, such as Biolog, immunological methods, and antibiotic resistance analysis.

Antibiotic resistant bacteria can develop in animals and humans as a result of treatment with antibiotics. The antibiotic resistance analysis method is based on the rationale that because antibiotics exert selective pressure on the fecal flora of the animals that ingest or are treated with the antibiotic(s), and because different types of animals receive differential exposure to antibiotics, different animals will have bacteria with different patterns of antibiotic resistance. Our laboratory uses enterococci as the test organism in identifying sources of fecal contamination. We chose enterococci because they survive well in natural waters and can be isolated from all potential sources of fecal pollution. The method also works well using E. coli as the test organism.

Correspondence should be addressed to:

Bruce A. Wiggins, Ph.D.
Professor, Department of Biology
James Madison University
312 Burruss Hall
Harrisonburg, VA 22807
Tel: 540-568-6196 Fax: 540-568-3333 E-mail: wigginba@jmu.edu

For antibiotic resistance analysis, the test bacteria are isolated from samples of known fecal sources using the membrane filter method onto m-Enterococcus agar. Isolates are picked from the filters and are then grown on agar plates containing various concentrations of 11 different antibiotics. The use of multiple concentrations of each antibiotic is a major component of antibiotic resistance analysis, and is what makes it much more successful than previous methods based on antibiotic resistance. For enterococci, we use three or four concentrations of each of the 11 drugs, for a total of 37 plates. Each isolate is scored as sensitive or resistant to each concentration of each drug. For known fecal samples, we generally test between 10 to 12 isolates per sample. The percentage of isolates with identical resistance patterns is low (usually about 15 percent).

The resulting resistance patterns of each isolate are then analyzed using discriminant analysis, a multivariate statistical method. Multivariate statistical analyses are another major component of antibiotic resistance analysis. Discriminant analysis classifies the bacteria based on shared patterns of antibiotic resistance, and the results are pooled to form a "known library" of antibiotic resistance patterns from different fecal sources. The library is summarized in the form of a classification table. The average rate of correct classification is the average rate that known isolates are correctly classified, and is used to measure the reliability of the known library. The Minimum Detectable Percentage for each source type is determined by averaging the percentages of other source types that are misclassified as that type. This value is the minimum percentage for each particular source that can be detected in a stream sample. Once the known isolates are classified, then the resistance patterns of isolates from natural waterways are compared with this known library to determine the source(s) of fecal pollution in that waterway. We generally collect 46 to 48 isolates for each stream sample, which gives a precision of approximately 2 percent.

An example of this analysis is presented in the accompanying tables. The isolates shown here were collected as part of a total maximum daily load study of Moores Creek in Albemarle County, Virginia. When the known sources (human sewage, dogs, livestock [including beef cattle, goats, and horses] and wild animals [including geese]) were grouped as either "human" or "animal" (a two-way analysis), the isolates were correctly classified at an average rate (of correct classification) of 85 percent (Table 1). The Minimum Detectable Percentage for this analysis was 7 percent for human, and 22 percent for animal sources. While not 100 percent successful, the average rate of 85 percent is well above the 50 percent rate that would be expected based on random classification. When each type of source is classified separately (a four-way analysis), the average rate of correct classification falls to 72 percent (Table 2), but is still well above the random rate of 25 percent. The Minimum Detectable Percentages for these sources range from 7 to 11 percent.

Table 1. Classification of 980 Isolates of Enterococci From Known Animal and Human Sources in the Moores Creek Watershed

	Number (and Percent) of Isolates Classified As:	
Source	Animal	Human
Animal (n = 846)	**785 (93)**	61 (7)
Human (n = 134)	30 (22)	**104 (78)**
Minimum Detectable Percentage	22	7

Correctly classified isolates are shown in bold.
The average rate of correct classification for this analysis is 85 percent.

Table 2. Classification of 980 Isolates of Enterococci from Known Dog, Human, Livestock, and Wild Sources in the Moores Creek Watershed

	Number (and Percent) of Isolates Classified As:			
Source	Dog	Human	Livestock	Wild
Dog (n = 174)	**123 (71)**	28 (16)	10 (6)	13 (7)
Human (n = 134)	15 (11)	**97 (72)**	10 (8)	12 (9)
Livestock (n=434)	37 (8)	14 (3)	**311 (72)**	72 (17)
Wild (n=238)	12 (5)	6 (2)	47 (20)	**173 (73)**
Minimum Detectable Percentage	8	7	11	11

Correctly classified isolates are shown in bold.
The average rate of correct classification for this analysis is 72 percent.

In addition to isolates from known sources, enterococci were isolated from stream samples collected along the length of Moores Creek. Samples collected on September 19, 2000, were tested and classified based on the library of known isolates. When the two-way library was used, the majority of the samples contained enterococci from animal origin and four samples also contained bacteria of human origin (Table 3). Based on the four-way library, all of the samples contained bacteria from wild sources (Table 4). Seven of the samples also contained livestock isolates, four samples contained human isolates, and six had bacteria from dogs. From these results, we concluded that the watershed was polluted by both human and animal sources, and that wild sources and livestock were the most prevalent.

Table 3. Two-way Classification of Sources of Fecal Pollution In Moores Creek on September 19, 2000.

Site #	# of Isolates	% Animal	% Human	# FC per 100 mL
1	46	**89**	**11**	4,850
2	46	**93**	7	4,300
3	46	**9**	4	4,300
4	46	**52**	**48**	16,200
5	46	**83**	**17**	4,050
6	46	**67**	**32**	22,500
7	46	**100**	0	21,000
8	46	**96**	4	23,500
9	46	**98**	2	19,500
Average		**86**	**14**	10,273

Values in bold are above the Minimum Detectable Percentage.

Table 4. Four-way Classification of Sources of Fecal Pollution In Moores Creek on September 19, 2000.

Site #	# of Isolates	% Dog	% Human	% Livestock	% Wild	#FC per 100 mL
1	46	2	9	**22**	**67**	4,850
2	46	**9**	4	**50**	**37**	4,300
3	46	4	0	**24**	**72**	4,300
4	46	**9**	**33**	6	**52**	16,200
5	46	**19**	**9**	**20**	**52**	4,050
6	4	**9**	**9**	0	**82**	22,500
7	46	**22**	0	**26**	**52**	21,000
8	46	7	4	**37**	**52**	23,500
9	46	**19**	2	**33**	**46**	19,500
Average		**11**	**8**	**24**	**57**	10,273

Values in bold are above the Minimum Detectable Percentage.

Antibiotic resistance analysis is a simple, rapid, and inexpensive method of microbial source tracking. Any laboratory that is equipped for basic microbiological work can easily perform antibiotic resistance analysis. No advanced or expensive special equipment is needed, and no special training is required for personnel. Additionally, throughput for this method is quite high. Finally, the method, although somewhat labor-intensive, is very inexpensive to perform.

Antibiotic resistance analysis is a published method that has been independently validated (and is currently being used) in at least six other laboratories in Virginia, Florida, Texas, and Oregon. It is being used for total maximum daily load development in (at last count) 10 polluted watersheds in Virginia, and in others throughout the country. If you are interested in getting more information about the procedures and protocols of antibiotic resistance analysis, contact me at wigginba@jmu.edu.

BRUCE WIGGINS has taught in the Department of Biology in the College of Science and Mathematics at James Madison University since receiving a doctorate from Cornell University in 1987. His research focuses on developing improvements to the antibiotic resistance analysis method for determining the sources of non-point fecal pollution in streams, groundwater, and estuarine waters. Wiggins received a B.S. in Microbiology from Pennsylvania State University, and both a B.S. in Microbiology/Ecology and Ph.D. in Microbiology/Ecology/Environmental Toxicology from Cornell University. His graduate research focused on the ecology of bacterial viruses and the effects of protozoan predation on the biodegradation of nitroaromatic compounds.

Animal Source Tracking: A Complement to Microbial Source Tracking

R.D. Ellender, B.L. Middlebrooks, Shiao Wang,
Dawn Rebarchik, and D. Jay Grimes
College of Science and Technology
University of Southern Mississippi

Background

The identification of species-specific sIgA bound to fecal bacteria or food particles is proposed as a possible alternative method to determine the origin of recent fecal waste in the aquatic environment. Immunoglobulin A is found in bodily secretions that have contact with the external environment. Part of its role is to minimize the attachment to and penetration of microbes into the epithelial lining. Exocrine secretions of the lacrimal and salivary glands, nasal discharges, genitourinary, seminal, and intestinal fluids all contain this antibody class. In secretions, dimers of IgA linked by a polypeptide known as the J-chain are associated with a polypeptide chain known as the secretory component (SC), which is instrumental in the secretion of IgA. sIgA found in the lumen of the intestine can be transported directly from gut epithelial cells when bound to the SC, or from the blood by lymphatic drainage of the lamina propria. The sIgA secreted into the intestine functions in the control of pathogenic and normal microbial constituents and is related to the complexing of food allergens. A major role of sIgA appears to be preventing microorganisms from attaching to the epithelium, thereby preventing colonization and penetration.

SIgA is produced by all mammals. In the absence of overt infection or immunization, gastrointestinal antibodies are apparently formed against a variety of bacteria, viruses, or other parasites. For example, evidence suggests that anti-non-pathogenic Escherichia coli sIgA is present in intestinal fluids in the absence of disease. Additional evidence demonstrates that the presence of SC protects IgA from cleavage by digestive enzymes of the gastrointestinal tract. The half-life of 7S IgA and the IgA2-J-SC dimer (11S) is approximately the same (4.5 to 6 days), but varies depending on the subgroup. The half-life of sIgA in the gastrointestinal tract is not known. Proteolytic enzymes, such as trypsin, chymotrypsin, or pepsin, digest the molecules producing 5S F(ab)2-like and 3.5S F(ab) fragments; however, the rate and degree of IgA digestion is less than that of IgG. Fc fragments are not produced from human sIgA by the above enzymes, but an enzyme isolated from Streptococcus sanguis produces this fragment. The enzyme can be isolated from colonic secretions along with Fc-like fragments and may play an important, although presently undefined, role in the ecology of secretory immunoglobulins.

The universal presence of sIgA in the fecal wastes of all mammals, its association with easily concentrated particulate matter, and its apparent relative resistance to enzymatic degradation, together with the availability of immunological reagents and procedures to detect, identify by species, and quantify sIgA, suggested its possible value as an indicator of the fecal pollution of water.

Correspondence should be addressed to:

Rudolph D. Ellender, Ph.D.
Associate Dean for Research and Development
College of Science and Technology, University of Southern Mississippi
P.O. Box 5165, Hattiesburg, MS 39406-5165
Tel: 601-266-4887 Fax: 601-266-5829 E-mail: Rudolph.Ellender@usm.edu

Research

We have successfully demonstrated that sIgA has a half-life measured in days in most water samples tested, and that it is resistant to ranges of pH, temperature, salinity, and other parameters considerably beyond the norm in water samples. We also successfully applied immunohistochemical and enzyme immunoassays using standard indicators (fluorescein, alkaline phosphatase/p-nitrophenyl phosphate) in the detection and quantitation of bound and eluted IgA and IgA fragments. It has the potential to serve as an indicator of recent pollution. Following the initial studies, the method was applied to field studies. What follows are certain critical aspects of the procedure.

Standard Assay

A standard enzyme linked immunosaorbent assay (ELISA) protocol was developed using 96 well polystyrene assay plates (Falcon 3915) (Figure 1) coated with 50 microliters (µL) (8 micrograms per milliliter [µg/mL]) of goat anti-human IgA (alpha chain, kappa chain, or secretory component specific, depending on the experiment) (Sigma) in coating buffer (12.8 millimolar [mM] Na2CO3, 35 mM NaHCO3, 3 mM NaN3, pH 9.6).

Figure 1. 96 well polystyrene plate.

Plates were held at room temperature (22 to 27° C) for 16 hours. The wells were washed three times with 250 µL of washing buffer (8 mM Na2HPO4, 1.46 mM KH2PO4, 137 mM NaCl, 3 mM NaN3, 2.7 mM KCl, 5 mL/liter Tween 20, pH 7.4). A 100-µL volume of 5-percent bovine serum albumin in coating buffer was added to each well and incubated at 37°C for 30 minutes.

