Appendix D

Overview of Microbial Source Tracking

Note: This discussion is an excerpt from: *Potential Application of Microbial Source Tracking Methods to the Dungeness Watershed and Bay, Clallam County, WA Technical Report, July 2003.*
By: Dana Woodruff, Battelle Marine Science Laboratory

Fecal Indicators

Traditionally, monitoring methods used for detection of pathogenic microorganisms (such as *Vibrio cholerae*, *Salmonella* spp., *Shigella* spp., hepatitis a virus, *Giardia* spp. *Cryptosporidium* spp.) in aquatic systems have been based upon cultivation and enumeration of fecal indicator bacteria (i.e. fecal coliforms, *Escherichia coli*, and fecal enterococci). Pathogenic microorganisms are generally few and difficult to identify and isolate. Therefore public health officials and scientists typically monitor nonpathogenic or indicator bacteria that are closely associated with pathogens. To function as an indicator of fecal contamination in surface water the organism should 1) be easily detected using simple inexpensive laboratory tests, (2) generally not present in unpolluted waters (3) appear in concentrations that can be correlated with the extent of contamination, and (4) have a die-off rate that is not faster than the die-off rate of pathogens (U.S. Environmental Protection Agency, 2001). While there is some controversy about the effectiveness of the traditional fecal indicators in terms of correlation with contamination and the die-off rate, there are other issues as well. Indicator bacteria are associated with fecal material from both humans and other warm- blooded animals, leading to many situations where one needs to determine the source host of the indicator bacteria.

Source Identification

Since the early 1900's there have been many attempts to discriminate between human and non-human sources of fecal coliform bacteria in surface waters, though they have been largely unsuccessful (Sinton et al., 1998). However, in recent years there has been increased interest in distinguishing between the many sources of fecal coliforms in surface waters. This has been partially driven by the Clean Water Act of 1972 and the inability of many states to meet the goal of "fishable, swimable" waters defined in the Clean Water Act. Section 303(d) of the Act addresses these waters by requiring states, territories, and authorized tribes to identify and list impaired waters every two years and to develop Total Maximum Daily Loads (TMDLs) for pollutants in these waters with oversight from the U.S. Environmental Protection Agency. TMDLs are primarily conducted for two reasons: 1) to establish the maximum pollutant load that a water body can

receive and still meet water quality standards, and 2) to allocate pollutant load reductions among point and non-point sources in order to provide a basis for establishing water quality controls and implementing one or more pollution control remedies (e.g. best management practices, BMPs) (U.S. Environmental Protection Agency, 2001; Simpson et al. 2002). Non-point fecal coliform contamination that is generated by agricultural practices, wildlife, and urban and rural development are particularly problematic sources of impairment to water quality because it is extremely difficult at times to determine the origin of these contaminants within and between these categories. Furthermore, to alleviate the overall fecal pollution problem, it is critical to determine whether the source of bacterial contamination is farm animals, pets, wildlife, or humans. Management and remediation of these sources would be more cost-effective if they could be correctly identified and remediation efforts allocated appropriately to those sources that can be controlled (Simpson, 2002).

Of the designated beneficial uses of water listed in section 303(c) of the Clean Water Act, protection from pathogenic microbe contamination is clearly most important for waters used for recreation, public water supplies, aquifer protection, and protection and propagation of fish, shellfish and wildlife (Simpson, 2002). While microbiological impairment of water is usually assessed by monitoring the presence of indicator bacteria as part of a TMDL implementation plan, tracking these bacteria to their source is being viewed with increasing interest in light of current TMDL requirements. For this reason, MST is being viewed as one approach for determining sources of fecal contamination affecting water bodies (Simpson, 2002). MST is based on the assumption that using an appropriate method and indicator bacteria, sources of microorganisms and associated target hosts can be found and characterized as to either: (1) human or animal origin, (2) general class of animal origin (e.g. livestock, wildlife, companion animals), or (3) species origin.

In February 2002, a Microbial Source Tracking Workshop sponsored by the EPA, the California State Water Resources Control Board, the Southern California Coastal Water Research Project (SCCWRP), and the National Water Research Institute was held in Irvine California. The workshop recognized that MST methods are potentially powerful tools that show promise but are still in the early stages of development. The workshop also brought together nationally recognized experts to define the state of knowledge, standardize field and laboratory methods for the most well developed approaches, and draft a protocol for a National Source Tracking Methods Comparison study to help inform local decision makers about the reliability of the different methods (Bernstein et al. 2002). Results from that study are expected to be available by the end of 2003. Another review of MST methods was recently completed by Meays et al. (2004).

MST approaches can be broadly categorized based on several sets of features. The first are those methods that are based directly on the genetic makeup of the indicator microorganism (genotypic methods), secondary characteristics of the indicator organism such as antibiotic resistance (phenotypic methods), and chemical tracers associated with human sources such as optical brighteners and caffeine (chemical methods) (Bernstein et. al, 2002). The second broad category are those methods that require the development of a background library of source isolates against which to compare a sample, and those that do not need a library for comparison (Table A-1). All of these methods are discussed in greater detail below.

Table A-1: A selection of microbial source tracking methods and their dependence on a source isolate library.

	Library Dependent	**Library Independent**
Genotypic (Molecular)	Ribotyping Pulsed Field Gel Electrophoresis (PFGE) Rep-PCR	Bacteroides/Prevotella markers Enterotoxin biomarkers (PCR)
Phenotypic (Biochemical)	Antibiotic Resistance Analysis (ARA) Carbon source profiling	F+ coliphage serotyping
Chemical Tracers		Caffeine detection Optical brighteners

Microbial Source Tracking Methods

Background

MST is clearly an emerging science, under continuous development, review, and refinement. Although these approaches are experimental and still fall under the umbrella of "research tools", they are considered by some to be the most promising approaches and "best available science" for determining sources (i.e. host organism) of fecal contamination in surface waters when the source is not apparent. Table A-2 summarizes the various MST methods.

