

## **Appendix E**

### **Description of Microbial Source Tracking Methods**

**As developed by Dr. Mansour Samadpour**

## Microbial Source Tracking

**Note: The following is excerpted from a project proposal by Dr. Mansour Samadpour:**

During the past decade tremendous advances have been made in developing rapid sensitive microbial pathogen detection systems. Agencies such as the United States Department of Defense, Department of Energy, and Department of Agriculture have funded a large number of projects for rapid, automated detection of microbial pathogens and indicators in various matrices, including water, wastewater, and food. The biggest gap in knowledge and methodologies remains in the area of identification of the sources of microbial pollution. While the field of microbial source tracking is still in its infancy, advances in this field are needed to elevate the field of environmental microbiology to its next level and focus the efforts and resources toward control of sources of pollution.

...

During the past decade several methods have been proposed for identification of the sources of microbial pollution in the environment. Currently there are several research groups that conduct source tracking and source identification studies, each using a different method and different target organisms. The methodologies that have been used to determine the sources of microbial contamination in the environment range from the use of phenotypic based methods such as antimicrobial resistance profiles (Wiggins, 1996), to genotypic based methods including ribotyping (Parveen et al. 2000), macrorestriction fingerprinting using pulsed field gel electrophoresis (Edberg et al, 1994), and polymerase chain reaction based methods (Dombek et al. 2000).

For the past eleven years we have worked on developing approaches and methods that would allow for identification of the sources of microbial contamination in the environment. The work has led to development of the "Microbial Source Tracking" (MST) method. The MST method relies on a specific sampling plan designed on the basis of a sanitary survey of the watershed of interest, and the types of questions that are to be answered. The source identification portion of the method relies on generating genetic fingerprints of *Escherichia coli* strains isolated from the contaminated site(s) and comparing of the fingerprints to those of *E. coli* strains isolated from potential sources of pollution. The method that is currently in use, in our laboratory, for generating the genetic fingerprints of the isolates for the MST studies is ribosomal RNA typing using two restriction enzymes (Eco RI and Pvu II). To date we have subtyped more than 65,000 *E. coli* strains during the course of our studies.

...

### Background

Numerous human pathogens are spread by fecal contamination of water. Examples include *Vibrio cholera*, *Salmonella typhi*, *Giardia lamblia*, *Cryptosporidium parvum* and Hepatitis A. These pathogens can be a risk to human health even at very low concentrations. Due to difficulties in the detection, identification, and enumeration of specific human pathogens in environmental and food samples, the concept of indicator organisms and related methodologies were developed and implemented in the late 1800's. Indicator organisms are used to assess the potential for the presence of pathogens due to fecal contamination. These organisms must be prevalent in feces, found in higher concentrations than pathogens, be more resistant to disinfectants, more persistent in the environment, and they must be easy to quantify. The group of bacteria referred to as fecal coliforms meet these criteria. Fecal coliforms are facultative anaerobic bacilli that ferment lactose

with the production of gas within 48 hours at a temperature of 44.5°C. A prevalent and well-studied member of this group is *Escherichia coli*.

The concept of indicator organisms is the principal component of regulatory microbiology. The major limitation of this concept is that it is an oversimplification of the complex dynamics of microbial ecology, physiology, and genetics. The utility of the indicator organism concept is further limited by its inability to track organisms associated with fecal contamination to their potential sources. Each year millions of dollars are spent on fecal and total coliform assays to determine the extent of bacterial and fecal pollution of aquatic environments and to satisfy increasingly rigid regulatory requirements concerning the microbiological quality of water. Knowing the sources rather than just monitoring the level of microbial pollution of source water enables water quality professionals and watershed managers to design and implement programs control pollution and protect source water.

The inability to conclusively identify the contributing sources of microbial contamination in watersheds has led to an over-reliance on treatment processes to insure a safe supply of drinking water. In many instances, the lack of effective source control programs has resulted in a deterioration of the microbiological quality of source waters, which in turn results in an increased likelihood of waterborne outbreaks of gastroenteritis in instances of treatment failure.