The plates were washed as before, and 50 µL of the IgA samples in incubation buffer (10 mM Na2HPO4, 10 mM KH2PO4, 137 mM NaCl, 5 mL/liter Tween 20, pH 7.0) were added to the wells. The plates were incubated at 37°C for 1 hour.

After washing, 50 µL (dilution depended upon the specific lot) of alkaline phosphatase conjugated goat anti-human IgA (alpha chain, kappa chain, or secretory component specific) in incubation buffer was added to each well.

The wells were again washed, and 50 µL of a chromogenic substrate (2 milligrams per milliliter [mg/ml] p-nitrophenylphosphate pNP in diethanolamine buffer [923 mM diethanolamine, 0.5 mM MgCl, pH 9.8]) was added and allowed to react at room temperature for 30 minutes. Fifty µL of three normal sodium hydroxide (3N NaOH) was added to stop the reaction, and the change in absorbance (405 nanometers [nm]) was measured with a microwell spectrophotometer (Bio-Tek EL 307) (Figure 2).

Unless otherwise indicated, the reported absorbances are the average of three replicate wells with the background subtracted.

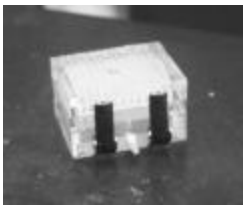

Figure 2. Microwell spectrophotometer.

Field Studies

Field samples were conducted using 1 liter (L) (point source) or 380 L (unknown source) volumes. Particulates in suspension were collected by centrifugation or by filtration. A 1 L sewage sample was centrifuged at 15,000 x g for 30 minutes to pellet the particulates. The pellet was resuspended in 100 mL of elution buffer (1M NaCl pH 7.0) and shaken for 5 minutes on the orbital shaker to elute the IgA. The elution mixture was then centrifuged again for 30 minutes at 15,000 x g to pellet the particulates, and the supernatant containing the IgA in elution buffer was concentrated by forcing it through a Millipore 30,000 nominal molecular weight limit (NMWL) cellulosic disc using an Amicon Ultrafiltration System. The sample was concentrated to 1 to 2 mL and removed to be assayed. The membrane was washed with 1.0 mL incubation buffer to remove any remaining IgA; the fluid was kept for assay.

The amount of human IgA present was determined using the standard ELISA. Commercial human sIgA was used in this assay along with the test sample as a set of standards with which to develop a standard concentration curve for quantitation of IgA in samples. Each of the concentrated samples was tested undiluted as well as dilutions of 1/2, 1/4, 1/8, and 1/16 in incubation buffer. The amount of human IgA present in the membrane wash was calculated to be 0.65 µg per mL of sewage. In the membrane supernatant, the amount was found to be 0.36 µg of IgA per mL of sewage; therefore, a total of 1.01 µg of IgA per mL was present in the sewage sample. Thus, the applicability of the methodology under development to assay of a "real world" sample was demonstrated.

The 1L field samples were shown to consistently contain sIgA. This type of experiment was used to study the effects of:

- Freezing sewage particulates on the recovery of IgA.
- Degradation of human IgA in sewage over time.
- Sewage treatment on the removal of IgA from waste particulates.
- Assay of sewage particulates for other classes of antibodies.
- Salinity on the recovery of IgA from sewage particulates.
- Levels of sIgA at regional sewage treatment plants (non-lagoon).
- Elution and concentration of sIgA from a series of lagoon samples taken over a 2-month period.

Studies were also conducted in waters of unknown or suspected sIgA content (i.e., waters likely contaminated with septic tank runoff.) Large (380 L) samples of water from various locations in a

few selected bayous were obtained, eluted, concentrated, and assayed for the presence of human sIgA. Sample water was obtained from just below the surface in hopes of retrieving recently released sewage particulates. Samples were routinely pumped through a double filter (Figure 3).

Figure 3. 380-liter particulate concentrator.

The first filter was of a large pore size, which retained particles that would clog the second (0.2-micrometer [µm]) filter. The filters/particulates were held on ice and transported to the lab for processing. Filters were cut into pieces approximately 1-inch square in size, and the pieces were eluted with 100 mL of one molar sodium chloride (1M NaCl) elution buffer and agitated for 5 minutes. The filter pieces were then removed and this suspension was centrifuged (15,000 ´ g for 30 minutes) to remove remaining particulates. The supernatant was then concentrated with the Amicon Stirred Ultrafiltration System (Figure 4) containing a Millipore filter. The samples were assayed using the standard plate ELISA.

Figure 4. Amicon ultrafiltration device.

Studies to Improve the Assay

Studies to improve the assay system were conducted. To facilitate the detection of sIgA in the environment, four indicator amplification systems were adapted to the "standard" ELISA and their sensitivity was evaluated. All four used the standard method of binding capture antibodies to polystyrene plates and adding the IgA-containing samples. The differences lie in the labeled antibodies and the enzyme reactions that led to color or light production. They included:

- Avidin-biotin amplification system.
- Amplified indicator system using NADH+ recycling and diaphorase.
- Amplified indicator system using firefly luciferase.
- Amplified indicator system using bacterial luciferase and the recycling of NADH+.

Results

Sewage samples were frozen to determine if IgA could survive freezing (Table 1). Results indicated that the freezing process had little or no effect on recoverable IgA; therefore, it should be possible to freeze samples from other locations, permitting them to be stored and/or shipped,

pending assay. The results of IgA degradation studies using preparations of fecal and sewage samples indicated that IgA persisted in detectable levels for at least 2 to 3 days. Such a persistence time is satisfactory for a potential indicator of human fecal pollution, being neither too long nor too brief. Too brief a survival time would make it unlikely that the indicator would be detected, and too long a time would make it impossible to determine when or where pollution occurred. The persistence time observed in this study means that the detection of IgA in water would be indicative of recent pollution and would facilitate detection of the source.

Table 1. The Effects of Freezing Raw Sewage Samples on the Recovery of Particulate-associated Human sIgA

Trial	Unfrozen	-70°C, 24hr
1	0.66	0.64
2	0.78	1.07
3	0.99	1.04

All figures represent µg IgA recovered per gram of particulates after elution.

The detection of human sIgA in raw sewage confirms that IgA is excreted into sewage and that it persists for sufficient time to survive the transport of the sewage to the point of entry into sewage treatment facilities. IgA was detected at the entry point of several local facilities, but it could not be detected at several different stages of the sewage treatment process (Table 2). IgA was found to be completely removed by proper sewage treatment; thus, IgA detected in waters after treatment would be a good indicator of malfunctioning treatment facilities or premature releases of sewage. The fact that IgA dropped below detectable levels in most waters after several days, and in all properly operating sewage facilities, means that the detection of human IgA in a body of water indicates recent contamination by untreated human fecal material. Isotypes of human immunoglobulins other than IgA (IgG, IgM) were not be detected in sewage (data not presented).

Table 2. Persistence of Detectable Human sIgA In Raw Sewage Incubated at 25°C

Trial	0 hr	24 hr	48 hr	72 hr
1	0.70	0.37	Not Done	0.00
2	1.03	0.48	0.56	0.00

Salt elutes IgA from fecal particles, and it is expected that in estuarine or marine waters, sIgA attached to particulates will elute into the water column (Table 3). Our studies concluded that 1M NaCl produced the most satisfactory and reproducible elution results; hence, this concentration of NaCl was used in all subsequent elutions.

Table 3. Recovery of Human IgA from Raw Sewage Samples Previously Exposed to Various Salinities

Trial	Control	5 ppt	20 ppt	35 ppt
1	3.18	1.23	1.18	0.63
2	2.20	1.36	1.59	0.29
3	1.19	1.23	0.59	0.46

ppt = Parts per trillion.

Concentrations of human sIgA were detected in 1 L sewage samples obtained from various stages of sewage treatment at three Mississippi Gulf Coast sewage treatment facilities (table not included). Site 1 treated waste with an oxidation ditch (activated sludge). At this site, the starting level of HsIgA on two sampling trials was 0.32 and 0.21 micrograms per grams (µg/g), respectively. No sIgA could be detected in the final effluent. Similar results were found at other locations using trickling filters and activated sludge. A small amount (0.03 µg/g) was found in the effluent of the trickling filter, but none was detected in any sample of water treated by activated sludge.

These data (Table 4) indicate the concentrations of human sIgA in lagoon influent over a 2-month period (range 0.59 to 3.18 µg/g). The lagoons have an approximate retention time of 42 days. SIgA was not found in the effluent of the lagoon complex.

Table 4. Amounts of Human IgA Recovered from a Series of 1 L Raw Sewage Samples Taken Over a 2-month Period

Trial	µg IgA/ Gram Particulates	Trial	µg IgA/ Gram Particulates
1	0.70	6	0.59
2	1.14	7	0.74
3	0.66	8	1.22
4	0.78	9	1.03
5	0.99	10	3.18

Human sIgA was demonstrated in a series of water samples obtained from various locations on bayous feeding into Biloxi Bay (Mississippi). These sites were selected based on information from government agencies. Although there was no known raw sewage influx, all of the residences in the area relied on septic tank systems. As may be seen, low concentrations of human IgA were detected on three occasions from two of the sampling sites. As Table 5 shows, the detection of IgA was intermittent. Sufficient data points were not obtained to determine if the intermittency was due to infrequent contamination by the septic systems or to fluctuations of IgA levels below the sensitivity of the indicator system used in this study. The methods were also tested for the ability to discriminate between human and bovine fecal contamination of water (Table 6). These data show that bovine sIgA was recovered from 100 gallons of water receiving bovine contamination. The same concentrates did not contain human sIgA.

Amplified assay methods developed were capable of detecting IgA in samples in concentrations as low as 1 nanogram per milliliter (ng/mL). Concentration/elution protocols developed resulted in a potential concentration of over 500,000 fold; thus, if these methods are combined with the amplified indicator systems in ELISAs, it would be possible to detect IgA in concentrations as low as 0.5 femtograms/mL in water samples. The demonstration of human IgA on several

Table 5. Amount of Human sIgA Recovered from Particulates Filtered from 100-gallon Water Samples From Sites on Three Bayous Feeding into Biloxi Bay, Mississippi

Site	Trial 1	Trial 2	Trial 3	Trial 4
1	0.188	0.023	0.00	0.00
2	0.00	0.023	0.00	0.00
3	0.00	0.00	0.00	0.00

Figures represent µg of IgA per 100 gallons of water.

Table 6. Amount of Bovine and Human IgA Recovered from Particulates Filtered from Water Samples from Three Sites on a Freshwater Stream Receiving Runoff from a Livestock Sale Barn

Site	Bovine Trial 1	Human Trial 1	Bovine Trial 2	Human Trial 2
Site #1	0.243	0.000	0.071	0.000
Site #2	0.116	0.000	0.065	0.000
Site #3	0.049	0.000	0.031	0.000

Figures represent µg of IgA per 100 gallons of water. Assay by standard ELISA.

occasions in field samples clearly demonstrated the efficacy of the protocols. The most reliable assays were the bioluminescent assay incorporating firefly luciferase and the avidin-biotin system. Both of these were capable of detecting IgA alpha chains at levels of 0.01 µg/mL as compared to 0.05 µg/mL for the standard ELISA. The amplification system employing the recycling of nicotinamide adenine dinucleotide (NAD+) and reduction of p-iodo tetrazolium violet by diaphorase was the most sensitive, detecting approximately 1 x 10-3 µg/mL; however, greater experimental variation was encountered with this assay than with other assays. The final system utilizing NAD+ recycling and bioluminescence was, initially, considered to possess the highest potential for amplification. Many problems were encountered with this assay due to the considerable number of variables innate in the series of reactions involved. In the end, the highest sensitivity that could be constantly attained was 1 x 10-3 µg/mL, the same as that of the diaphorase system. Furthermore, this level was only obtainable with samples assayed immediately after the reaction mixture was prepared. The major problem was traceable to the degradation of flavin adenine dinucleotide, and it proved difficult to prepare two different solutions of this reagent that produced the same reaction levels. If the sensitivity of this method can not be increased to a level considerably higher than that of the diaphorase assay, it would certainly not be a feasible method, considering the substantial increase in expense and difficulties in achieving standardization; however, if problems associated with the stability of the various components of the indicator system can be eliminated, it seems that it has the potential to produce very high sensitivity.