The methods presented below are not all-inclusive, but represent techniques showing the most promise for success throughout the country, or have been used in the past with some success. Some of these methods could be applied to the Sinclair-Dyes Inlet watershed specifically. The methods have been grouped into three general categories: (1) chemical (anthropogenic) markers; (2) biochemical (phenotypic) methods; and (3) molecular (genotypic) methods. In a broad sense, they are also presented in order of increasing technical complexity, increasing cost, and greater resolution in terms of source tracking specificity.

Table A-2: Listing of microbial source tracking methods and degree of host source specificity and relative cost.

Method	Basis	Degree of Host Source Specificity Currently Achievable	Relative Cost *
Molecular (genotypic)			
Ribotyping	Ribosomal RNA are isolated to create distinctive bands of fingerprints for different sources	Numerous species	High
Pulsed Field Gel Electrophoresis	Distinguishes bacterial DNA using low-voltage, oscillating electrical current to separate bands	A few species	High
Rep-PCR	Identifies unique polymorphisms within the DNA of fecal bacteria	A few species	High
Bacteroides/Prevotella marker	PCR primers amplify specific genes related to individual species	A few species	Moderate/High
Biochemical (phenotypic)			
Antibiotic Resistance Analysis	Bacteria from different hosts have unique patterns of resistance to various antibiotics	Human, grouped animals, some species	Moderate
Carbon Source Profiling	Bacteria utilize carbon and other nutrients differentially for growth and energy	Human, grouped animals	Low/Moderate
F+ coliphages	Viruses that infect E. coli – various serotypes associated with animal groups	Human/grouped animals	Moderate
Chemical			
Optical Brighteners	Laundry detergent additives are usually associated with septic systems	Human only	Low
Caffeine Detection	Coffee/soft drink additive associated with septic/CSO waste	Human only	Moderate

* Relative cost is based on the cost per sampled analyzed, plus the number of samples that would need to be taken to conduct an MST study.

Chemical (Anthropogenic) Biomarkers

Chemical methods do not detect fecal bacteria, but rather are designed to detect chemical compounds that are commonly associated with human or anthropogenic activities. These biomarkers are usually found in sewage wastewaters; hence their presence in other water bodies may indicate a source of human-related fecal contamination.

<u>Optical Brighteners</u>

Optical brighteners or fluorescent whitening agents are dyes added to most laundry detergents, and labeled as "whiteners" or "brighteners". They are designed to adsorb to fabrics and "brighten" washing by fluorescing in the blue region of the visible spectrum, when exposed to ultraviolet radiation in sunlight (Sinton et al. 1998). They have a high affinity for cotton and while these dyes are invisible to the naked eye, they appear as a bright glow under black light (Waye, 2000). In a laboratory setting, when exposed to UV light (350 nm) they emit a blue light that can be measured with a scanning spectro-fluorometer (Sargeant, 1999; Fletcher, 1999). Several studies have used optical brighteners in helping to identify faulty septic systems, sewage exfiltration, and storm drain cross-connections (Sargent and Castonguay, 2003; Kerfoot and Skinner, 1981; Thrailkill et al., 1985) however, other methods such as dye testing and charcoal packets are more likely to detect failing on-site systems (Sargeant, 1999). In Virginia, Waye (2000) successfully used this technique to screen storm drain outfalls for sewage in the Four Mile Run watershed. In this study composite samples were taken during the summer months (300 outfalls) for a cost of less than $7,000. However, organic matter in natural water samples can also fluoresce and has been know to interfere with analyses (Sinton et al. 1998). A study in surface estuarine waters of a Virginia watershed measured uniform low levels of fluorescence that was due to natural background organic matter (Kator and Rhodes, 1996). In addition, the chemicals (primarily phosphorus based) are persistent in the environment and may not reflect sources of recent pollution.

Advantages and Disadvantages:

- Useful screening tool for failing on-site wastewater systems, sewage exfiltration, storm drain cross-connections.
- Optical brighteners persist in the environment, hence may not be reflective of recent pollution.
- Natural background fluorescence can interfere with analysis in some areas.
- Detection of human target only, not animal
- Cost – low, sample turnaround time is quick.

Caffeine Detection

This method has been proposed as an indicator of human sources of pollution and is currently under development. Caffeine passes through the human digestive system and its presence in the environment can be indicative of fecal contamination. Several studies have used this as in indicator in metropolitan areas. King County Metro used caffeine to investigate wastewater problems in combined sewer overflows (CSOs) in Duwamish River and Elliot Bay (Sargeant, 1999). It was detected in close proximity to CSOs and found in some outfall samples. The levels need to be relatively high to be detectable. A dilution of more than 200:1 makes detection difficult (Sargeant, 1999). Also, some other plants have significant levels of caffeine (e.g. watermelon), and caffeine is easily degraded by soil microbes (Hagedorn, 2003b). Both of these issues can easily confound sample results.

Advantages and Disadvantages:

- Useful for CSO problem detection.
- Need high level of caffeine to be detectable, best used in metropolitan area.
- Cost is moderate/high (~ $100/sample, Hagedorn, 2003b).
- Detection of human target only, not animal
- Natural background presence can interfere with results

Biochemical (Phenotypic) methods

Biochemical or phenotypic methods identify bacteria based on a physical expression of some unique characteristic of the bacterium. These can include metabolic measurements, immunological measurements, or resistance to antibiotics.

Antibiotic Resistance Analysis

Antibiotic resistance analysis (ARA) uses antibiotic resistance patterns of fecal *E. coli*, streptococci and enterococci to identify fecal sources. It is based on the premise that human fecal bacteria will have a greater resistance to antibiotics than those from animal sources. The fecal bacteria of animals should have significantly less and different resistance to the wide variety and concentrations of antibiotics used commercially (Simpson et. al., 2002).