Until a few years ago, the identification of nonpoint sources of microbial pollution was an impossible task. However, advances in molecular biology and molecular epidemiology have resulted in the development of molecular subtyping methods that can be used to assess the impact of suspected sources of microbial pollution in rivers, lakes, and water reservoirs. Once the sources of microbial pollution are identified, appropriate control measures can be devised to reduce or eliminate their impact.

### **Principles of Microbial Source Tracking**

Bodies of water are impacted by large numbers of sources of microbial pollution in their watersheds. In a given watershed, potential sources of microbial pollution include soil, vegetation, and the entire human and animal population residing in the watershed. Determination of the sources of microbial pollution in a watershed is not an easily accomplished task. It requires establishing a large collection of bacterial isolates of a specific species from the impacted body of water that is representative of the genetic diversity of that bacterial species in the watershed. Identification of the sources of the microbial pollution is then achieved by subtyping the water isolates and matching the subtypes to a collection of bacterial isolates of the same species from known sources, which include humans and various animal species.

...

Our laboratory's MST method has been developed on the basis of the principles of microbiology, epidemiology, molecular epidemiology, microbial population genetics, sanitary engineering, and hydrogeology. 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, 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.

### **Underlying Assumptions of Microbial Source Tracking**

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 descendants 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 shows that using our subtyping method (ribosomal RNA typing using two restriction enzyme reactions) more than 96% of *E. coli* strains are seen in only one host species (or group of related species) (Mazengia, 1998).

Given that bacterial population structure is clonal and within each bacterial species different clones have adapted to specialized environments, it should be possible to:

- Study a collection of bacterial isolates from a contaminated site (e.g. receiving water) and from possible sources of contamination
- Divide the isolates into groups of clonal origin
- Match the isolates from the contaminated site to their sources
- Identify the contributing sources

### **Subtyping Methods Used in Source Tracking Studies**

Another important factor in determination of clonality is the methodological issues. 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 8 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 sanitary survey of the study area which is representative of the population of a given species in a watershed would require the analysis of thousand 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 fields of population genetics, molecular systematics, and molecular epidemiology. In 1986 Grimont, et al. showed that DNA probes corresponding to specific regions of the rRNA operon can be used to speciate bacteria. Stull, et al. (1988) and Lipuma, et al. (1988) used the rRNA operon to study the molecular epidemiology of several species of bacteria. In order to trace the indicator bacterium, *E. coli*, from the water to its specific source, the bacterial strain must first be uniquely identified. Populations of *E. coli*, like other bacteria, are composed essentially 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.

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, 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.

The pattern of DNA fragments corresponding to the rRNA operon is referred to as the ribotype. Ribotyping has been useful in many studies to differentiate between bacterial strains that would have otherwise been difficult or impossible to distinguish. Fisher, et al. (1993) followed the transmission of *Pseudomonas cepia* from environmental sources to and between cystic fibrosis patients and discovered the majority of cases contracted cystic fibrosis from one of two treatment centers. Moyer, et al. (1992) used ribosomal RNA typing to identify the *Aeromonad* strains responsible for several waterborne gastroenteritis episodes in a community and was able to trace the contamination to specific locations in water treatment and distribution systems. Baloga and

Harlander (1991) compared several typing methods for distinguishing between strains of *Listeria monocytogenes* implicated in a food-borne illness and found that ribotyping was the preferred method due to its precision and reproducibility. Atlas and Saylor (1988) described the technology of ribotyping as applicable to the tracking of genetically engineered microorganisms (GEMs) in the environment.

...

### Ribotyping Using *Eco* R1 and *Pvu* II

Our initial source tracking studies were all conducted using a single enzyme (*Eco* R1) ribotyping protocol. The choice of *Eco* R1 was the result of a large scale screening of enzymes to determine the differentiative power of each of the available enzymes. *Eco* R1 showed the most differentiative power followed by *Pvu* II. Although in the beginning the single enzyme system was showing close to 100% residency among our source isolates, as our database grew and more isolates were studied we began to notice a sharp increase in the incidence of transients among the source isolates. At that time, we hypothesized that a single enzyme lacks the sensitivity to effectively separate the isolates into groups of clonal origin. We then conducted a study to test this hypothesis.