This study demonstrated the feasibility of using the immunoclassification of wastewater particulates (i.e., isolation, concentration, detection of sIgA) to determine the animal source of water contamination. Further refinement of the methods could yield procedures that define the percent contribution of animals that contaminate water.

RUDOLPH ELLENDER has been with the University of Southern Mississippi since 1976, teaching courses as varied as nursing and public health microbiology to the diseases of aquatic animals and pathogenic microbiology. Currently, he serves as Associate Dean for Research and Development for the College of Science and Technology at the University of Southern Mississippi. Research interests include environmental microbiology; teleost and crustacean virology and tissue culture; in vitro carcinogenesis; the molecular epidemiology of viruses; tissue culture of marine plants; bioremediation; and the microbial decolorization of pulp mill wastewater. In addition, he holds a patent on the decolorization of pulp mill effluent with Dwaine Braasch. Ellender received both a B.S. and M.S. in Microbiology from Louisiana State University and a Ph.D. in Veterinary Microbiology from Texas A&M University.

Carbon Source Profiles, Pulsed-field Gel Electrophoresis Patterns, and Antibiotic Resistance Analysis

Charles Hagedorn
Professor, Department of Crop and Soil Environmental Sciences
Virginia Polytechnic Institute and State University

Water-quality improvement projects at the watershed level were initiated in Virginia in 1995. Funding by state agencies stimulated the development of methodologies to identify sources of fecal pollution in water (Hagedorn et al., 1999; Wiggins, 1996) and predated the United States Environmental Protection Agency's total maximum daily load program (U.S. Environmental Protection Agency, 1999). Carbon source profiles, based on the commercial Biolog system, pulsed-field gel electrophoresis (Simmons et al., 1995; Simmons et al., 2000), and antibiotic resistance analysis (Bowman et al., 2000; Hagedorn et al., 1999; Harwood et al., 2000; Wiggins et al., 1999), have been employed as phenotypic (antibiotic resistance analysis and carbon source profiles) and genotypic (pulsed-field gel electrophoresis) fingerprinting methods to determine sources of *Escherichia coli* and *Enterococcus* in water.

Antibiotic Resistance Analysis

Antibiotic resistance analysis was the first source-tracking method to be widely applied in Virginia and, to date, has been employed in some 23 total maximum daily load projects (Bowman et al., 2000; Graves et al., in press; McClellan et al., 2000). The Page Brook Watershed has been monitored since 1995 (Hagedorn et al., 1999), and an antibiotic resistance analysis of Enterococcus isolates identified cattle as the major source of fecal bacteria in the impaired stream segment. Stream fencing was installed on 12 of 17 farms in the watershed from 1996 to 1997. At two sampling sites where complete fencing occurred, fecal coliforms were reduced by an average of 84.5 percent (PB10) and 98 percent (PB12) over 4 years post-fencing (Table 1). At site PB16, where stream fencing was not complete, coliform reductions averaged 58.8 percent over 4 years post-fencing. The remaining predominant source of fecal bacteria in the stream is wildlife, and fecal coliform populations from wildlife are slowly increasing as riparian zones inside the fenced areas become more attractive wildlife habitats over time.

Table 1. Low-flow Fecal Coliform Populations from the Three Most Contaminated Sites (PB10, PB12, and PB16) in the Page Brook Watershed

Sampling Site	Average Fecal Coliforms (Colony-forming Units [cfu] per 100 mL)								
	Pre-fencing cfu	Post-fencing cfu and % Reduction							
	1997	1998		1999		2000		2001	
	Cfu	cfu	%	cfu	%	cfu	%	cfu	%
PB10	3,103	610	88.8	320	89.7	726	76.6	712	83.5
PB12	42,400	1,596	96.2	467	98.9	297	99.3	764	98.2
PB16	2,347	934	60.3	63.0	1,070	1,070	54.4	995	57.6

Correspondence should be addressed to:

Charles Hagedorn, Ph.D.
Professor, Department of Crop and Soil Environmental Sciences
Virginia Polytechnic Institute and State University
401 Price Hall, Blacksburg, VA 24061-0404
Tel: 540-231-4895 Fax: 540-231-3431 E-mail: chagedor@vt.edu

The known source Enterococcus library developed for Page Brook from 1995 to 1996 has been tested over the past 4 years by adding additional known source isolates each year to examine changes in the library as a function of time (Table 2). After an initial decline in rates of correct classification between 1996 and 1998 for all sources, the rates of correct classification increased in 1999 (except wildlife) and have remained very stable since then. Such consistent results demonstrate the utility of antibiotic resistance analysis as a fecal source-tracking tool when a library is built from intensive sampling of known sources over a relatively small watershed.

*Table 2. Rates of Correct Classification for the Page Brook Known Source Library Over Time**

	1996	1998	1999	2000	2001	Average
Livestock	85.0%	78.4%	86.2%	81.3%	82.6%	82.7%
Wildlife	87.3%	81.5%	80.6%	83.3%	81.9%	82.9%
Human	93.2%	86.3%	87.5%	91.6%	89.8%	89.7%
Average	88.5%	82.1%	84.8%	85.4%	84.8%	85.1%

*There are no differences between values across rows at P<0.01.

Carbon Source Profiles

Carbon source profiling is the most recent bacterial source-tracking method developed in our laboratory and is based on the commercial Biolog system (Holmes et al., 1994). Three hundred sixty-five Enterococcus isolates were collected from known sources in five geographical areas and identified to species. Discriminant analysis identified 30 of 95 wells in the GP2 MicroPlate that best classified the 365 isolates by source. The average rate of correct classification was 92.3 percent for a human versus non-human two-way classification when isolates from all regions were combined into one library (Table 3). The corresponding average rate of correct classifications for other classification schemes were 80.3 percent for a four-way classification of human versus livestock versus wildlife versus domestic pets, and 82.1 percent for a three-way classification without human isolates. The average rate of correct classification was 76.7 percent for a 13-way classification of individual sources and was 85.1 percent for a 12-way classification without human isolates.

Table 3. Rate of Correct Classification by Source, Based on Discriminant Analysis of the Enterococcus Library (365 Isolates)

Source	Number of Isolates	% Correctly Classified	
		Two-way Classification	
Human	105	93.3	
Non-human	260	91.3	
Average	365	92.3	
		Three-and Four-way Classification	
		With Human	No Human
Human	105	88.6	—
Livestock	105	78.5	81.9
Domestic Pets	50	84.6	92.3
Wildlife	105	69.7	72.0
Totals	365	80.3	82.1

Geographic variability among the isolates was demonstrated by making five individual libraries for the regions where the isolates were obtained. In this case, the average rate of correct classification (based on classification of the specific sources) was 100 percent for three of the libraries and over 92 percent for the other two. Of the 365 isolates, 323 were identified among six species by Biolog (*E. casseliflavus, faecalis, faecium, gallinarum, hirae, and mundtii*). When six individual libraries were made based on classification of the specific sources by Enterococcus species, the average rate of correct classification was 92.9 percent for one library and over 95 percent for the other five. Although the Enterococcus library is small, the results to date demonstrate the potential for carbon source profiles to be developed and used as a phenotypic method for determining sources of fecal pollution in water. One advantage of carbon source profiling is the use of a plate reader that eliminates the judgment decisions inherent to most other source tracking methods reported to date. Biolog now offers plates with specific sources of nitrogen, phosphorus, and sulfur in addition to carbon sources in the GP2 and GN2 plates. These plates could be screened, as we have done with the GP2 plate, to find additional nutrients that could be used to obtain even better rates of correct classifications with Enterococcus and *E. coli*.

Pulsed-field Gel Electrophoresis

Pulsed-field gel electrophoresis was first developed by Simmons as a source-tracking tool for use with *E. coli* (Simmons et al., 1995, Simmons et al., 2000). In the Four Mile Run Watershed in Virginia, acceptable matches were obtained for 278 (51.6 percent) of 539 isolates considered. For the isolates that could not be analyzed, 133 (24.7 percent) were found to have no matching record in the known source library. For the 278 that were matched, the largest categories were waterfowl, human, raccoon, and deer (Simmons et al., 2000), very typical results for a suburban watershed. When molecular methods are used as source-tracking tools, obtaining a substantial number of non-matching isolates appears to be a common result. Using a multivariate approach, such as logistic regression, provides a DNA band pattern library where fewer isolates cannot be matched, but also produces rates of correct classifications that are generally comparable to those obtained with antibiotic resistance analysis and carbon source profiles.

Method Testing

Pulsed-field gel electrophoresis is being used in our lab primarily as a cross-validation DNA fingerprinting method and is performed on 5 to 10 percent of the isolates from known and unknown sources that are source-classified in watershed studies by antibiotic resistance analysis. In comparison, tests where known source isolates of E. coli were treated as unknowns, the average rate of correct classifications between antibiotic resistance analysis, carbon source profiles, and pulsed-field gel electrophoresis, to date, were 90.3 percent for wildlife, 92.1 percent for livestock, and 93.6 percent for human isolates. Individually, antibiotic resistance analysis and carbon source profiles have higher levels of agreement with pulsed-field gel electrophoresis than with each other. In comparison tests with stream (unknown source) E. coli isolates, the three methods correctly identified the isolates as being from predominantly human, livestock, or wildlife origin at stream sites chosen for their obvious contamination by one type of source (Table 4).

*Table 4. Method Comparisons for E. coli Isolates from Sites Chosen for Source of Contamination**

Source	ARA	PFGE	CSP	Average
Wildlife	84.1%	88.4%	86.7%	86.4%
Livestock	82.6%	87.5%	84.3%	84.8%
Human	93/2%	96.3%	94.5	94..7%
Average	86.6%	90.7%	88.5%	88.6%

*Based on 192 isolates for ARA, and 50 each for PFGE and CSP.
ARA = Antibiotic resistance analysis.
CSP = Carbon source profiles.
PFGE = Pulsed-field gel electrophoresis.

Conclusions

The goal of our laboratory is to develop and test bacterial source tracking methods in a cost-effective manner, and then provide these methods as options to agencies and communities that are involved in projects where source tracking would prove beneficial. Experience in some 23 total maximum daily load projects in Virginia indicates that antibiotic resistance analysis, pulsed-field gel electrophoresis, and carbon source profiles all work reasonably well and the source tracking results made sense for the watersheds where they were used.

References

Bowman, A.M., C. Hagedorn, and K. Hix (2000). "Determining sources of fecal pollution in the Blackwater River watershed." *Abstracts, Virginia Water Research Symposium 2000, VWRRC Special Report SR-19-2000*, T. Younos and J. Poff (eds.), Blacksburg, Virginia.

Hagedorn, C., S.L. Robinson, J.R. Filtz, S.M. Grubbs, T.A. Angier, and R.B. Reneau, Jr. (1999). "Using antibiotic resistance patterns in the fecal streptococci to determine sources of fecal pollution in a rural Virginia watershed." *Applied and Environmental Microbiology*, 65: 5,522-5,531.

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." *Journal of Environmental Quality*, accepted (July-August issue).

Harwood, V.J., J. Whitlock, and V.H. Withington. (2000). "Classification of the antibiotic resistance patterns of indicator bacteria by discriminant analysis: Use in predicting the source of fecal contamination in subtropical Florida waters." *Applied and Environmental Microbiology*, 66: 3,698-3,704.

Holmes B., M.M. Costa, S.L.W. Ganner, and O.M. Stevens (1994). "Evaluation of Biolog system for identification of some gram-negative bacteria of clinical importance." *Journal of Clinical Microbiology*, 32: 1,970-1,975.

McClellan, P., V.O. Shanholtz, B. Petrauskas, and J. Kern (2000). "Bacterial source tracking: A tool for total maximum daily load development." *Abstracts, Virginia Water Research Symposium 2000, VWRRC Special Report SR-19-2000*, T. Younos and J. Poff (eds.), Blacksburg, Virginia.

Simmons, G.M., S.A. Herbein, and C.M. James (1995). "Managing nonpoint fecal coliform sources to tidal inlets." *Water Resource Update*, 100: 64-74.

Simmons, G.M., D.F. Waye, S. Herbein, S. Myers, and E. Walker (2000). "Estimating nonpoint fecal coliform sources in Northern Virginia's Four Mile Run watershed." *Abstracts, Virginia Water Research Symposium 2000, VWRRC Special Report SR-19-2000*, T. Younos and J. Poff (eds.), Blacksburg, Virginia.