Test bacteria are isolated from samples of known fecal sources and grown on media containing various concentrations of a variety of commercial antibiotics (~10 to 20). Each isolate is scored as sensitive or resistant to each concentration of each drug. The resulting resistance patterns of each isolate are

then analyzed using a statistical approach called discriminant analysis, which is a major component of ARA. Discriminant analysis classifies bacteria based on shared patterns of antibiotic resistance, where the results are pooled to form a "known library" of patterns from different fecal sources. The average rate of correct classification is the average rates that known isolates are correctly classified, and is used to measure the reliability of the isolate library. Once the known isolates are classified, the resistance patterns from the unknowns are compared to determine the sources (Harwood, 2002).

ARA has been used in a number of studies with a variety of sources for comparison, including humans, sewage, cattle, chickens, turkeys, dogs, pigs, raccoons, wild bird, and wild animals (Harwood et al., 2000; Wiggins 1996; Wiggins et al. 1999; Graves et al. 2002; Parveen et al. 1997). Typically, ARA methods result in an average rate of correct classification ranging between 62 and 94 percent when individual species were compared. Values increase by about 10 percent when comparing human and non-human. However, when unknown source isolates are compared with the libraries, the correct classification decreases to approximately 72 percent (Simpson et. al., 2002). A recent study by Wiggins et al. (2003) compared the size and composition of various libraries and found that small libraries tended to have higher average rates of correct classification for known sources, but were much less able to correctly classify non-library isolates.

Advantages and disadvantages:

- Better suited for small simple watersheds, not larger complex systems.
- Can distinguish human and animals, animal groupings.
- Library dependent, need large library.
- Dependent on antibiotic usage, antibiotic resistance can be transferred.
- High throughput of samples (i.e. many samples can be analyzed in a short time).
- Relatively easy to perform.
- Moderate to high cost depending on whether a source isolate library already exists.

Carbon Source Profiling

Carbon source profiling is based on how carbon and nitrogen sources are utilized for energy and growth by different bacteria. The method utilizes a commercially available BIOLOG ™ system allowing a user to rapidly perform, score, and tabulate 96 carbon source utilization tests per isolate (Simpson, 2002). It has been widely used in the medical profession for identification of isolates.

This technique is quite new in terms of its application to MST studies. A study conducted recently in Virginia tested its ability to correctly classify human, livestock, domestic pets, and wildlife (Hagedorn et al. 2003a). 365 *Enterococcus* isolates were collected from known sources in five geographic regions. The average rate of correct classification for human versus non-human was 92 percent when isolates from all regions were combined into one library. A four-way classification between each major group of animal resulted in a correct classification rate of 80 percent. While carbon source profiling has been discussed as a promising new tool, no other applied studies were found in the literature on this method.

Advantages and disadvantages:

- Very new, method needs further development for MST
- May separate human from large animal groupings
- Stability of markers needs to be evaluated
- Library dependent
- Easy to perform and interpret.
- Inexpensive.

Molecular (Genotypic) methods

Molecular methods are sometimes referred to as "DNA fingerprinting" and are based on small differences in the genetic makeup of different strains of fecal bacteria. The key to this method is finding small differences against a large background of similarity. It is thought that the small distinctions between fecal bacteria from different animals occur because the intestinal environments (selective pressures) are not the same, and those differences can be related to source (Hagedorn, 2003b).

Molecular methods can also be categorized as "library" dependent (source database required) or "library" independent (source database not required). A library refers to a collection of microorganisms from different potential human and animal specie sources, particularly from the watershed under study (Simpson et. al., 2002). Most of the MST methods developed to date are dependent on the development of comprehensive libraries. The approach requires the culturing of unknown bacterial source samples to obtain pure isolates that will be compared to those of the library. Although the libraries often include several hundred isolates per potential source impacting a particular watershed (e.g. human, livestock, wildlife, pets), the minimum number of isolates needed to perform statistically sound studies has not been determined. In general, most methods have not been thoroughly tested, although some have been successfully applied (Simpson et al., 2002).

Several new molecular methods have been developed recently that do not require the use of a library. The first method described here does not require a source database or library: *Bacteroides-Prevotella* markers; whereas the other three methods described do require a source isolate library: ribotyping, PFGE, and rep-PCR.

Bacteroides-Prevotella marker

Bacteroides is a group of anaerobic bacteria that make up one-third of human fecal flora, far outnumbering coliforms (Holdeman et al. 1976). *Bacteroides,* and a close relative *Prevotella,* are found exclusively in feces and other cavities within humans and animals (Paster et al. 1994). When found in water they are invariably diagnostic of pollution. Since fecal anaerobes do not survive for long periods in the environment, their presence represents recent fecal contamination events. This group of bacteria has traditionally not been used as an indicator because of the difficulty of growing anaerobic bacteria. However, since they make up such a high proportion of the fecal flora, they are easier to detect by PCR (described below) from water than the less abundant coliforms (Field, 2002).

Polymerase chain reaction (PCR) is a molecular method that allows one to amplify genetic markers by making multiple copies of a target gene sequence in a test tube. 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. Length heterogeneity - polymerase chain reaction (LH-PCR) (Suzuki et al. 1998) and terminal restriction fragment length polymorphism (T-RFLP) (Liu et al. 1997) are two techniques that analyze differences in lengths of gene fragments due to insertions and deletions, and estimate the relative abundance of each fragment (Bernhard and Field, 2000). Recently *Bacteroides-Prevotella* sequences have been used to design PCR primers that specifically amplify 16S ribosomal RNA genes from Bacteroides-Prevotella bacteria (Field et al., 2003; Bernhard and Field, 2000a; Bernhard and Field, 2000b).