Table 1. Collection of *E. coli* strains which were used to evaluate the use of single versus double enzymes for ribotyping.

Source Type	No. of Isolates	Total Ribotypes <i>Pvu</i> II	Total Ribotypes <i>Eco</i> R1	Total Ribotypes <i>Eco</i> R1 / <i>Pvu</i> II
Human Sources <sup>1</sup>	813	265	348	381
Bovine	325	48	92	153
Horse / Llama	342	43	66	104
Avian <sup>2</sup>	183	55	72	107
Canine <sup>3</sup>	194	38	53	72
Feline <sup>4</sup>	73	22	30	33
Deer / Elk	53	18	14	21
Farm Animals <sup>5</sup>	100	25	36	41
Wild Animal <sup>6</sup>	59	13	21	25
<b>TOTAL</b>	<b>2142</b>	<b>527</b>	<b>732</b>	<b>873</b>

Table 2. Grouping of the 2142 *E. coli* strains by each of the two restriction enzymes and by the combination of the two enzymes.

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%)

*Pvu* II divided the 2142 *E. coli* strains from various sources into 514 groups. 49% of the groups were resident clones and 51% were transient. *Eco* R1 divided the 2142 isolates into 723 groups, 56% of which were resident clones and 44% were transient. When we combined the results of the two enzymes we found that the 2142 isolates were divided into 873 clonal groups, 96% of which were resident clones and only 4% were transient. This was very convincing evidence that lead us to move towards a two enzyme ribotyping system. In order to increase the level of specificity of our source identification, we took an additional step of identifying and tagging the transient clones in the database. The transient clones are not used for source identification. The practical implication of a single versus double enzyme ribotyping protocol is shown in Table 3. While the 14 isolates from 14 different sources are seen as one with *Pvu* II, *Eco* R1 separates them into 14 different groups, allowing for their use in source identification.

Table 3. Illustration of the advantage of the use of a double enzyme system over single enzyme ribotyping. While *Pvu* II identifies the 14 isolates as transient, *Eco* R1 separates them into 14 resident clones.

Isolate #	PVU II	ECO RI	Source Type
20977	Z	A	BOVINE
21610	Z	B	HORSE
20699	Z	C	DOG
12069	Z	D	CAT
21696	Z	E	SEPTAGE
12805	Z	F	SANITARY SEWER
13076	Z	G	BOBCAT
8450	Z	H	DUCK
21464	Z	I	LLAMA
14328	Z	J	RAW SLUDGE
1601	Z	K	DEER
21438	Z	L	PIG
21641	Z	M	GOAT
14781	Z	N	BEAVER

## **MST Approach**

We have developed the following approach in the MST studies conducted in our laboratory:

- A. Interview with stakeholders, watershed managers, and local agencies that have been monitoring water quality of the study area
- B. Sanitary Survey of the study area
- C. Determine questions to be answered by the study
- D. Design a sampling plan to answer the questions. The sampling plan is designed specifically to answer all the questions that are raised regarding the study site. The sampling plan is put together to reflect influences such as seasonality, storm events, landuse, recreational use, and regrowth. Another important element in sampling design is the total project budget.
- E. Field Work (collecting water and source samples according to the sampling plan). This is often performed by collaborators at the site. The water samples are collected and processed locally (in certified laboratories). Source samples are shipped directly to our laboratory.
- F. Processing of water samples (to determine levels of fecal coliforms/E. coli and isolate E. coli strains) by local laboratories with subsequent shipment of the plates containing the organisms to our laboratory. The source samples are directly shipped to our laboratory. They are processed upon arrival to isolate E. coli strains.
- G. Logging the samples and cultures drives from each sample in our sample and data logs (both hard copy and computer database). The samples and cultures are logged together with pertinent epidemiological information such as: Isolate number (our log number) study ID, provider sample ID, provider ID, sampling date, sampling site (complete address), and source type.
- H. Establishing pure cultures of E. coli from primary water and source plates. Verification of speciation. Freezing the cultures in or permanent collection of isolates.
- I. Subtyping of the isolate collection. We currently are using ribosomal RNA typing as using two restriction enzymes (Eco RI and Pvu II) as our subtyping method. On selective bases we use one or two additional restriction enzymes. Our long standing policy on the choice of subtyping method is that as soon as a better method becomes available we will subtype our collection with the new method and change our database accordingly with the results of the new method.