U.S. Environmental Protection Agency (1999). Total maximum daily load (TMDL) program. [Online] http://www.epa.gov/owowwtr1/tmdl/tmdlfs.html.

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." *Applied and Environmental Microbiology*, 62: 3,997-4,002.

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). "Identification of sources of fecal pollution using discriminant analysis: Supporting evidence from large datasets." *Applied and Environmental Microbiology*, 65: 3,483-3,486.

CHARLES HAGEDORN has taught courses on soil and aquatic microbiology in the Department of Crop and Soil Environmental Sciences at Virginia Polytechnic Institute and State University since 1986. He has conducted research for the past 27 years in both academia and the private sector, and has published 67 journal articles, been awarded over $2.6 million in competitive grants, and received several outstanding service awards from state and federal agencies and professional organizations. His expertise has been sought by the U.S. Environmental Protection agency to serve on their Biotechnology Science Advisory Committee, by the Virginia Center for Innovative Technology on their Biotechnology Institute Advisory Committee, and as a former member of the European Biotechnology Safety Committee. Current research interests include bacterial source tracking, alternative on-site wastewater technologies, and the fate and transport of microbes in soil and water. Hagedorn received a B.S. in Biology from Bridgewater College in Virginia, and both an M.S. and Ph.D. in Microbiology from Iowa State University.

Coliphage Tracking to Identify Sources of Fecal Contamination

Mark D. Sobsey
Professor, Department of Environmental Sciences
University of North Carolina

Coliphages and Their Taxonomy

Coliphages are viruses infecting *Escherichia coli* bacteria. There are two main groups of coliphages: somatic and male-specific (Figure 1). Somatic coliphages infect host bacteria by attaching directly to the outer cell wall (outer cell membrane). The male-specific or F+ coliphages infect only male strains of bacteria by attaching to the hair-like appendages projecting from the cell surface, called pili or fimbrae, that are the characteristic male trait.

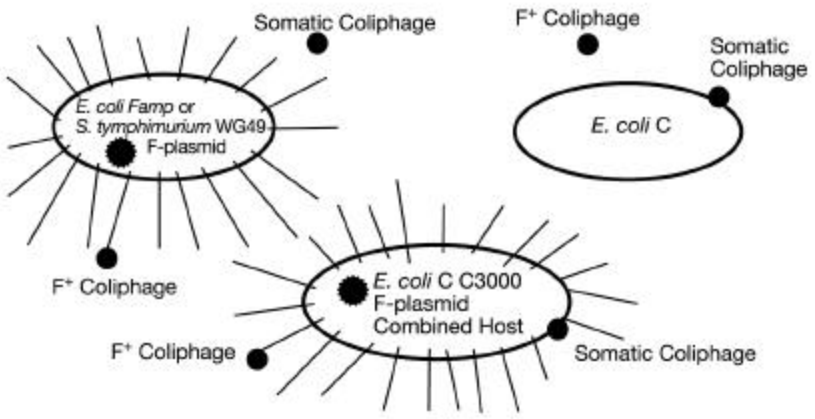

Figure 1. Somatic and Male-speific Coliphages and Their Typical Host Bacterai

The somatic and male-specific coliphages belong to several taxonomic groups of bacteriophages (Figure 2). There are only two main taxonomic groups of male-specific coliphages (Leviviridae and Inoviridae), and four main groups of somatic coliphages (Myoviridae, Styloviridae, Podoviridae, and Microviridae). All of these coliphage taxonomic groups are non-enveloped, but otherwise their morphology and properties vary.

The male-specific or F+ coliphages are either Leviviridae or Inoviridae. The Leviviridae are small (~25 nanometers [nm] diameter), non-enveloped, and contain single-stranded RNA; therefore, they superficially resemble many of the human enteric viruses (enteroviruses, caliciviruses,

Correspondence should be addressed to:

Mark D. Sobsey, Ph.D.
Professor, Department of Environmental Sciences
University of North Carolina
CB# 7400, McGavran-Greenberg Building, Room 4114a
Chapel Hill, NC 27599-7400
Tel: 919-966-7303 Fax: 919-966-7911 E-mail: MarK_Sobsey@unc.edu

hepatitis A and E viruses, and astroviruses). The Inoviridae are helical (filamentous), non-enveloped viruses, about 7 x 1,000 to 2000 nm, containing single-stranded, circular DNA. Both groups of F+ coliphages are present at high concentrations in sewage, with typical concentrations of 100 to 10,000 infectious units milliliter (mL). F+ RNA coliphages usually predominate, but both groups are present.

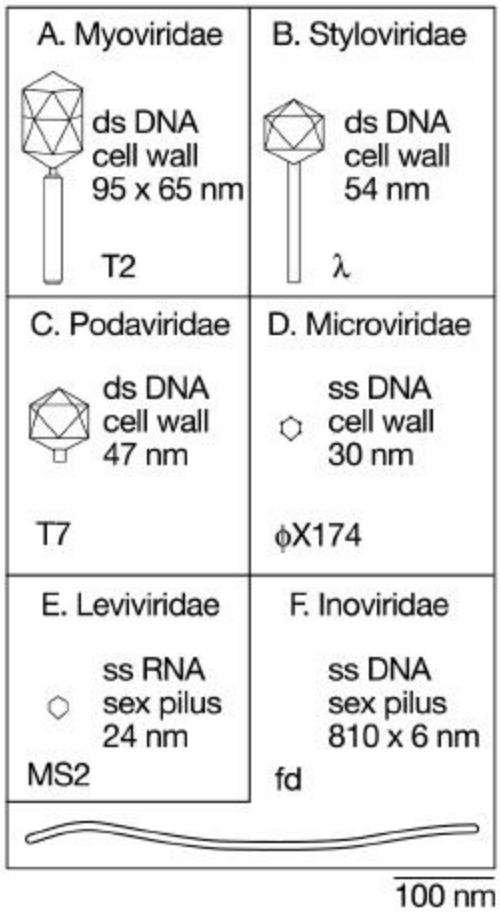

Figure 2. Taxonomy and Properties of Coliphages

The somatic coliphages, which include Myoviridae, Styloviridae, Podoviridae, and Microviridae, are all non-enveloped and contain DNA. Three of the groups (Myoviridae, Styloviridae, and Podoviridae) have tails for attaching to host cells and contain double-stranded DNA. The coliphages of the fourth group, the Microviridae, are small and icosahedral; they contain single-stranded, circular DNA. All groups of somatic coliphages are present in sewage, and the overall concentrations of somatic coliphages are similar to those of F+ coliphages at 100 to 10,000 infectious units per liter. The relative concentrations of the four different groups in somatic coliphages in wastewater and sewage are not well characterized, but all four groups have been detected in sewage and fecally contaminated water.

Source Tracking and Identification Using Coliphages

Of the various coliphages present in sewage and fecally contaminated water, source tracking has been based on detecting the presence of the four sub-groups of the F+ RA coliphages, the

Leviviridae. Although there are genetic differences between and within the other taxonomic groups of coliphages, these differences have not been sufficiently well characterized or employed to distinguish among or discriminate sources of coliphages in feces, sewage, water, or other environmental media. Currently, research is in progress to determine the genetic properties and variability of F+ DNA coliphages and whether or not these differences are consistently associated with different F+ DNA coliphage sources; however, this research is in its early stages. Further investigation is needed.

There are four subgroups of F+ RNA coliphages: Groups I, II, III, and IV. These groups can be distinguished by antigenic differences using serological techniques (neutralization of virus infectivity by group-specific antisera) or by genetic differences using gene probes (hybridization with oligonucleotide probes). A number of studies have reported that F+ RNA coliphages of:

- Group I are present in both human and animal fecal contamination and sewage.
- Groups II and III are predominantly or exclusively associated with human fecal contamination and domestic or municipal sewage.
- Group IV are predominantly associated with animal fecal contamination or animal sewage.

Hence, it is possible to broadly distinguish human from non-human animal fecal contamination based on the presence and prevalence of the different groups of F+ RNA coliphages. A number of independent studies have been done that document the ability of F+ RNA coliphage grouping to detect and distinguish between human and animal fecal contamination in fresh and marine waters. Distinguishing human from animal fecal contamination by grouping F+ RNA coliphages appears to be reliable when only one of these two main sources is present or predominates. When both human and animal sources of fecal contamination are present in water, the ability to determine the relative contributions of both sources is uncertain at this time. This is because some studies have reported differential survival in water of the different sub-groups of F+ RNA coliphages in water at higher temperatures (about 15°C or higher). Hence, if one group dies off faster or slower than another, a quantitative estimation of source strength and relative contribution may not be possible. Further studies are needed to determine the extent to which such differential die-off of F+ RNA coliphages may interfere with source identification and tracking.

Based on currently available genetic information and analytical methods, it is not possible on a practical basis to further resolve F+ RNA coliphages into genogroups or other sub-groups that identify specific animal fecal contamination sources. It is not yet possible to identify animal fecal contamination on the basis of animal type (swine, cattle, geese, etc.) and to more conclusively identify human fecal contamination as distinct from that of animals based using F+ RNA coliphage grouping; however, research is in progress and more is in the planning stages to better characterize F+ RNA coliphages based on genetic differences and to determine if such differences are sufficiently characteristic to distinguish a specific animal host of origin and, thereby, more conclusively identify sources of fecal contamination in water and other environmental media.

Detection, Quantification, and Isolation of Coliphages in Water and Other Media

The typical methods to detect coliphages in water and other samples are based upon the ability of

the coliphages to infect host bacteria, which results in the intracellular proliferation of the coliphages and lysis of the host bacteria. This a widely used approach to detect coliphages. The lysis of the host bacteria is visualized as a zone of lysis or clearing of the bacteria as a discrete, circular area (called a lysis zone or plaque) in a confluent layer (or "lawn") of host bacteria in a solid nutrient medium. The medium is solid because it contains agar.

The vast majority of coliphages detected using recommended standard *E. coli* hosts and assay conditions infect only *E. coli* and no other host bacteria. The range and characteristics of coliphages detected using currently recommended hosts and assay conditions are relatively well defined. Coliphages are present at high concentrations in sewage and other fecal wastes, and they are indicators of fecal contamination in water and other environmental media, such as biosolids and foods. The conventional method to detect coliphages is by their ability to infect host cells in which they replicate (proliferate), producing large numbers of progeny viruses and lysing (killing) the host cells in the process. It is this killing and lysis of host cells that forms the basis of most virus infectivity assay methods, including those employed for coliphage source tracking. Because some host bacteria are infected only by somatic coliphages, others only by male-specific coliphages, and yet others by both groups of coliphages, the choice of host bacterium is critical for detecting the appropriate coliphages for source tracking.

Double Agar Layer Plaque Assay. The so-called double agar layer (DAL) plaque assay method has been widely used for enumerating coliphages and other bacteriophages. In this method, a sample volume of <1 to about 5 mL is supplemented with host bacteria, then combined with molten agar medium and poured into a Petri plate containing a bottom layer of agar medium. After the top agar layer containing the sample and host bacteria hardens, the plate is incubated at 37°C overnight for the development of plaques. Plaques are counted and the coliphage concentration is computed as plaque-forming units (PFU) per unit volume of sample. A major limitation of the DAL method is the relatively small sample volume assayed per plate. The application of this method to environmental waters often results in non-detects (negative results for the sample volume assayed) because of the relatively low concentration of coliphages in the sample.

Methods for Coliphage Detection in Large Sample Volumes. Because the DAL method cannot be easily adapted to detecting coliphages in large (100 mL or more) volumes of water and other media, large volume coliphage assay methods have been developed. These large volume coliphage detection methods include filter adsorption-elution methods, U.S. Environmental Protection Agency (EPA) Method 1601 (two-step enrichment) and U.S. EPA Method 1602 (single agar layer). Typical sample volumes are up to 1,000 mL for the filter adsorption-elution and enrichment methods and 100 mL for the single agar layer (SAL) method. Use of these methods improves the detection of coliphages in water when their concentrations are relatively low.