Several recent studies have applied this methodology to Tillamook Bay in Oregon and several of its major tributaries (Bernhard et al., 2003; Bernhard and Field, 2000a; Bernhard and Field, 2000b). This area has been impacted by fecal pollution, causing economic losses to the shellfish industry and leading to closures of the bay for recreation and fisheries. Potential contributors include farm animal waste, sewage treatment plants, and septic systems. Genetic markers for cattle and human were developed and successfully used in these studies. Markers for pigs, elks, dogs/cats, gulls, chickens, ducks/geese, gulls and seals are currently under development.

Advantages and disadvantages:
- New method, currently under development.
- No library needed
- More representative indicator of fecal pollution than *E. coli* in some situations, especially marine.
- Markers developed for human and cow, others under development
- Technically demanding
- Cost – high initially, reduced as markers are developed.

Pulsed Field Gel Electrophoresis

Pulsed field gel electrophoresis (PFGE) is a methodology that has been commercially developed by BioRad™. Pure cultures of bacterial cells are placed in agarose plugs where DNA is digested using a series of restriction enzymes. The digested plugs are then imbedded into specialized gels and an electric current passed through the gel with alternating currents in different directions for an extended period of time (several days) (Simpson et. al., 2002). The DNA restriction fragments (bands) are separated and compared to each other and to a library of known sources. The method has been used extensively in clinical microbiology, however it is time consuming and the number of isolates that can be processed at one time is limiting (Simpson, 2002). It has been used less frequently for environmental isolate identification; however Simmons et al., (2000) used PFGE on isolates of *E.coli* DNA in the Four Mile Run watershed in Virginia. Fifty one percent of the isolates collected were matched to an isolate from one of several library databases. The remaining isolates were grouped as unknown (30 percent could not be matched to a library database), inconclusive (based on statistical evidence), non- *E.coli* fecal coliforms, or unusable. The remaining isolates were matched to waterfowl, human, raccoon, deer, canine and others.

Advantages and disadvantages:
- Gold standard for genotyping in clinical work.
- Few watershed applications to date.
- High discrimination and reproducible, however may be too sensitive for MST.
- Dependent on source library.
- Very time consuming, limited simultaneous processing.
- Expensive.

Rep-PCR

Rep-PCR is a molecular technique where PCR amplification of the DNA between adjacent repetitive extragenic elements is used to obtain strain-specific DNA

fingerprints that can easily be analyzed with pattern recognition computer software (Dombek et al. 2000). This technique has been used routinely for clinical studies of human and animal pathogens (Simpson et al, 2000), and has recently been applied to an MST study in Minnesota (Dombek et al. 2000). A source library of approximately 2,500 *E.coli* isolates from thirteen sources was developed, including: cows, pigs, sheep, goats, turkeys, chickens, ducks, geese, deer, horses, dogs, cats, and humans (Johnson et. al., 2002).   The source library was then compared to fecal bacteria isolated from four impacted Minnesota waterways. Overall, about 63 percent of environmental *E. coli* isolates were found to match those in the DNA library based on a 90 percent statistical similarity threshold. It was concluded that the library database of known sources needs to be expanded.

Advantages and disadvantages:

- Highly reproducible with good discrimination.
- Dependent on source library.
- Few watershed applications to date.
- Potential for quick turnaround time
- Expensive

Ribotyping

Ribotyping refers to a method of fingerprint pattern generation of DNA fragments based on the genes that code for ribosomal ribonucleic acids (rRNA) of *E. coli* or other bacteria*.* To trace the indicator bacterium, *E. coli*, from water to its specific source, the strain of *E. coli* must be uniquely identified. Populations of *E. coli*, like other bacteria, are essentially composed of a mixture of strains of clonal descent. A clone is a population of cells that is descended from a single cell. Due to the relatively low rates of recombination, the clones remain more or less independent (Selander et al., 1987). These clones, or strains, 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 (Samadpour, 2002). The subtle genetic differences between bacterial strains are the basis for ribotyping.

The ribotyping method involves multiple steps and the development of a source library. Water samples and source samples are collected and cultured to isolate pure cultures of *E. coli*.   DNA is extracted from these cultures and cut into fragments using restriction enzymes. Several different restriction enzymes can be used to increase the discriminatory power of the analysis. The fragments are then separated using agarose gel electrophoresis and transferred to nylon membranes. A technique called Southern blot hybridization is performed using radio-labelled rDNA probes that results in a pattern of bands corresponding to a

ribotype. These band patterns can then be used to uniquely identify the bacterial strain. Bacterial strains are identified by comparing the unknown sample isolate pattern against the knowns in the source library (Simpson et al., 2002; Samadpour, 2002).

Several investigators have applied ribotyping methodology to MST studies. Parveen et al. (1999) used ribotyping to distinguish fecal *E. coli* of human origin from pooled fecal *E. coli* isolates of non-human origin in estuarine waters of the Apalachicola National Estuarine Research Reserve. The average rate of correct classification (ARCC) was 82 percent. Carson et al. (2001) used ribotyping to distinguish between human and seven other nonhuman hosts. The average rate of correct classification was 97 percent between human and non-human, however the ARCC dropped to 74 percent when distinguishing between the eight species. A relatively small number of known source isolates (287) was used in this study, which largely accounts for the low numbers of correctly identified isolates.

Two of the major questions regarding ribotyping and other methods that rely on a source library are how large a database needs to be, and how transferable a database is from one geographic region to another. The latter question was partially addressed in a study (Hartel et al. 2002) that compared geographic variability of 586 *E. coli* isolates from cattle, horses, swine and chickens located in Idaho and Georgia. Results showed that some variability existed for *E. coli* isolates between the regions for the four species. While some *E. coli* subspecies were found in more than one location (i.e., cosmopolitan), other subspecies were found in only one location (i.e., endemic). This would suggest that while there may be some applicability of a host source database from one region to another, source isolates collected from the watershed/estuary of interest will greatly add to the success of a study. It was suggested by Mansour Samadpour of the University of Washington, in a recent article in Science (Malakoff, 2002) that a library of approximately 500,000 *E. coli* strains collected from a range of vertebrates could cover the country if the data were updated regularly.