## **MST Data Analysis**

The subtyping data for each isolate is analyzed (please refer to the materials and methods section for details), and entered into the database (Microsoft Access). Our MST Database contains detailed information regarding the E. coli strains in our collection. The source E. coli in the database are divided into two categories. Resident Clones (RC), and the Transient Clones (TC). The RCs are defined as clones that are unique to a particular host species (human, cow, etc.), or a group of closely related host species (dogs and coyotes), TCs are defined as clones which are seen in more than one unrelated host species. Only RCs are used for source tracking (assignment of source to water isolates). Using the current regiment of subtyping methods, more than 96% of all the host isolates fall into the RC category. The database is constantly updated to insure that the TCs are tagged and are not used for source identification purposes. We have found that when using additional enzymes we can eliminate more than half of the TCs and change them into RCs.

## **MST Utility**

The data resulting from an MST study can be used in:

- Understanding the sources, distribution, and movement of microbial populations in watersheds, source waters, swimming beaches, fisheries resources, etc.
- Conducting risk and exposure assessment studies of the potential human health effects associated with the presence of microbial pollution in source waters
- Identifying human pathogens that have established reservoirs in watersheds
- Determining the impact of various types of land use on water quality
- Identification of the sources of microbial pollution and quantification of the impact of each source
- Designing and implementing source control programs
- Studying the effects of control measures
- Environmental litigation.

## Materials and Methods

**Bacterial strains and culture conditions.** Water and sewage grab samples will be processed by membrane filtration for fecal coliforms (**Standard Methods for the Examination of Water and Wastewater, 20<sup>th</sup> ed.**). After incubation at 44.5°C for 24 hours they will be read. Appropriate colonies will be chosen morphologically (round, blue, and flat) and streaked for isolation on MacConkey media, then incubated at 37°C for 24 hours. Fecal samples will be collected from representative animal species in the Duwamish Watershed. They will be streaked on MacConkey plates and incubated at 37°C for 24 hours. Non-mucoid colonies that fermented lactose on MacConkey will be re-streaked on Trypticase Soy Agar (TSA) plates. Five *E. coli*-type colonies per sample will be isolated. Biochemical analysis will be done to positively identify *E. coli*. Two *E. coli* strains from each water sample and one *E. coli* strain from each source sample will be added to our study collection. These isolates will be assigned an isolate number and stored in LB-15% glycerol freezing media at -70°C.

**Genomic DNA isolation and restriction endonuclease digestion.** Confluent growth will be scraped with a sterile flat-headed toothpick and suspended in 200 µl 50mM Tris, 50mM EDTA (pH 8.0). Then another 600 µl 50mM Tris, 50 mM EDTA will be added and the suspension will be mixed well by pipetting up and down. Next, 45 µl 20% sodium dodecyl sulfate (SDS) followed by 10 µl proteinase K (20 µg/ml; Pharmacia, Piscataway, NJ) will be added. They will be incubated at 40°C for 1 hour. Equal volume of phenol will be added, samples will be vortexed, then centrifuged for 5 minutes. The top layer will be extracted and an equal volume of chloroform will be added. The prep will be vortexed again, centrifuged, and extracted. Two and a half volumes absolute ethanol will be added; the DNA will precipitate out and be spooled onto a sterile glass capillary pipette. The DNA will be washed with a few drops of absolute ethanol, dried, and re-suspended in 50 µl dH<sub>2</sub>O.

Restriction endonuclease digestions will be set up using *Eco* RI and *Pvu* II (in separate, individual reactions), 10 u/µl (Boehringer Mannheim, GmbH, Germany) as instructed by the manufacturer and 2 µgr DNA. They will be incubated at 37°C overnight. The samples will be centrifuged and .5µl pure enzyme will be added. The samples will be re-incubated at 37°C for a minimum of three hours. They will be centrifuged again and 3 µl stop dye will be added.