Filter Adsorption-elution Methods. Two main forms of filter adsorption-elution methods are available. In both forms, the coliphages are adsorbed to membrane or other adsorbent, microporous filters, usually by facilitating coliphage adsorption to the filter with the addition of magnesium ions (in the form of magnesium chloride) to the water sample prior to filtration. In the liquid elution form of the method, the adsorbed coliphages are eluted from the filter with a small volume (5 to 10 mL) of aqueous medium, such as beef extract, and the eluate is assayed for coliphages by the DAL plaque assay or an alternative method. In the direct membrane filter method, cellulose membrane filters with adsorbed coliphages are placed face down on the surface of plates containing agar medium and host *E. coli* cells. The coliphages desorb (elute) from the cellulose membrane, and transfer to the agar medium, where they infect host cells and form

plaques after overnight incubation. The filter methods have the advantage of detecting coliphages in sample volumes of up to about 1,000 mL, and the coliphages are detected as discrete units (plaques) for easier purification and subsequent identification.

U.S. EPA Method 1601 for Two-step Enrichment. U.S. EPA Method 1601 is a so-called two-step "enrichment" method in which coliphages infect *E. coli* in broth medium, and the presence of the coliphages in the broth medium is confirmed by a second analytical step. In the first step of this method, liquid bacterial medium, magnesium chloride (to promote coliphage attachment to the host bacteria), and the *E. coli* host are added to the water sample, making a liquid (broth) culture for coliphage infection of the *E. coli* host bacteria. The enrichment culture is incubated overnight for coliphage infection and lysis of the host bacteria. As the second step, a small volume (several microliters) of the enrichment culture is placed as a spot on the surface of a Petri dish of agar medium containing *E. coli* host bacteria. If the applied sample contains coliphages able to infect the host bacteria, a circular zone of host cell lysis (clearing) develops after several hours of incubation in the spot where the sample was applied. Such a lysis zone in the spot indicates coliphage presence in the enrichment broth and is a positive result. If no such lysis zone develops in the sample spot on the plate, the enrichment culture of the sample is considered negative for coliphages. The appearance of a zone of lysis in the enrichment spot indicates coliphage presence. When U.S. EPA Method 1601 is applied to a single sample volume, the analysis determines the presence or absence of coliphages in the sample volume analyzed. If the method is applied to multiple sample volumes, each in separate enrichment cultures, the method is capable of giving an estimation of the concentration of coliphages in the water sample, based on which sample enrichment volumes become positive and negative for coliphages. A limitation of the enrichment method for coliphage source tracking is that more than one type of coliphage may be present in the same enrichment culture. It has proven difficult to separate mixtures of coliphages for their subsequent identification. This is because two or more coliphages present as a mixture in the enrichment are also present in the lysis zone spot from which coliphages are picked for subsequent identification; however, a number of studies have reported that the enrichment method is more sensitive than other methods, such as the DAL, SAL, and filter methods, in detecting low levels of coliphages in water and other media. Therefore, coliphages can be detected in water by this method when other methods fail to detect them

U.S. EPA Method 1602 for Single Agar Layer Plaque Assay. U.S. EPA Method 1602 is a so-called single agar layer method for the enumeration of coliphage plaques (discrete clear zones of lysis of host bacteria) developing in a culture of host bacteria in an agar medium in a Petri dish. A 100-mL sample of water is supplemented with magnesium chloride and the host bacteria, and then combined with molten agar medium. The mixture is then distributed into several Petri plates, the agar medium is allowed to solidify, and the plates are incubated overnight for the development of coliphage plaques, which are clear, circular zones of lysis, each produced by a separate or individual coliphage. The plaques are then counted to determine the total number of coliphages in the sample, assuming each plaque arose from an infectious coliphage. The SAL method has the advantage of detecting coliphages in sample volumes of up to about 100 mL, and the coliphages are detected as discrete units for easier purification and subsequent identification.

References

Grabow, W.O.K., and P. Coubrough (1986). "Practical direct plaque assay for coliphages in 100-mL samples of drinking water." *Applied and Environmental Microbiology*, 52(3): 430-433.

Guelin, A (1948). "Etude quantitative de bacteriophages typhiques. VI. Dans les eaux." *Annales de l'Institut Pasteur*, 79: 447.

Hsu, F.C., T.R. Handzel, G. Lovelace, J.R. Stewart, M.D. Sobsey, S.S. Thompson, and M.V. Yates (1998). "Improved methods to detect low levels of coliphages in water by enrichment presence-absence and membrane filter methods." *Proceedings, Water Quality Technology Conference Proceedings*. American Water Works Association, Denver, Colorado.

Hsu, F.C., Y.S.C. Shieh, J. Van Duin, M.J. Beekwilder, and M.D. Sobsey (1995). "Genotyping male-specific RNA coliphages by hybridization with oligonucleotide probes." *Applied and Environmental Microbiology*, 61(11): 3,960-3,966.

Sobsey, M.D., D.A. Battigelli, T.R. Handzel, and K.J. Schwab (1995). *Male-specific coliphages as indicators of viral contamination of drinking water*. American Water Works Association Research Foundation, Denver, Colorado.

U.S. Environmental Protection Agency (2001). "Method 1601: Male-specific (F+) and Somatic Coliphage in Water by Two-Step Enrichment Procedure." EPA 821-R-98-030. Office of Water, Washington, DC.

U.S. Environmental Protection Agency (2001). "Method 1602: Male-specific (F+) and Somatic Coliphage in Water by Single Agar Layer Procedure." EPA 821-R-98-029. Office of Water, Washington, DC.

MARK SOBSEY joined the faculty at University of North Carolina in 1974. Currently, he is Professor of Environmental Microbiology in the Department of Environmental Sciences and Engineering, School of Public Health at the University of North Carolina. He has extensive research, teaching, and service experience in environmental health microbiology and virology, with more than 250 published papers and reports. His research focuses on the detection, characterization, occurrence, environmental survival/transport/fate, treatment, health effects, and risk assessment of viruses, bacteria, and protozoans of public health concern in water, wastewater, biosolids, soil, air, and food. Sobsey is a scientific advisor to the World Health Organization, U.S. Environmental Protection Agency, the State of North Carolina, and other international, national, and state entities. He received both a B.S. in Biology and a M.S. in Hygiene from the University of Pittsburgh, and a Ph.D. in Environmental Health Sciences from the University of California at Berkeley.

Source Tracking of Fecal Waste Material in Environmental Waters Using a Biomarker Based on Enterotoxin Genes in *E. coli*

Betty H. Olson, Leila A. Khatib, and Eunice C. Chern
Department of Environmental Analysis
University of California, Irvine

Robin Oshiro
Office of Water
U.S. Environmental Protection Agency

Increasing pressures upon surface-water supply from population growth and the implementation of total maximum daily load standards set by the U.S. Environmental Protection Agency through the amendments to the Clean Water Act of 1972 forced agencies to assess point and non-point waste sources entering watersheds. On the one hand, regions containing both agricultural and urban areas often find fecal coliforms of little value in establishing the identity of fecal sources in their watersheds, and may have difficulty meeting the 200 colony-forming units (CFU)/100 milliliter (mL) fecal coliform standard. During the early 1980s, although pathogens could be identified, no better means than fecal coliforms existed to identify animal or human waste sources that could pose health threats to humans. By the early 1990s, it was clear that fecal source identification could be an extremely useful tool to regulatory and water agencies.

The microbiological- and molecular-based methods proposed to identify the fecal sources in water, to date, form the basis of this workshop. Before describing the toxin biomarker approach to fecal source differentiation, we will briefly describe the criteria that my laboratory believes are important for reliable and adequate source differentiation (Tables 1 and 2). Table 1 shows the

Table 1. Optimum Characteristics of Viable Source Differentiation Methods

Level of Performance	Specificity	Sensitivity	Geographical Stability	Temporal Stability	Sampling for Identification	Age of Pollution	Viability of Target Organism
Optimum	Identifies the source correctly 100 percent of the time	Identifies the trait when it is rare in the popula-tion tested or at low concen-trations	One data base can be used through the United States	The identifying characteristic does not change over time-years and is endemic	All representa-tives of a population can be screened	Target organism has a defined survival time in the environ-ment	Pathogens viable/ infective

Correspondence should be addressed to:

Betty H. Olsen, Ph.D.
Professor, School of Social Ecology
University of California, Irvine
1368 Social Ecology II, Irvine, CA 92697-7070
Tel: 949-824-5281 Fax: 949-824-2056 E-mail: bholson@uci.edu

Table 2. Optimum Characteristics of Source Differentiation Method for Users

Level of Performance	Certainty of Negatives and Unknown	Simplicity of Technique	Cost of Technique	Ease of Transferring Technique to Other Laboratories	Ease of Automation
Optimum	None or very well defined criteria	Can be done by microbiologists with minimal training	Utilizes equipment that is relatively inexpensive	Any laboratory could carry out the technique	Technique can be modified to handle large number of samples

characteristics of the accuracy a method should have, while Table 2 contains information that is useful to users from regulatory or water agencies that might want to use a technique. All of the methods tested to date work with varying degrees of specificity, sensitivity, and geographical and temporal stability for the populations screened. We have tried to develop a method that fits as many of these characteristics (Tables 1 and 2) as possible.

Because we wanted a simple positive or negative test, which also allowed us to query the entire *E. coli* population in a sample, we selected a method that would not involve fingerprinting or ribotyping. The enteropathogenic *E. coli* appeared ideal because their toxins were described both phenotypically and genotypically, and we had isolated a number from treated and raw drinking waters (Martens et al., 1993). The biomarker method is based on the theoretical premise that several of the enterotoxigenic *E. coli* and avian pathogenic genes carried by *E. coli* are host specific. The distribution of these genes amongst animals and humans is shown in Table 3. The host-specificity of these toxins is determined by the associated receptor recognition genes (colonization factors). Twenty different colonization factors genes have been identified in *E. coli* containing STh and a number of others in pigs and cattle. The host receptor genes must be present for the strain to be pathogenic. Each host group (cattle, pigs, humans, birds, etc.) has its own colonization factors, which allows for host specificity. Although the data shown in Table 3 seems contradictory, as STh has been reported in buffalo, LTIIa in humans, and STII in humans, dogs, and cattle, a closer examination of that literature shows that STII, LTIIa, and STh have been isolated infrequently (only once) from other than the target organism.

For example, in Table 3, STII has been reported to have been isolated from four hosts. We found only one report in the literature for each of those hosts. STII was reported isolated from 0.75 percent of humans in one Japanese study. Similar results were found for STII isolated from dogs and cattle. Both are reported in only one article each, and the percentage of hosts positive was very low with the maximum being 4.4 percent for diarrheic cattle. Although not perfect, the method held the promise of being species-specific. We developed primers to three traits (STh-human, STII-pig, and LTIIa-cow) and began to screen for cross-reactivity with *E. coli* isolated from a number of species.

Given this information, we screened STII as a biomarker pig waste, STh as a biomarker for human waste, and LTIIa as a biomarker for cattle waste. We needed to establish that the traits were species-specific, endemic, and occurred in a high enough frequency to be detected using polymerase chain reaction (PCR). A total of 172, 30, 139, 108, and 12 fecal samples and E. coli

isolates from farm animals, domestic animals, birds, humans, and wild animals, respectively, were screened for cross-reactivity using STh. A total of 214 fecal samples and isolates from the

Table 3. Enterotoxigenic E. coli and Avian Pathogenic E. coli Occurrence in Humans, Animals, and Birds

Toxin	STI		STII	LTI		LTII		TSH
Classification of Toxins	Heat Stable Toxins			Heat Labile Toxins				Associated with the colician V plasmid
	Methanol soluble, assayable in infant mice assay		Methanol insoluble, detection. Detection in ileal Loop of pigs	Antigenically distinct from LTII		Antigenically distinct from LTI		
Alleles and (Alternative Nomenclature)	STIa (StaP, STp	STIb (StaH STh)	STb	LTp (LTIp)	LTh LTIh	LTIIa Genes a and b	LTIIb Genes a and b	
Animals Affected	human, cattle, pig, sheep, dog, buffalo	human, buffalo	human, cattle, pig, dog	cattle Pig	human chicken dog	cattle, water buffalo human pig	humans	Exotic birds, sea gulls, mallards

various categories stated above were screened for cross-reactivity with pig and cow primers. The bird primers are currently being tested.