A number of ribotyping studies have been conducted in the Pacific Northwest using techniques developed by M. Samadpour. He currently has a source library that contains over 90,000 strains of *E. coli* (M. Samadpour, pers. comm.), and has been involved in over 80 source tracking studies (Malakoff, 2002). Although the results of these studies vary somewhat, the source library has increased through the years, and in general the frequency of positive identification of unknowns has increased. A 1991 study conducted in the Pipers Creek watershed in northwest Seattle showed a 43 percent match of unknowns to known source samples, predominantly cat, dog and duck (Herrera Environmental Consultants, 1993). A study conducted between 1993 and 1995 at Little Soos Creek in southeast King County identified approximately three-fourths of the fecal

coliform contamination, with the primary sources determined to be cows, dogs, and horses (Samadpour and Chechowitz, 1995).

In 2000-2001 a study was conducted in the Henderson Inlet area in Puget Sound in response to the worsening water quality of the marine shellfish growing areas (Samadpour et al., 2002). Four stations were sampled including one marine station and three creeks that drain into Henderson Inlet. A total of 943 isolates were typed from sediment, shellfish, fresh and marine waters. At the time of the study, Samadpour's source isolate library had approximately 65,000 fingerprints. An additional 100 source isolates were added to the library from the Henderson Inlet area through this study. A relatively high percentage of unknown samples were matched in this study - 86 percent. The highest percent of matches were in a creek area (92 percent) and the lowest matches were in the marine water (82 percent). Overall, a total of 27 source types were identified including: avian, beaver, bovine, canine, cat, deer, dog, duck, duck-goose, feline, goose, horse, human, marine mammal, multiple species, muskrat, opossum, otter, porcupine, poultry, rabbit, raccoon, rodent, seagull, sea lion, seal, and unknown. The high number of source matches in this study was attributed to the size of the library as well as the rigor of the study design. The study design included a comparison of wet weather vs. dry weather conditions and ebb vs. flood tide conditions at the marine site.

Advantages and disadvantages:

- Highly reproducible with good discrimination
- Dependent on source library
- One of most widely used methods for MST
- A number of successful applied studies conducted to date
- Extensive source isolate available
- Expensive (~$75 per isolate)

Summary of MST methods

The methods presented above clearly show a range of applications, costs, and usefulness. The chemical (anthropogenic) biomarkers include optical brighteners and caffeine detection. In both cases they can potentially provide evidence of human sources, but not other animals. These methods are by far the least expensive. Optical brighteners should be considered as a screening tool only for failing septic systems or storm drain cross-connections. Optical brighteners are based on the premise that laundry facilities are connected to on-site septic systems, which is not always the case. In addition, natural background levels can give false positive readings depending on the geographic region where they are applied. Caffeine detection is best applied to highly populated areas since

detection is difficult when not close to the source. Their primary usefulness appears to be addressing problems associated with CSOs in metropolitan areas.

Phenotypic or biochemical methods are more expensive than chemical methods, however they have the potential of identifying some non-human sources. Most of these methods require the development of a source isolate library for comparison of knowns with unknown sources. Since phenotypic profiles are considered by some to be less stable than genotypic profiles, these methods are best applied to smaller less complex watersheds. Carbon source profiling shows promise as a useful method, however it is still in the developmental stages and has been applied to very few microbial source tracking studies to date. ARA, on the other hand, has been applied to a number of microbial source tracking studies with varying success. It appears to be best applied to smaller, less complex watersheds.

Molecular or genotypic methods rely on identifying small differences in the genetic makeup of different strains of fecal bacteria. While most of these methods have not been thoroughly tested, some have been successfully applied to source tracking studies. The *Bacteroides/Prevotella* biomarker does not require a source isolate library for comparison purposes, however a number of the molecular methods do require a library. The methods are generally more expensive than phenotypic methods, however can provide a higher level of source identification. The *Bacteroides/Prevotella* biomarker shows great promise for the future because it is library independent and may be more representative as a fecal indicator than *E. coli*. The current drawback is that relatively few markers have been developed and tested (primarily human and cow), however additional markers are currently being developed. Several other methods, PFGE and Rep-PCR, show promise but need further development and application to watershed studies. In recent years, the ribotyping method has been applied by several investigators to watershed studies. The success of this library dependent method depends in great part on the size of the library database. If a library database is sufficiently large, it can provide a fairly high level of sophistication in terms of species identification. The drawback of this method is the time consuming analysis and relatively high cost.

The MST methods discussed above can loosely be grouped into three categories:

- Chemical - human source contributions
- Phenotypic – human/animal or broad groups (e.g. wildlife, pets, livestock)
- Genotypic – human/grouped animals with some species specificity

Chemical markers that identify human contributions, while inexpensive, do not provide the level of source identification needed in the Dungeness area.

Phenotypic methods provide some added level of specificity (i.e. major animal groups). The most widely applied phenotypic method is ARA. This method is less expensive than genotypic methods and has been used to identify major animal groups (e.g. wildlife, pets, livestock) with varied success. While identification of major animal groups would be useful, there are several drawbacks of this method. First, the success of the ARA method is dependent partially on the size is the watershed (i.e. it is most successfully used in watersheds that are smaller than the Dungeness). Second, the method has not been applied to a watershed as complex as the Dungeness in terms of the possible source contributions (e.g. marine mammals, exotic animals). Finally, because this method has not been used in this geographic region before, a source library would need to be developed which would involve collection of thousands of source isolate samples.

Genotypic methods can provide a species level of specificity in some cases. There are several methods that show promise for the future, including the *Bacteroides/Prevotella* marker, that do not require a source library. This method, however, still needs further development since the number of species specific markers available at this time is very limited. The ribotyping method is, by far, the most widely used method for MST studies.