**Gel electrophoresis and Southern Blot hybridization.** Samples will be run on a 0.8% agarose gel in 1X Tris-borate-EDTA at 22 volts and 17 milliamps, for 17 hours. λ *Hind* III will be used as a size standard along with an *E. coli* isolate with a distinct ribotype pattern designated 3915. The DNA fragments will be then transferred to a Nitran filter (Schleicher & Schuell, Keene, NH), baked at 80°C for one hour and probed. P32 labeled copies of *E. coli* ribosomal RNA will be made by extension of random hexanucleotide primers (Finberg, et al.) using Avian Myeloblastosis Virus reverse transcriptase (Stratagene, La Jolla, Ca) under conditions specified by the supplier. Hybridization will be done in 5X SSC (1X SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 0.1% SDS, 1mM EDTA, and 50% formamide at room temperature overnight. Salmon sperm DNA and blocking reagent, (Boehringer Mannheim GmbH, Germany) will be used to block non-specific binding. Three washes will be done with a solution of 2X SSC and .1% S.D.S., once at 25°C for 20 minutes and twice at 65°C for 20 minutes to wash off low-homology, non-specific binding. Blots will then be exposed with an intensifying screen to X-ray film (Kodak, Rochester, N.Y.) for 24 hours at -70°C. Two to three exposures will be done to ensure detection of all possible bands.

## **QUALITY ASSURANCE: Measurement & Data Acquisition**

Sampling QA/QC Water samples: Water samples will be collected using sterilized sampling bottles, by grab sampling method as described in **Standard Methods for the Examination of Water and Wastewater** (APHA, 1997). Samples will be delivered (on ice) to the Molecular Epidemiology Inc. Laboratory and analyzed within eight hours from collection. All the sample bottles will be labeled with sampling station identification number, sampling date and time, sample number, source identification number, and sampler's initials. All the sample information will be entered into the field log, and Chain of Custody forms. Both the sampler and the receiving laboratory will sign the Chain of Custody form.

Water samples will be analyzed by the mFC method (Standard Methods for Examination of Water and Wastewater, APHA, 1997). To ensure aseptic conditions, blank samples will be filtered to determine whether our filtering apparatus, dilution blanks, and other equipment are free of contamination by fecal coliforms. Prior to filtering each sample, a blank sample (containing only dilution water) will be filtered. This will allow testing of the sterility of our filter tower and dilution water. Following the filtering of the prescribed number of dilutions, a final blank sample will be filtered. This will allow us to determine whether our rinsing method between individual dilutions was adequate enough to prevent contamination from previous filtrations. After incubation the results will be entered into result forms, and the forms will be entered into a database. QC records will include positive and negative controls with each batch of water samples filtered, media preparation and performance characteristics documentation, incubator and waterbath temperature documentation. Method performance will be documented by performing duplicate samples and routinely confirming both positive and negative colonies at a frequency of one in twenty samples. Acceptable reproducibility range will be an RPD of 20% when cfu is in the countable range.

Source samples: Fresh animal fecal samples will be collected aseptically into sterile containers and delivered to the environmental microbiology laboratory, on ice. Animal fecal samples are only collected when they are positively identified as belonging to a given animal species. No more than three samples will be collected from the members of the same animal species from a given location. Only a single sample will be collected from an individual animal. All sample containers will be labeled with the following information: sample type, host species, sample date and time, sample location, and sampler's initials. All the sample information will be logged into the field log. After collection of the samples, samples are delivered to the lab where they will be given a sample number and will be logged into the permanent sample log.

**MST QA/QC** The goal of the MST project is to identify the sources of fecal coliforms that are present in water samples. Two types of samples will be received for this study: water and fecal samples. Our laboratory analysis includes:

- a. Sample arrival, and logging.
- b. Filtration and quantification of FC.
- c. Isolation and purification of *E. coli* strains from water and fecal samples.
- d. Growing pure cultures of *E. coli* strains for freezing (long term storage), and isolation of DNA.
- e. Restriction enzyme digestion and Agarose gel electrophoresis of DNA samples.
- f. Southern blot hybridization using radio labeled cDNA probe for rRNA genes.
- g. Exposure of autoradiograms.
- h. Analysis of the data.