Each trait was examined in field tests to validate the method. Ten sewage treatment plants from California, varying in size from 15 to 2 million gallons per day (mgd), were sampled one to 10 times over a yearlong period. DNA to be tested was extracted from centrifuged raw, primary, and secondary waste effluents from publicly owned treatment works. Single PCR was used. Cattle and pig samples were filtered and grown on mTEC agar according to Dufour et al. (1981). Additionally, seven septic tank wastes from the Midwest were tested also using single PCR and direct extraction. Thirty-three dairies and ranches, and 33 pig farms were used to validate the cow and pig biomarkers, respectively. Farms were located throughout the United States. Cattle samples came as far away as New York and Hawaii and pig samples from North Carolina. Seasonal samples over a 1-year period were tested using single PCR.

Results showed that 95 percent of sewage treatment effluent samples were positive, and 88 percent of cattle and 96 percent of pig samples were positive. It appeared from the data that all farms and treatment plants were positive, but our method was not sensitive enough to detect the trait at low prevalence rates. We decided to use nested PCR to increase our sensitivity. The results from a year's sampling at four cattle farms using nested PCR were 100-percent positive; in addition, seven pig samples from various farms outside of California were all positive as well as stream samples downstream of these farms. The human samples from plants ranging from 242 to 1.3 mgd were 90-percent positive.

We also carried out a three-dilution, five-replicate MPN test to determine the prevalence of the biomarker traits in the E. coli populations using nested PCR (Table 4), subjecting DNA base extracts from these samples to nested PCR. A few samples changed from negative to positive, but

the most dramatic change was the increase in prevalence of the traits when nested PCR was used. The prevalence of the traits varies from farm to farm and from sewage treatment plant to sewage treatment plant as well as seasonally; however, traits are always present. In seeded environmental

Table 4. Prevalence of LTIIa, STII, and STh Biomarkers in the Total E. coli Population in Each Sample from Spring 1999 to Spring 2000

	LTIIa-cow	STh-human	STII-pig
Average	A:1,626	1:614	1:5
Median	1:44	1:224	1:6
Geometric Mean	1:89	1:220	1:9
Range	1:11,364 to 1:0.6	1:4,600 to 1:19	1:3 to 1:19

water samples, we were able to detect 0.1 target cells/liter because our control strain has 15 to 20 copies of LTIIa per cell. We also found dramatic ranges of prevalence of the trait in stream waters ranging from 1:2 to 1:542 for STII. We have detected LTIIa-cow biomarker when total *E. coli* concentrations in stormwaters was in excess of >2,400/mL. From prevalence work we have developed, ranges of *E. coli* screened can be used to frame negative results, such as zero pig traits were found amongst 10,000 *E. coli* (hypothetical example).

Although nested PCR was a big improvement over single PCR, we felt further improvements in sensitivity or detection could be attained. To achieve this, we used magnetic bead hybridization. Magnetic bead hybridization eliminates extraneous DNA and greatly increased PCR efficiency. Using Single PCR, we were able to detect the trait LTIIa-cow using 2.5 attogram (ag) of total DNA from the hybridizations. With magnetic bead hybridization, we were able to detect 1 to 1.5 traits of LTIIa. Using nested PCR, we could detect 1.5 to 2 traits of LTIIa and STII, while single PCR detected ag quantities of the traits. One trait has a molecular weight of approximately 10 to 19 grams.

Recently, we have begun testing an avian pathogenic *E. coli* marker to identify bird waste. This test looks promising with many types of birds testing positive, but we are in the early phases of development. We also have identified toxins for both the cat and dog and will be testing those soon. The level of performance for all biomarker tests is shown in Table 5.

Table 5. Level of Performance on the Biomarker Method

Level of Performance	Specificity	Sensitivity	Geographical Stability	Temporal Stability	Sampling for Identification	Age of Pollution	Viability of Target Organism
	cow 100% pig 100% human 95% bird 100%	10^{19g}g	Positive across U.S. cow pig human bird- Southern California	cow, pig human endemic Bird?	Entire E. coli community screened. Sub-samples not required	1 to 3 weeks in environmental waters, depending on temperature, light, and predation	Viable-growth or nonviable-direct extraction

We feel that we have developed a simple test to differentiate human and animal waste in watersheds. This test is highly sensitive, finding one target *E. coli* amongst 1,000,000 non-target *E. coli*. Many of the current methods would not detect such a rare occurrence; therefore, this could be important in tracking intermittent human waste pollution.

References

Dufour, A.P., and V. Cabelli (1981). "Membrane filter method for enumerating Escherichia coli." *Applied and Environmental Microbiology*, 41: 1,152-1,155.

Martens, M.T., I.G. Rivera, D.L. Clark, M.H. Stewart, R.L. Wolfe, and B.H. Olson (1993). "Distribution of uidA gene sequences Escherichia Coli isolates in water sources and comparison with the expression of b-glucuronidase activity in 4-methylumbelliseryl-b-D-glucuronide media." *Applied and Environmental Microbiology*, 59: 2,271-2,276.

A water-quality specialist, BETTY OLSON is both Professor in the School of Ecology and Professor in the School of Medicine at the University of California, Irvine, where she has taught and conducted research since 1980. She is the author or co-author of over 146 publications on microbiology and biotechnology, has received numerous awards and honors for her research, and has received over $3 million in funding since 1988 for projects with the U.S. Department of Agriculture, California Department of Water Resources, and National Water Research Institute, among others. Olson received a B.S. in Biology (emphasis in Microbiology) from the University of California, Irvine, and both an M.S. in Environmental Health Science and Ph.D. in Environmental Health Science (minors in Ecology and Microbiology) from the University of California, Berkeley.

Detection of Enteroviruses Using PCR-based Techniques for Source Identification and Assessment of Microbiological Water Quality

Rachel T. Noble
Assistant Professor, Institute of Marine Science
University of North Carolina

Introduction

Microbiological water quality of the coastal zone is currently determined by using bacterial indicators, such as total coliforms, fecal coliforms, and enterococci, to infer the presence of microbial pathogens. The use of bacterial indicators is based upon the assumption that they are present in waters with fecal contamination and not in those without; however, the use of bacterial indicators is limited in that they sometimes fail to successfully predict the presence of all types of pathogens (bacterial, viral, and protozoan). Studies have revealed that several dangerous virus types can be contracted by swimming in contaminated ocean waters (Cabelli et al., 1982; Seyfried et al., 1985a, 1985b; Haile et al., 1999), and that outbreaks of gastroenteritis have been caused by swimming in water with acceptable coliform counts (Cabelli et al., 1982). Also, bacterial indicators are not always a useful predictor of the presence of human fecal contamination. Gerba et al. (1979) showed that bacterial indicators fell short of predicting the presence of enteroviruses. Both Noble and Fuhrman (2001) and Jiang et al. (2001) present the lack of a strong predictive relationship between the presence of enteroviruses and adenoviruses as related to currently used bacterial indicators, as determined by the polymerase chain reaction (PCR) for detecting human pathogens.

In the last decade, a number of new molecular and microbiological techniques have been applied to the field of coastal water quality. These methods have largely been used to provide further information on the presence of pathogens and/or to separate sources of fecal contamination as either "human" or "non-human," as fecal indicator bacteria (and fecal contamination) can come from a variety of different sources (dogs, cats, livestock, birds, humans). Furthermore, some of the indicator bacteria can naturally be found in freshwater and marine environments, making it difficult to link sources with high indicator results. Some of the newer approaches offer rapid and reliable results and might, in the future, be useful for day-to-day management decisions. Some of the methods are "database oriented," where a watershed's fecal contamination sources are characterized over a period of time to provide information for mitigation actions, such as total maximum daily loads. All of the methods are still considered research methods and are constantly being changed and optimized to increase the accuracy, sensitivity, and rapidity of results. The aim of this abstract will be to introduce and review the theory, application, and results of using PCR to identify sources of fecal contamination, and to place the use of this method into the larger context of "source-tracking methods." The author wholeheartedly believes in a "toolbox approach," combined with careful delineation of the particular goal or question at hand, in using these methods.

Correspondence should be addressed to:

Rachel T. Noble, Ph.D.
Assistant Professor, Institute of Marine Science
University of North Carolina
3431 Arendell Street, Morehead City, NC 28557
Tel: 252-726-6841 E-mail: rtnoble@email.unc.edu

Source Tracking Methods

The following methods have been recently applied to source identification:

- Ribotyping (Parveen et al., 1999).
- Pulse-field gel electrophoresis (Parveen et al., 2001).
- PCR (Pina et al., 1998).
- Reverse transcription-PCR (RT-PCR) (Tsai et al., 1993).
- Quantitative-PCR (Q-PCR) (Monpoeho et al., 2000).
- Determination of antibiotic-resistance patterns (ARP) (Hagedorn et al., 1999; Harwood et al., 2000).
- Terminal restriction fragment length polymorphism analysis (Bernhard and Field, 2000).
- F-specific coliphage analysis (Hsu et al., 1995).

These methods all have distinct advantages and limitations. Each of these methods differ in regard to cost, analysis time, required training, sampling logistics, quantitation, specificity, reproducibility, and geographical applicability. As the methods used for source identification are relatively new, comparative tests between methods need to be performed to better understand the limitations and advantages inherent to each method. In addition, it is vital to use the appropriate statistical approach to analyze the data from each method.

PCR-based Methods

PCR is a primer-based molecular biology technique, which can be used to detect specific groups or types of viruses based upon complementarity of primer sequences with conserved sequences in the genomes of bacteria, viruses, and protozoans. The primers are typically 17 or 18 base pairs in length and designed using known sequence information. PCR offers high sensitivity and reproducibility, and it permits the screening of a large number of samples at one time. PCR is used for detecting and amplifying those organisms that contain DNA, whereas RT-PCR is used for detecting RNA (the RT step involves the transcription of the RNA into a cDNA copy). For example, Tsai et al. (1993) outlined a method for successfully detecting enteroviruses and hepatitis A in seawater using RT-PCR. In the past, PCR-based methods have often been considered costly because of the amount of training required to perform them in a standardized fashion; however, improvements in sample concentration, extraction, and sample cleanup (to remove PCR inhibitors found in environmental samples) have occurred over the past few years. It should be remembered that the interpretation of PCR methods is limited to detecting specific types of indicators and pathogens in environmental samples.

Quantitative PCR

Q-PCR is a novel primer-based molecular technique that combines the specificity of "traditional" PCR with the quantitative measurement of fluorescence for determining the presence of specific types of nucleic acid in environmental samples. For example, Monpoeho et al. (2000) used Q-PCR to quantify the 5' noncoding region of enteroviruses in sewage sludge. Currently, there are two commonly used applications for Q-PCR: TaqManTM or Molecular BeaconsTM (Tyagi and Kramer, 1996). One real-time PCR technique in use today is the TaqManTM, or 5'-nuclease assay, which employs a dual-fluorescent labeled oligonucleotide probe to quantitatively measure the accumulation of target molecules during each cycle of PCR. The oligonucleotide probe bears a "reporter" fluorescent dye at the 5'-end and a second "quencher" fluorescent dye at the 3'-end. The 3'-dye is chosen to absorb light at a longer wavelength than the emission spectra of the 5'-

reporter dye. This configuration allows for fluorescence resonance energy transfer to occur and fluorescence of the 5'-dye is "quenched" by the 3'-dye. The sequence of the oligonucleotide probe is chosen to be complementary to a site within the desired PCR amplicon or nucleic acid sequence of interest. During the course of the PCR reaction, the probe anneals to its target sequence and is subsequently degraded by the 5'-nuclease activity of Taq DNA polymerase. This physically separates the reporter dye from the quencher dye and releases the reporter dye from quenching of the fluorescence, allowing the fluorescence emission of the reporter to be detected and quantitatively measured (Figure 1).