Although ribotyping has had varied success in the past, recent advances and increased library size have improved its success rate (Samadpour, 2002). The study in Henderson inlet is a case in point, which had a source identification rate of 86 percent, and included a marine and upland component. The library used for this study contained over 65,000 known E. coli isolates, with 100 of those contributed from the specific study area. (Samadpour, 2002). The source isolate library currently contains over 90,000 isolates, with many of those from the Puget Sound region. Based on the success of these recent studies and a source isolate library that has already been developed, we would recommend the use of this particular ribotyping method for the Dungeness Bay and watershed. Although additional source isolates would need to be collected to increase the level of percent matches, this method provides the level of species specificity that can identify controllable sources in the Dungeness area.

Ribotyping methods can provide information on the predominant source types and the frequency of occurrence at a given site, however they cannot easily determine the fecal loading of specific sources which is one element of TMDL development and implementation. While one source may occur more frequently, it may individually contribute less fecal matter (Samadpour, 2002). In general, a sampling scheme designed to collect the number of isolates necessary to determine fecal loading would be prohibitively expensive. However, a study design to determine predominant source types affecting an area would help develop realistic implementation plans for remedial efforts to restore water quality. MST methods are evolving and advancing at a rapid pace. While

ribotyping may be the most developed approach at this time based on the state of the science, this could change as methods evolve or additional information is obtained.

To successfully implement an MST study, the following points should be kept in mind:

- <u>The method should fit the question</u> – In other words, the level of source specificity needed to answer the question should determine the method chosen. If the question is simply "Are there human or animal sources here?" then phenotypic methods may be adequate, however if greater source specificity is required, then molecular methods may be necessary.

- <u>Sample design is critical</u> – A thoughtfully, well designed study can make a critical difference between success and failure. Input for the sample design should include regional health professionals, watershed specialists, and scientists who would be conducting the MST sample analysis. In general, working from upstream sources to downstream receiving waters is the most efficient approach for most situations.

- <u>Understand the method limitations</u> – MST is not a "magic bullet". It is one tool, with its own limitations, that can provide a piece of the puzzle in understanding sources of non-point fecal contamination.

- <u>Keep an open mind about potential sources</u> – Many presumptions are made regarding assumed sources. MST studies have yielded many "surprises" about unexpected sources, hence the need to carefully select sites to sample.

- <u>Review the available methods</u> – If an MST study is undertaken, the current methods should be reviewed for updates prior to selecting a method(s). This is a rapidly changing field with advances occurring regularly.

- <u>Implement remedial actions based on the results of the MST study.</u> Over the long term, MST studies will only be as successful as the implementation of remedial actions and BMPs based on the findings of the study, especially regarding sources that can be controlled.

In addition, be aware of some of the drawbacks and limitations of MST with regard to TMDLs:

- Ranges of literature value for animal loading rates vary widely, with no attempt to calculate uncertainty ranges.

- Expectations by some TMDL developers and reviewers that MST results will generate realistic animal populations through back-calculations with literature-value loading rates are misplaced, as the uncertainty inherent in these rates and the fate and transport of bacteria populations is poorly understood.

- MST analysis is often inadequate as a result of too-few samples with unknown uncertainty ranges. There is frequently a sole reliance on an overly-limited number of MST "snapshot" grab samples to apply to animal loadings in continuous simulation TMDL modeling.

- The vast majority of MST analyses have been applied only to dry weather baseflow conditions, but critical TMDL conditions are often from very high flow events, which may have a different pattern of bacteria sources.

- Bacteria are nearly always simulated in complex TMDL watershed models as a chemical governed by a first-order die-off, rather than as competing colonial microbes that stick to surfaces and create bio-films. While these models are "complex" in terms of user-friendliness and data inputs, they are often built on simplistic and limiting assumptions about how bacteria behave in urban environments.

- There is very limited research on the factors that influence bacteria transport and die-off from pervious and impervious surfaces. In addition, there are no literature values or guidance on species-specific animal loadings on impervious surfaces.

- In short, more research and better bacteria modeling assumptions and routines are needed to adequately simulate bacteria population dynamics.

Microbial Source Tracking References

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. *App. Environ. Microb.* 66(4):1587-1594.

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. *Appl. Environ. Microbiol.* 66:4571-4574.

Bernstein, B.B., J.F. Griffith, S.B. Weisberg. 2002. Microbial Source Tracking Workshop: Summary of Proceedings. U.S. EPA workshop. Feb. 5-7, 2002, Irvine, California.

Brion, G.M., J.S. Meschke, M.D. Sobsey. 2002. F-specific RNA coliphages: occurrence, types and survival in natural waters. *Water Res*. 36: 2419-2425.

Carson, C.A., B.L. Shear, M.R. Ellersieck, and A. Asfaw. 2001. Identification of fecal *Escherichia coli* in humans and animals by ribotyping. *Appl. Environ. Microbiol.* 67(4): 1503-1507.

Clean Water Workgroup. 2002. Clean Water Strategy for addressing bacterial pollution in Dungeness Bay and Watershed. Clallam County, WA.

Dombek, P.E., L.K. Johnson, S.T. Zimmerley, M. J. Sadowsky. 2000. Use of repetitive DNA sequences and the PCR to differentiate *Escherichia coli* isolates from human and animal sources. *Appl. Environ. Microb*. 66(6):2572-2577.

Dureth, S., H. Reimer, and K. Pecher. 1986. Tracing faecal pollution by coprostanol and intestinal bacteria in an ice covered Finnish lake loaded with both industrial and domestic sewage. *Water, Air, Soil Pollution.* 28:131-149.

Field. K.G. 2002. Fecal Source Tracking with *Bacteroides.* Abstract from Microbial Source Tracking workshop. Feb. 5-7, 2002 Irvine, CA pp. 17-21.