### **Sample arrival and logging**

All samples upon arrival are inspected for damage to sample containers or microbiological plates, and signs of contamination. Sample identifiers are also checked against the Chain of Custody forms. Samples are logged into our log book noting the provider's sample identification number, provider ID, sample type, study ID, sample site, sample collection date and sample arrival date. Compromised samples are noted and appropriately discarded.

### **Isolation and purification of *E. coli* strains from water and fecal samples**

Fecal samples are plated on MacConkey agar and incubated at 35 ° C, overnight. The next day 3-5 lactose-fermenting, non-mucoid colonies are picked and replated on MacConkey agar for purification.

Five non-mucoid blue colonies picked from mFC plates corresponding to each water sample are plated on MacConkey agar for purification.

At this stage each of the colonies picked from a given sample bears the Sample ID number and an accession letter. A single, well isolated, non-mucoid colony is picked from each MacConkey plate and is plated on Tryptic Soy Agar. After overnight incubation at 35 ° C, each culture is tested by a spot indol test. Indol positive cultures are further tested for the ability to utilize citrate using Simmon's Citrate medium. Isolates with the combination of indologenesi s, and citrate non-utilization are identified as *E. coli* and are given isolate numbers. Appropriate positive and negative controls are incorporated when testing the biochemical reactions.

### **Growing pure cultures of *E. coli* strains for freezing (long term storage)**

A portion of each *E. coli* strain isolated, identified and retained from the samples will be stored at -80° C, in nutrient broth plus 15% glycerol.

### **Phenotypic and genotypic characterization of *E. coli* isolates**

All methodologies for characterizing *E. coli* isolates will follow standardized protocols and will have the following QC documentation:

All record entries will include the analyst's initials, and the date.

All reagents, media and buffers are prepared according to written and approved SOPs. Each batch prepared is tested for sterility as appropriate and undergoes a performance test with positive and negative controls or with the previous batch prior to use in production. Commercially available sensidiscs impregnated with antibiotics will be tested against standard strains to ensure conformance with current designations of "sensitive, intermediate, and resistant" status.

Each batch of enzymatic reactions is performed with a positive control strain and is performance checked in an analytical procedure (eg, electrophoretic gel).

Incubation, electrophoresis, PCR conditions are all standardized for each method and documented for each run. Documentation includes: agarose gel concentration and volume, buffers, pH, mA, V, and run times. Each methodology will have its own logbook to track the isolates included in each run as well as the appropriate controls and their performance.

Any runs with control reactions out of normal response range will be noted and corrective action taken. Corrective action may include repeating the procedure and will be documented in a corrective action log.

All record books will be audited monthly for completeness, and technicians will be involved in the audit review.

### **Isolation of DNA, restriction enzyme digestion and agarose gel electrophoresis of DNA samples**

Genomic DNA is isolated from each *E. coli* strain using a standard protocol. Every batch of restriction enzyme reaction contains two reactions with our positive control strain, which will be included on two lanes on each gel. Each agarose gel is assigned a number, and when more than one gel is run, the position of the first standard reference strain is changed in each gel (1<sup>st</sup> lane on the first gel, to the Nth lane on the Nth gel). After electrophoresis, gels are stained in ethidium bromide; the two gels are each stained in a single container. One of the two gels is placed in the same container. The corner of the gel with the higher number is clipped and the label for each gel is also transferred to the staining container. Each gel is then photographed and a hard copy of the print is labeled with a gel sheet (containing the isolates' ID number loaded on each lane, the enzyme used to cut the DNA, as well as the date, gel number, voltage, mA, gel strength, buffer strength, and electrophoresis time). This information is kept in the gel book.

### **Southern blot hybridization using radio labeled cDNA probe for rRNA genes**

Southern blotting is performed according to the protocol detailed in our SOP. After photography each gel is returned to the same staining container. Gels are denatured for Southern blotting in the same container. Each blotting apparatus is set in a separate container that is labeled with the gel number. Each membrane filter is labeled with the gel number, restriction enzyme designation, date, and technician's initials.

### **LITERATURE CITED**

APHA, AWWA, and WPCF. 1992. *Standard Methods for the Examination of Water and Wastewater*, 18<sup>th</sup> ed. American Public Health Association, Washington, D.C.