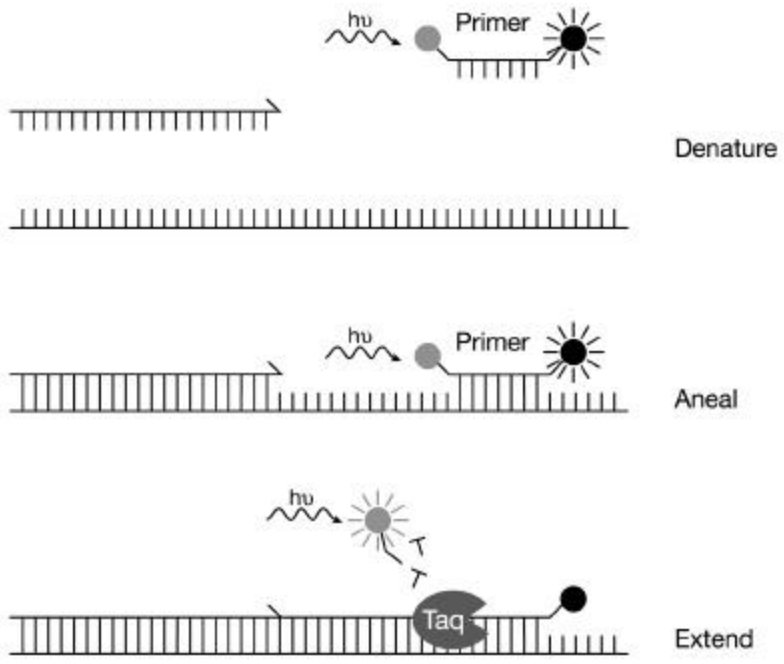

Figure 1. Schematic representation of real-time PCR with TaqMan™ primers. In the intact TaqMan™ probe, energy is transferred from the short-wavelength fluorophore on one end (grey circle, on left end of primer) to the long-wavelength fluorophore on the other end (black circle, on right end of primer), quenching the short-wavelength fluorescence. After hybridization, the probe is susceptible to degradation by the endonuclease activity of a processing Taq polymerase. Upon degradation, the fluorescence from the short-wavelength fluorophore is increased and the fluorescence from the long-wavelength fluorophore is decreased. Diagram structured after those at www.idtdna.com.

Another quantitative technique used to measure levels of a target sequence employs dual-labeled oligonucleotide probes of a different design called Molecular BeaconsTM. Molecular BeaconsTM bear a 5'-fluorescent reporter dye and a "dark" quencher group in the 3'-position. The probe has a central domain designed to specifically hybridize to a unique target sequence. Short sequences are added to the 5'- and 3'-ends, which are unrelated to the target but which allow the probe to form a hairpin. In this configuration, the 5'-reporter dye is brought into proximity of the 3'-Dabcyl and is quenched. When the probe hybridizes to its target sequence, this hairpin structure is disrupted and the 5'-reporter is physically separated from the 3'-quencher, allowing fluorescence emission to be detectable and measured in a quantitative fashion (Figure 2).

Q-PCR offers speed and "real-time" quantification, allowing for the detection of particular DNA or RNA sequences with specific probes that fluoresce only when the target sequence are present, thus removing the need to run a gel (Heid et al., 1996). Samples are prepared with reagents in a single tube, similar to standard PCR as mentioned above (Monpoeho et al., 2000). To make Q-PCR fully quantitative, we have developed an internal standard, which is necessary for determining the efficiency of amplification by PCR. Our currently used internal standard is Poliovirus Sabin Strain 1, enumerated by epifluorescence microscopy and used in our Q-PCR reactions. We are working to develop other internal standards for RT-PCR that involve the use of a T7 promoter to create RNA transcripts of the same gene amplified with our PCR primers. For DNA viruses, we will create plasmids from a suitable cloning vector. The internal standards will then be added, in specific amounts, to the PCR reaction, along with fluorophore-labeled primer/probe combination (of a different color than for the analysis reaction itself) for amplification. A standard graph of the threshold amplification values obtained from the serially diluted internal standard is then constructed, plotted on a standard curve, and the copy number is calculated by the Q-PCR software.

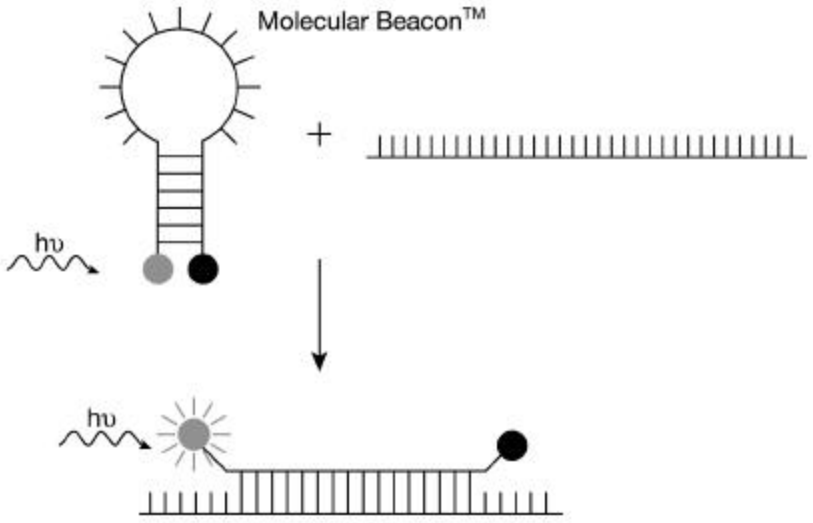

Figure 2. Schematic representation of the Molecular BeaconTM. In the hairpin loop structure, the quencher (black circle, on right of hairpin structure) forms a nonfluorescent complex with the fluorophore (grey circle, on left of hairpin structure). Upon hybridization of the Molecular BeaconTM to the complementary sequence, the fluorophore and quencher are separated, restoring the fluorescence. Diagram structured after those at www.idtdna.com.

Application of PCR

In the late 1990s, our research in the field of water quality focused on the use of traditional PCR for detecting specific types of viral and bacterial pathogens in seawater and shellfish harvesting waters (e.g., Jiang et al., 2001; Noble and Fuhrman, 2001). Our research demonstrated that PCR was an effective method for detecting human pathogenic virus genomes. The optimization of methods for sample filtration, cleanup, and nucleic acid extraction have been important components of these studies and others (e.g., Monpoeho et al., 2000; Atmar et al., 1993). For example, in 50 coastal seawater samples (Noble and Fuhrman, 2001), low correlations were seen

between enteroviruses and total coliforms, fecal coliforms, and enterococci (r values ranged from 0.14 to 0.34), suggesting that bacterial indicators are not necessarily good predictors of the presence of human pathogens in coastal waters. We suggest that a direct analysis for viral pathogens might be advisable at certain locations (e.g., high-use beaches) or during certain seasons of the year, even when bacterial indicator levels appear to be low. We are currently optimizing Q-PCR techniques for the analysis of environmental samples taken from stormwater-impacted beaches on both the Atlantic and Pacific coasts, and we are working towards detecting other viral pathogens via Q-PCR for source identification of fecal contamination.

Conclusions

We have reviewed the theory and application of PCR-based methods for source tracking of fecal contamination. Other methods that have not been mentioned, but are currently being developed, include the analysis of toxin biomarker genes, randomly amplified polymorphic DNA analysis, DNA microarrays, and plasmid profiling. In addition, there are other chemical methods available (coprostanols, caffeine), and new molecular methods are being developed all the time. It is apparent that there is a wealth of resources available to federal, state, and local regulatory agencies to determine sources of fecal contamination in coastal waters. Some of these tools are the best options for future implementation of total maximum daily loads in coastal watersheds. As mentioned, there is no single source identification method that provides the answer to every question. There have been some issues in the past that bring up the question of whether molecular-based methods are better than classic microbiological methods. It is important to remember that whether genotypically, phenotypically, or quantitatively based, all of the methods reviewed have appropriate uses, and the best demonstration of their successful use is through the publication of the method in a peer-reviewed journal. Of particular importance for the successful application of a certain method is the careful delineation of the scientific question at hand.

References

Atmar, R.L., T.G. Metcalf, F.H. Neill, and M.K. Estes (1993). "Detection of enteric viruses in oysters by using the polymerase chain reaction." *Applied and Environmental Microbiology*, 59(2): 631-635.

Bernhard, A.E. and K.G.A Field (2000). "PCR Assay to discriminate human and ruminant feces on the basis of host differences in Bacteroides-Prevotella genes encoding 16S rRNA." *Applied and Environmental Microbiology*, 66(10): 4,571-4,574.

Cabelli, V.J., A.P. Dufour, M.A. Levin, and L.J. McCabe (1982). "Swimming-associated gastroenteritis and water quality." *American Journal of Epidemiology*, 115(4): 606-616.

Gerba, C.P., S. M. Goyal, R. L. LaBelle, I. Cech, and G. F. Bodgan (1979). "Failure of indicator bacteria to reflect the occurrence of enteroviruses in marine waters." *American Journal of Public Health*, 69(11): 1,116-1,119.

Hagedorn, C., S.L. Robinson, J.R. Filtz, S.M. Grubbs, T.A. Angier, and R.B. Runeau (1999). "Determining sources of fecal pollution in a rural Virginia watershed with antibiotic resistance patterns in fecal streptococci." *Applied and Environmental Microbiology*, 65(12): 5,522-5,531.

Haile, R.W., J.S. Witte, M. Gold, R. Cressey, C. McGee, R.C. Millikan, A. Glasser, N. Harawa, C. Ervin, P. Harmon, J. Harper, J. Dermand, J. Alamillo, K. Barrett, M. Nides, and G.Y. Wang

(1999). "The health effects of swimming in ocean water contaminated by storm drain runoff." *Epidemiology*, 10(4): 355-363.

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." *Applied and Environmental Microbiology*, 66(9): 3,698-3,704.

Heid, C.A., J. Stevens, K.J. Livak, and P.M. Williams (1996). "Real time quantitative PCR." *Genome Research*, 6(10): 986-994.

Hsu, F.C., Y.S.C. Shieh, J. Van Duin, M.J. Beekwilder, and M.D. Sobsey (1995). "Genotyping male-specific RNA coliphages by hybridization with oligonucleotides probes." *Applied and Environmental Microbiology*, 61(11): 3,960-3,966.

Jiang, S.J., R.T. Noble, and W.P. Chu. (2001). "Human adenoviruses and coliphages in urban runoff-impacted waters of Southern California." *Applied and Environmental Microbiology*, 67: 179-184.

Monpoeho, S., A. Dehee, B. Mignotte, L. Schwartzbrod, V. Marechal, J.C. Nicolas, S. Billaudel, and V. Ferre (2000). "Quantification of enterovirus RNA in sludge samples using single tube real-time RT-PCR. Short tech. reps." *BioTechniques*, 29: 88-93.

Noble, R T., and J.A. Fuhrman (2001). "Enteroviruses detected in the coastal waters of Santa Monica Bay, CA: Low correlation to bacterial indicators." *Hydrobiologia*, 460: 175-184.

Parveen, S., N.C. Hodge, R.E. Stall, S.R. Farrah, and M.L Tamplin (2001) "Phenotypic and genotypic characterization of human and nonhuman Escherichia coli." *Water Research*, 35(2): 379-386.

Parveen, S., K.M. Portier, K. Robinson, L. Edmiston, and M. Tamplin (1999). "Discriminant analysis of ribotype profiles of Escherichia coli for differentiating human and nonhuman sources of fecal pollution." *Applied and Environmental Microbiology*, 65(7): 3,142-3,147.

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." *Applied and Environmental Microbiology*, 64(9): 3,376-3,382.

Seyfried, P.L., R.S. Tobin, N.E. Brown, and P.F. Ness (1985a). "A prospective study of swimming-related illness. I. Swimming-associated health risks." *American Journal of Public Health*, 75: 1,068-1,070.

Seyfried, P.L., R.S. Tobin, N.E. Brown, and P.F. Ness (1985b). "A prospective study of swimming-related illness. II. Morbidity and the microbiological quality of water." *American Journal of Public Health*, 75: 1,071-1,074.

Tsai, Y., M.D. Sobsey, L.R. Sangermano, and C.J. Palmer (1993). "Simple method of concentrating enteroviruses and hepatitis A virus from sewage and ocean water for rapid detection by reverse transcriptase polymerase chain reaction." *Applied and Environmental Microbiology*, 59(10): 3,488-3,491.

Tyagi, S., and F.R. Kramer (1996). “Molecular BeaconsTM - Probes that fluoresce upon hybridization.” *Nature Biotechnology*, 14(3): 303-308.