Field, K.G., A.E. Bernhard, and T.J. Brodeur. 2003. Molecular approaches to microbiological monitoring: Fecal source detection. *Environ. Monit. And Assess*. 81: 313-326.

Fletcher, M., P. Verity, M. Frischer, K. Maruya, and G. Scott. 1999. Microbial indicators and phytoplankton and bacterial communities as evidence of contamination caused by changing land use patterns. Baruch Institute, University of South Carolina. 98 pp.

Graves, A., C. Hagedorn, A. Teetor, M. Mahal, A. Booth, R.B. Jr. Reneau, 2002. Antibiotic resistance profiles to determine sources of fecal contamination in a rural Virginia watershed. *J. Environ. Qual.* 31:1300-1308.

Hagedorn, C., S.L. Robinson, J.R. Filtz, S.M. Grubbs, T.A. Angier, R.B. Reneau, Jr. 1999. Determining sources of fecal pollution in a rural Virginia watershed with antibiotic resistance patterns in fecal streptococci. *Appl. Envir. Micro*. 65(12):5522-5531.

Hagedorn, C. 2002. Carbon source profiles, pulsed-field gel electrophoresis patterns, and antibiotic resistance analysis. Abstract from Microbial Source Tracking workshop. Feb. 5-7, 2002, Irvine, CA pp. 55-58.

Hagedorn, C. J.B. Crozier, K.A. Mentz, A.M. Booth, A.K. Graves, N.J. Nelson, R.B. Reneau, Jr. 2003a. Carbon source utilization profiles as a method to identify sources of faecal pollution in water. *J. Appl. Microb*. 94(5): 792-799.

Hagedorn, C. 2003b. BST Methodologies. Internet address: http://soils1.cses.vt.edu/ch/biol_4684/bst/BSTmeth.html. Accessed June, 2003.

Hartel, P.G., J. D. Summer, J.L. Hill, J.V. Collins, J.A. Entry, and W.I. Segars. 2002. Geographic variability of *Escherichia coli* ribotypes from animals in Idaho and Georgia. *J. Environ. Qual*. 31:1273-1278.

Harwood, V.J., J. Whitlock, V. Withington. 2000. Classification of antibiotic resistance patterns of indicator bacteria by discriminant analysis: Use in predicting the source of fecal contamination in subtropical waters. *Appl. Envir. Micro*. 66(9): 3698-3704.

Herrera Environmental Consultants. 1993. Pipers Creek bacteriological source tracking investigation. Report prepared for Seattle Engineering Department, Drainage and Wastewater Utility, Seattle, WA. 39 pp.

Jamestown S'Klallam Tribe. 2003. Restoring the Dungeness: An Overview of the Dungeness River Restoration Strategy. Prepared for the National Oceanic and Atmospheric Administration through the Pacific Coast Salmon Recovery Program. 118 pp.

Johnson, R.K., M.B. Brown, P.E. Dombek, M.J. Sadowsky. Source tracking fecal bacteria in the environments using rep-PCR DNA fingerprinting: Prospects and Problems. Abstract in Microbial Source Tracking workshop. Feb. 5-7, 2002, Irvine CA. pp. 37-41.

Kator, H. and M. Rhodes. 1996. Identification of pollutant sources contributing to degraded sanitary water quality in Taskinas Creek National Estuarine Research Reserve, Virginia. Final Report, Taskinas National Estuarine Research Reserve, OCRM #NA470R0199, College of William and Mary, Virginia Institute of Marine Science. 47 pp.

Kelsey, H., G. Scott, D. Porter, B. Thompson, and L. Webster. 2003. Using multiple antibiotic resistance and land-use characteristics to determine sources of fecal coliform bacterial pollution. *Environmental Monitoring and Assessment* 81(1-3):337-348.

Kerfoot, W.B. and S. M. Skinner. 1981. Septic tank leachate surveys for lakeside sewer needs evaluation. *J. Water Pollut. Control Fed*. 53:1717-1725.

Krahn, M. M., C.A. Wigren, L.K. Moore, and D.W. Brown. 1989. High performance liquid chromatographic methods for isolating coprostanol from sediment extracts. *J. Chrom*. 481:263-273.

Leeming, R. Ball, A. Ashbolt, N.J., G. Jones, and P.D. Nichols. 1994. Distinguishing between human and animal sources of faecal pollution. *Chemistry in Australia, water (supplement*): p. 434-435.

Lui, W.T., T.L. Marsh, H. Cheng, L.J. Forney. 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Appl. Environ. Microb*. 63:4,516-4,522.

Malakoff, D. 2002. Microbiologists on the trail of polluting bacteria. *Science*. 295:2352-2353.

Meays, C., K. Broersma, R. Nordin, and A. Mazumder. 2004. Source tracking fecal bacteria in water: a critical review of current methods. *Journal of Environmental Management* 73: 71-79.

Melvin, D.J. 2000. Sanitary Survey of Dungeness Bay, March 20, 2000. Prepared by the Washington State Department of Health Office of Shellfish Programs. 29 pp.

Montgomery Water Group. 1999. *The Dungeness River Water Users Association Comprehensive Water Conservation Plan, Volume 1*. Montgomery Water Group, Inc, Pacific Ground Water Group, The Environmental Company, Kirkland, WA.

Nichols, P.D., Leeming, R. Rayner, M.S., Latham, V. Ashbolt, N.J., Turner, C. 1993. Comparison of the abundance of the fecal sterol coprostanol and fecal

bacterial groups in inner-shelf waters and sediments near Sydney, Australia. *J. Chromatography*. 643:189-195.

Parveen, S. R.L. Murphree, L. Edmiston, C.W. Kaspar, K.M. Portier, and M.L. Tamplin. 1997. Association of multiple-antibiotic-resistance profiles with point and nonpoint sources of *Escherichia coli* in Apalachicola Bay. *Appl. Environ. Microbiol*. 63:2607-2612.