Atlas, R.M., G. Saylor, R.S. Burlage, and A.K. Bej. 1992. Molecular approaches for environmental monitoring of microorganisms. *BioTechniques* 12(5):706-717.

Baloga, A. O. and Harlander, S. K. 1991. Comparison of methods for discrimination between strains of *Lester monocytogenes* from epidemiological surveys. *Applied and Environmental Microbiology*. 57(8): 2324-2331

Dombek, P.E., Johnson, L.K., Zimmerley, S.T., and Sadowsky, M.J. 2000. Use of repetitive DNA sequences and the PCR to differentiate *Escherichia coli* isolates from human and animal sources. *Applied and Environmental Microbiology*. 66:2572-2577.

Krumperman, P.H. 1983. Multiple Antibiotic Resistance Indexing of *Escherichia coli* to Identify High-Risk Sources of Fecal Contamination of Foods. *Applied and Environmental Microbiology*. 46:165-170.

Fisher, M. C., LiPuma, J. J., Dasen, S. E., Caputo, G. C., Mortensen, J. E., McGowen, K. L., and Stull, T. L. 1993. Source of *Pseudomonas cepacia*: Ribotyping of isolates from patients and the environment. *Journal of Pediatrics*. 123: 745-747.

Grimont, F. and Grimont, P.A.D. 1986. Ribosomal ribonucleic acid gene restriction patterns as potential taxonomic tools. *Annales 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.

Harwood, V.J., Whitlock, J., and Withington V. 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*. 66:3698-3704.

Karch, H., Russmann, H., Schmidt, H., Schwarzkopf, A., Heesemann, J. 1995. Long-term shedding and clonal turnover of enterohemorrhagic *Escherichia coli* O157:H7 in diarrheal diseases. *Journal of Clinical Microbiology*. 33(6):1602-1605

Kaspar, C.W. and Burgess, J.L, Knight, I.T., and Colwell, R.R. 1990. Antibiotic Resistance Indexing of *Escherichia coli* to Identify Sources of Fecal Contamination in Water. *Canadian Journal of Microbiology*. 36:891-894.

Krumperman, P.H. 1983. Multiple Antibiotic Resistance Indexing of *Escherichia coli* to Identify High-Risk Sources of Fecal Contamination of Foods. *Applied and Environmental Microbiology*. 46:165-170.

Lipman, L.J.A., De Nijs, A., Lam, T.J.G.M., and Gaastra, W. 1995. Identification of *Escherichia coli* strains from cows with clinical mastitis by serotyping and DNA polymorphism patterns with REP and ERIC primers. *Veterinary Microbiology* 43:13-19.

Mazengia, T., E.F. Fritsch, and J. Sambrook. 1982. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Moyer, N.P., et al. 1992. Application of ribotyping for differentiating Aeromonads isolated from clinical and environmental sources. *Applied and Environmental Microbiology*. 58(6) 1940-1944.

Parveen, S., K. M. Portier, K. Robinson, L. Edmiston, and M. L. Tamplin. 2000. Discriminate Analysis of Ribotype Profiles of *Escherichia coli* for Differentiating Human and Nonhuman Sources of Fecal Pollution. *Applied and Environmental Microbiology*. 65:3142-3147.

Samadpour, M. 1990. Molecular epidemiology of Shiga-like toxin producing *Escherichia coli* in food. Ph.D. dissertation. University of Washington Department of Food Science and Fisheries.

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

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal Molecular Biology* 98:503.

Stull, T.L., et al. 1988. A broad-spectrum probe for molecular epidemiology of bacteria: ribosomal RNA. *Journal of Infectious Diseases*. 157(2): 280-286

Wiggins, B.A., Andrews, R.W., Conway, R.A., Corr, C.L., Dobratz E.J., Dougherty, D.P., Eppard, J.R., Knupp, S.R., Limjoco, M.C., Mettenburg, J.M., Rinehardt, J.M., Sonsino, J., Torrijos, R.L., and Zimmerman M.E. 1999. Use of Antibiotic Resistance Analysis to Identify Nonpoint Sources of Fecal Pollution. *Applied and Environmental Microbiology*. **65**:3483-3486.