RACHEL NOBLE is an Assistant Professor at the University of North Carolina, Institute of Marine Sciences. She recently held a joint appointment between the University of Southern California’s Wrigley Institute for Environmental Studies and the Southern California Coastal Water Research Project as a post-doctoral scientist, and focused her work on assessing water quality along the Southern California shoreline and detecting enteroviruses in stormwater-impacted areas of the coast. Currently, her research focuses on the use of molecular techniques, such a Quantitative Polymerase Chain Reaction (Q-PCR), for identifying sources of fecal material in estuarine and coastal environments, and for use in assessing microbiological water quality. Her primary interest is the field of microbiological water quality, especially in regard to relating the presence of known human pathogens to human health risk of those using the waters for recreational or food consumption purposes. Noble received a B.S. in Molecular Biology from Carnegie Mellon University and a Ph.D. in Marine Biology from the University of Southern California.

Adenovirus as an Index of Human Viral Contamination

Sunny C. Jiang
Assistant Professor, Environmental Analysis and Design
University of California, Irvine

Current water-quality standards based on fecal indicator bacteria suffer two major limitations:

- They are not capable of distinguishing human fecal contamination from other sources.
- They inadequately represent the survival of viruses and other human pathogens that are more resistant than indicator bacteria to water treatment processes and environmental degradation conditions.

This study argues that using human viruses as an index of water quality offers the advantages of specific detection of human fecal source contamination and an accurate representation of the viral quality of water.

There are over 100 different human viruses associated with human waste, and all are potentially water transmissible. It is impractical, if not impossible, to detect all such viruses; however, certain viruses may adequately serve as an index for the presence of human fecal contamination and the overall viral quality of water. One potential index virus, adenovirus, is a double-stranded DNA virus, belonging to the family adenoviridae. Over 40 human adenovirus serotypes have been described. Generally, most adenoviruses cause respiratory infections, but they can also be found in the tonsils or adenoid tissue of healthy people. Adenoviruses are frequently found in urban rivers (Castingnolles et al., 1998; Chapron et al., 2000; Tani et al., 1995) as well as polluted coastal waters (Pina et al., 1998; Puig et al., 1994; Jiang et al., 2001). Studies conducted in Europe have suggested using adenovirus as an index of human viral pollution since this virus was often detected in human fecal contaminated environmental samples (Pina et al., 1998). Adenovirus infections occur year-round, and there is little or no seasonal variation in shedding (Allard et al., 1990). They have been consistently found in greater numbers than enteroviruses in raw sewage around the world (Irving and Smith, 1981; Krikelis et al., 1985a, 1985b) and are substantially more stable than either poliovirus or hepatitis A virus in tap water and seawater (Enriquez et al., 1995); therefore, the objective of this study was to test if adenovirus could be used as an index for human viral contamination of water in Southern California.

Using nested and reverse-transcription (RT) polymerase chain reaction (PCR) methods, three types of human viruses, including adeno, hepatitis A, and enteroviruses, were assayed for in 11 Southern California rivers and creeks between July 10, 2000, and August 30, 2000. RT-PCR results were further confirmed using probes specific for hepatitis A and enteroviruses, respectively. Eighteen of the 21 sites examined were positive for at least one type of human virus (Table 1). A seasonal investigation of human viruses in coastal waters at the mouths of the Los Angeles, San Gabriel, and Santa Ana Rivers by repetitive sampling indicated that human viruses

Correspondence should be addressed to:

Sunny C. Jiang, Ph.D.
Assistant Professor, Environmental Analysis and Design
University of California, Irvine
Irvine, CA 92697-7070
Tel: (949) 824-5527 Fax: (949) 824-2056 E-mail: sjiang@uci.edu

Table 1. Detection of Human Pathogenic Viruses in Southern California coastal Rivers

Sampling Sites	Adenovirus	Enterovirus	Hepatitis A Virus
Malibu Creek I	—	—	+
Malibu Creek II	—	—	—
Ballona Creek I	+	+	+
Ballona Creek II	—	—	+
Los Angeles River I	+	+	+
Los angeles River II	—	—	+
San Gabriel River I	+	+	+
San Gabriel River II	+	+	—
San Ana River I	—	—	—
San Diego Creek I	—	+	+
San Diego Creek II	+	+	+
Aliso Creek I	+	+	+
Aliso Creek II	+	+	+
San Juan Creek I	—	—	+
San Juan Creek II	—	—	+
San Luis Rey River I	+	+	+
San Luis Rey river II	—	+	+
San Diequito River I	+	+	+
San Diequito River II	+	+	+
San Diego River I	+	+	—
San Diego River II	—	—	+

Table 2. Seasonal Investigation of Human Pathogenic viruses
In Southern California Coastal Waters

Sampling Dates	Adenovirus	Enterovirus	Adenovirus	Enterovirus
San Gabriel River	Point Zero		Surf Zone	
10/10/00	—	+	—	+
10/24/00	—	+	—	+
11/17/00	—	+	—	+
01/25/01	+	—	—	—
02/08/01	—	—	—	—
02/21/01	—	—	—	—
03/01/01	—	+	—	+
05/03/01	—	—	—	—
05/17/01	—	—	—	—
05/24/01	—	—	—	—
05/31/01	—	—	—	—
Santa Ana River	Point Zero		Surf Zone	
10/04/00	—	+	—	—
10/17/00	—	—	—	—
11/01/00	—	—	—	+
01/25/01	—	—	—	—
02/10/01	—	—	—	—
02/14/01	—	—	—	—
02/21/01	—	—	—	—
Los Angeles River	Point Zero			
01/25/01	—	+		
02/15/01	+	+		
02/12/01	—	+		
02/21/01	—	—		
02/27/01	—	+		
02/27/01	—	—		
04/17/01	—	—		
05/01/01	—	—		
05/01/01	—	—		
05/15/01	—	—		

were more often detected in the winter than in the spring (Table 2), suggesting that rain fall from winter storms may transport viruses from urban rivers into coastal oceans.

Although adenoviruses were detected in the majority of samples collected from urban waterways, entero- and hepatitis A viruses were also found at locations where adenoviruses were absent, suggesting that

adenoviruses alone cannot serve as an index for human viral contamination in this region. It is suggested that more than one human virus should be tested for to better reflect the contamination source and viral quality of the water. Comparisons of the presence of human viruses with the presence of standard indicator bacteria and somatic and F-specific coliphage indicated a poor correlation between the presence of a human viral signal and the concentration of current indicator bacteria, suggesting that indicator bacteria may not accurately reflect the viral quality of water.

In summary, a PCR-based assay frequently detected human viruses in Southern California urban waterways, suggesting the prevalence of human fecal contamination in urban runoff; however, the assay used in this study was unable to discern inactivated viruses from infectious ones. Therefore, the results presented here may be an overestimation of the risk of contracting a viral disease from Southern California urban rivers.

References

Allard, A., B. Albinsson, and G. Wadell (1992). "Detection of adenoviruses in stools from healthy persons and patients with diarrhea by two-step polymerase chain reaction." *Journal of Medical Virology*, 37: 149-157.

Castingnolles, N., F. Petit, I. Mendel, L. Simon, L. Cattolico, and C. Buffet-Janvresse (1998). "Detection of adenovirus in the waters of the Seine River estuary by nested-PCR." *Molecular and Cellular Probes*, 12: 175-180.

Chapron, C.D., N.A. Ballester, J.H. Fontaine, C.N. Frades, and A.B. Margonlin (2000). "Detection of astroviruses, enteroviruses, and adenovirus types 40 and 41 in surface waters collected and evaluated by the information collection rule and an integrated cell culture-nested PCR procedure." *Applied and Environmental Microbiology*, 66: 2,520-2,525.

Enriquez, C.E., C.J. Hurst, and C.P. Gerba (1995). "Survival of the enteric adenoviruses 40 and 41 in tap, sea, and wastewater." *Water Research*, 29: 2,548-2,553.

Jiang, S. C.R. Nobel, and W. Chu (2001). "Human adenoviruses and coliphage in urban runoff-impacted coastal waters of Southern California." *Applied and Environmental Microbiology*, 67: 179-184.

Irving, L.G. and F.A. Smith (1981). "One-year survey of enteroviruses, adenoviruses, and reoviruses isolated from effluent at an activated-sludge purification plant." *Applied and Environmental Microbiology*, 41: 51-59.

Krikelis, V., N. Spyrou, P. Markoulatos, and C. Serie (1985a). "Detection of indigenous enteric viruses in raw sewage effluents of the City of Athens, Greece, during a two-year survey." *Water Science and Technology*, 17: 159-164.

Krikelis V., N. Spyrou, P. Markoulatos, and C. Serie (1985b). "Seasonal distribution of enteroviruses in domestic sewage." *Canadian Journal of Microbiology*, 31: 24-25.

Pina, S., M. Puig, F. Lucina, 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." *Applied and Environmental Microbiology*, 64: 3,376-3,382.

Puig, M, J. Jofre, F. Lucena, A. Allard, G. Wadell, and R. Girones (1994). "Detection of adenovirus and enteroviruses in polluted waters by Nested PCR amplification." *Applied and Environmental Microbiology,* 60: 2,963-2,970.

Tani, N., Y. Dohi, N. Kurumatani, and K. Yonemasu (1995). "Seasonal distribution of adenoviruses, enteroviruses, and reoviruses in urban river water." *Microbiology and Immunology*, 39: 577-580.

SUNNY JIANG has been Assistant Professor at the University of California, Irvine since 1998. She has appointments in both the Department of Environmental Analysis and Design and the Department of Chemical and Biochemical Engineering and Materials Science. In addition, she was appointed as Associate Professor of the College of Life Sciences at Nankai University in Tianjian, China, in 2001. Her current research focuses on the investigation of the causes of coastal water contamination and the ecology of bacteria and viruses in the aquatic environment. Among her honors, Jiang received the Environmental Project Award from the Environmental Professionals Organization in 2001. She received a B.S. in Biochemistry from Nankai University, and both an M.S. and Ph.D. in Marine Science from the University of South Florida.

Application of Methods to Identify Coliform Pollution Sources Using Multiple Antibiotic Resistance , Selected Molecular Techniques and GIS Spatial Analysis

Geoffrey I. Scott
Brian C. Thompson
Laura F. Webster
Jan A. Gooch-Moore
Jill R. Stewart
NOAA/NOS, Center for Coastal Environmental Health and Biomolecular Research, Charleston, SC

Jill R. Stewart
School of Public Health, University of North Carolina, Chapel Hill, NC

David Chestnut
South Carolina Department of health and Environmental Control, Columbia, SC

Robert F.Van Dolah
South Carolina Department of Natural Resources, MRRI, Charleston, SC

Heath Kelsey
Dwayne Porter
School of Public Health., University of South Carolina, Columbia, SC

Furman Cantrell
Clemson Institute for Environmental Toxicology, Clemson University, Clemson, SC

An overview and summary of the methods presented at the EPA workshop in Irvine , CA will be presented along with case studies form coastal regions of the southeastern U.S. Discharges of wastewater from sewage treatment plants (STPs), septic tanks, farm animal operations (FMOs), urbanization and wildlife pollution sources may adversely affect estuarine water quality, often closing shellfish beds for harvesting and downgrading water quality classification in rivers and streams. Development of methods for differentiating human versus wildlife coliform bacterial sources is needed to properly manage bacterial pollution emanating from different sources. Methods presented at the workshop will be summarized. Case studies will emphasize several methods for differentiating human and wildlife coliform bacterial sources were evaluated including Multiple Antibiotic Resistance (MAR), Pulsed Field Gel Electrophoresis (PFGE), and Ribotyping (RT). Water samples were collected from several river and estuarine watersheds in SC and selected pollution sources (STPs, septic tanks, FMOs, and wildlife). Samples were enumerated for fecal coliform bacterial densities (MPNs or MF) and *E. coli* were isolated

by API biotyping . Samples were then analyzed by MAR, PFGE, and RT. Adjoining land use in several areas was further analyzed by GIS and multivariate statistics to predict significant land use metrics affecting fecal coliform densities and to identify human pollution sources. Results indicated that the % of *E. coli* comprising the coliform group and MAR was highest at sewage treatment plants and in urban areas adjoining sites with septic tanks or influenced by sewer discharges. Wildlife areas had negative MARs or resistance to only a single antibiotic and a lower % of *E. coli*. PFGE and RT provided DNA differentiation of bacterial pollution sources. Multivariate statistics and GIS provided methods to locate human pollution sources, identify land metrics affecting coliform MPNs and quantify presumptive Total Maximum Daily Load estimates of fecal coliform sources in shellfish harvesting areas. These findings indicate that these methods may be helpful in identifying different sources of fecal coliform bacteria.