Parveen, S., K.M. Porter, K. Robinson, L. Edmiston, M.L. Tamplin. 1999. Discriminant analysis of ribotype profiles of *Escherichia coli* for differentiating human and nonhuman sources of fecal pollution. *Appl. Environ. Microbiol*. 65(7):3142-3147.

Rensel, J.E. and T. Smayda. 2001. *Dungeness Bay Circulation and Fecal Coliform Assessment: (Phase 1 Study).* Report prepared for Jamestown S'Klallam Tribe. August 3, 2001 version.

Rensel, J.E. 2003. *Dungeness Bay Bathymetry, Circulation, and Fecal Coliform Studies: Phase 2*. Report prepared for Jamestown S'Klallam Tribe and U.S. EPA, Region 10. 94 pp.

Richter, K.E, 2004. CH3D Hydrodynamic Model Validation Draft Report. Space and Naval Warfare Systems Center, San Diego, CA. September 12, 2004

Samadpour M. and N. Chechowitz. 1995. Little Soos Creek microbial source tracking. Report to Surface Water Management Division. King County Dep. of Public Works, Seattle, WA. 35 pp.

Samadpour, M. 2002. Microbial Source Tracking: Principles and Practice. Abstract from Microbial Source Tracking Workshop. Feb. 5-7, 2002. Irvine, CA. pp.5-9.

Samadpour, M., L. Hofstad, S. Davis. 2002. Bacteriological contamination source identification, Henderson Inlet 1999-2001. Report prepared for Thurston County Public Health Environmental Health Division. 35 pp.

Sargeant, D. 1999. *Fecal contamination Source Identification Methods in Surface Water*. Washington Dept. of Ecology Report #99-345. 17 pp.

Sargeant, D. 2002. *Dungeness River and Matriotti Creek fecal coliform bacteria total maximum daily load study*. Environmental Assessment Program. WA Dept. of Ecology Pub. No. 02-03-014. Olympia, WA 46 pp.

Sargent, D., and W. Castonguay. 2003. Water Quality Sampling: An Optical Brightener Handbook. Internet address: http://www.naturecompass.org/8tb/sampling/index.html. Accessed June 2003.

Selander,R.K., D.A. Caugant, T.S. Whittam. 1987. Genetic 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.

Simmons, G. D. Waye, S. Herbein, S. Myers, E. Walker. 2000. Presented at Virginia Water Research Symposium 2000, Blacksburg, VA, Younos, T., Poff, J., Eds.; VWRRC Special Report SR-19-2000.

Simpson, J.M., J.W. Santo Domingo, and D.J. Reasoner. 2002. Microbial Source Tracking: State of the Science. *Environ. Sci. and Tech.* vol. 36 (24):5279-5288.

Sinton, L.W., R.K. Finlay, D.J. Hannah. 1998. Distinguishing human from animal faecal contamination in water: a review. *New Zeal. J. of Mar. and Fresh. Res*. Vol. 32: 323-348.

Sobsey, M. 2002. Coliphage tracking to identify sources of fecal contamination. Abstract from Microbial Source Tracking workshop. Feb. 5-7, 2002. Irvine, CA. pp. 59-63.

Suzuki, M. M.S. Rappe, 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. *Appl. Environ. Microb*. 64:4,522-4,529.

Thrailkill, J., R. Wiseman and B. Scanlon. 1985. Investigation of pollution in a Karst aquifer utilizing optical brighteners. Research Report #158. Water Resources Research Institute, Univ. of Kentucky, Lexington, Kentucky.

U.S. Environmental Protection Agency. 2001. Protocol for Developing Pathogen TMDLs. EPA 841-R-00-002. Office of Water (4503F), United States Environmental Protection Agency, Washington, DC. 132 pp.

Venkatesan, M.I., E. Ruth, I.R. Kaplan. 1986. Coprostanols in Antarctica marine sediments: a biomarker for marine mammals and not human pollution. *Mar. Poll. Bull.* 17: 54-557.

Wade, T.L., G.F. Oertal and R.C. Brown. 1983. Particulate hydrocarbon and coprostanol concentrations in shelf waters adjacent to the Chesapeake Bay. *Can. J. Fish. Aquat. Sci*. 40: 34-40.

Walker, R.W., C.K. Wun, W. Litsky. 1982. Coprostanol as an indicator of fecal pollution. *CRC Crit. Rev. Environ. Control* 10: 91-112.

Waye, D. 2000. A New Tool for Tracing Human Sewage in Waterbodies: Optical Brightener Monitoring. Abstract from Virginia Water Resources Research Symposium, Nov. 2000. Internet address: http://www.novaregion.org/4milerun/pr_illum_leaks.htm, accessed May 2003.

Webster, L.F., B.C. Thompson, M.H. Fulton, D.E. Chesnut, R.F. Van Dolah, A.K. Leight, and G.I. Scott. 2004. Identification of sources of *Escherichia coli* in South Carolina estuaries using antibiotic resistance analysis. *Journal of Experimental Marine Biology and Ecology* 298(2): 179-195.

Wiggins, B.A. 1996. Discriminant analysis of antibiotic resistance patterns in fecal streptococci, a method to differentiate human and animal sources of fecal pollution in natural waters. *Appl. Envir. Micro*. 62(22): 3997-4002.

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. Rinehard, J. Sonsino, R. Torrijos, M.E. Zimmerman. 1999. Use of antibiotic resistance analysis to identify nonpoint sources of fecal pollution. *Appl. Environ. Microbiol*. 65:3484-3486.

Wiggins, B.A. 2002. Microbial source tracking: Using antibiotic resistance analysis. Abstract from Microbial Source Tracking workshop. Feb. 5-7, 2002. Irvine, CA. pp. 43-46.

Wiggins, B.A. et al. 2003. Use of antibiotic resistance analysis for representativeness testing of multi-watershed libraries. *App. Envir. Micro*. 69(6):3399-3405.