SECTION 3
MICROBIAL SOURCE TRACKING OBJECTIVES AND METHODS

3.1 Introduction/Background

This section provides a basic background to the science of microbial source tracking.

3.1.1 Underlying Assumptions

The MST method is based on two principles: first, that the genetic structure of the bacterial population is clonal. This is a well-established element of microbial genetics. Bacteria reproduce by binary fission, or dividing in half. The two daughter cells generated as a result of this cell division are virtually identical in all aspects. All descendents of a common ancestral cell are genetically related to each other. Over time, members of a given clone may accumulate genetic changes which will cause them to diverge from the main lineage and 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 MST methods 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 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 the host's 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 reliable, characteristic signature of their source. A transient strain is a bacterial strain that is introduced into a new environment or host but which 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. Dr. Mansour Samadpour of Molecular Epidemiology, Inc. (MEI) accumulated considerable evidence to support this assertion for *E. coli*. Using this subtyping method (ribosomal ribonucleic acid [rRNA] typing using two restriction enzyme reactions), data show that more than 96 percent of *E. coli* strains are seen in only one host species, or a group of related species (Mazengia 1998). Thus, it appears that only about 4 percent of the *E. coli* strains are transient and not attributable to one specific source.

3.1.2 Ribotyping

The key methodological problem in tracing sources of microbial contamination in the environment used to be the lack of a universal single-reagent typing scheme for bacteria. This was overcome by the work of several investigators in the fields of population genetics, molecular systematics, and molecular epidemiology. In 1986, Grimont and Grimont showed that deoxyribonucleic acid (DNA) probes corresponding to specific regions of the rRNA operon could 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. To trace the indicator bacterium, *E. coli*, from water to its specific source, the bacterial strain must first be uniquely identified. Populations of *E. coli*, like other bacteria, are 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 inhabits another.

Ribosomal ribonucleic acids, 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, an rRNA 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 that the majority of patients contracted cystic fibrosis from one of two treatment centers. Moyer *et al*. (1992) used rRNA 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 moncytogenes* implicated in a food-borne illness and found that ribotyping was the preferred method due to its precision and reproducibility. Atlas *et al.* (1992) described the technology of ribotyping as applicable to the tracking of genetically engineered microorganisms in the environment.

Dr. Samadpour's MST method was developed on the basis of the principles of microbiology, epidemiology, molecular epidemiology, microbial population genetics, sanitary engineering, and hydrology. In any watershed, there are multiple contributing animal sources of microbial pollution, each of which has its own unique clones of bacteria that constitute their normal flora. Ribotyping is applied as part of an MST study in the following steps. First, collections of isolates from 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. Second, using an appropriate molecular subtyping method, all 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 water sample and a suspected animal source, the animal is implicated as a contributor of that specific strain of the bacteria to the polluted site. The relative contributions of various sources are quantified based on the fraction of isolates from a representative set

of ambient water samples that match ribotypes of resident strains from that source (human or nonhuman).

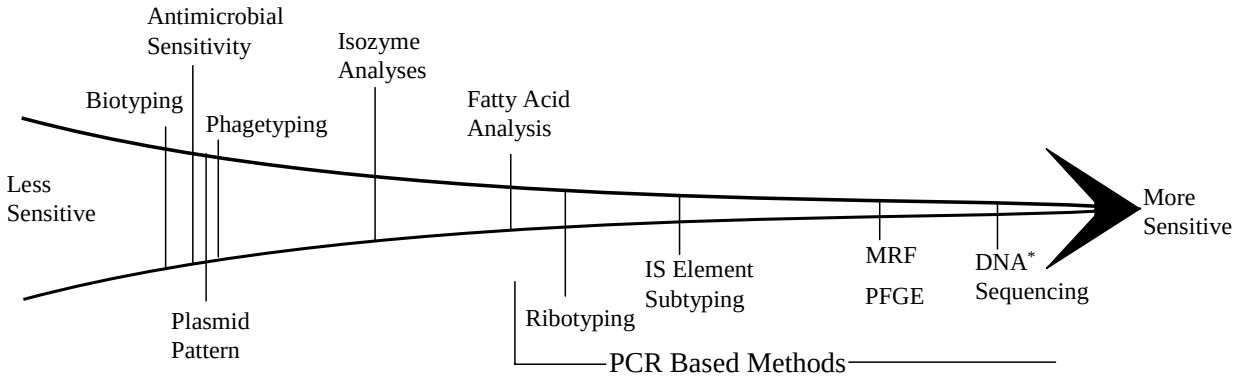
3.1.3 Antibiotic Resistance Analysis

Antibiotic Resistance Analysis (ARA) is another method commonly used in MST studies. The basic underlying assumption for ARA is that the use of therapeutic antibiotics in humans and both therapeutic and feed additive uses in animals would result in imposing selective pressure on the intestinal microflora of humans and animals. Due to differences in the types of antibiotics used in humans and animals, it should be possible to categorize the environmental isolates of bacteria, and on the basis of their antibiotic resistance profile (ARP), determine their host origin (Krumperman 1983; Kaspar *et al*. 1990; Wiggins 1996; Parveen *et al*. 1997; and Hagedorn *et al*. 1999).

Antibiotic resistance profiles are generated by measuring the ability of *E. coli* to grow in the presence of each of several antibiotics. In a manner similar to ribotyping, the antibiotic resistance profiles of *E. coli* isolates are compared to those from a library of known fecal sources. A given antibiotic resistance profile is seldom unique to a given animal or source. Instead, statistical methods are used to determine the most likely source. Also, the specificity of ARA is less than that of ribotyping; often, source classification is limited to human, livestock, and wildlife categories. However, ARA is less expensive than ribotyping.

Figure 3.1 MST Methods

Phenotype Based Methods



Genotype Based Methods

Source: Dr. Mansour Samadpour 2001.

Figure 3.1 above displays a conceptual sensitivity continuum of some of the widely used subtyping methods. Phenotypic based methods (methods based on the expression of phenotypes) are at the less sensitive domain of the continuum while genotypic based methods constitute the more sensitive end of the spectrum. The level of sensitivity depends upon the choice of gene(s) and the size of fragment(s) sequenced.

3.1.4 Study Design and Data Quality Objectives

A key factor influencing the study design and DQOs was the need to identify contributing sources of fecal coliform bacteria in the MRG watershed under wet-weather

conditions, and correlating these findings to the periods when elevated fecal coliform levels had been observed historically. The geographic scope of the project included a 40-mile stretch of the Rio Grande, from Angostura Diversion Dam near the southern border of the Santa Ana Pueblo downstream to the Isleta Diversion Dam near the northern border of Isleta Pueblo and contributing tributaries and watersheds.

During the first 2 years of the study, progress was hindered by a lack of rainfall. During the study, drought conditions intensified from "severe" to "extreme" or "exceptional." According to the National Weather Service, 2003 was the driest year since 1956. Rainfall sufficient to produce runoff was extremely rare and widely scattered. From June 2002 to December 2003, only 64 percent of the planned runoff-influenced ambient water samples were collected. Long-term forecasts indicated that the drought would likely continue. Given the limited amount of rainfall, the forecast of continuing drought, and other practical considerations to complete the necessary MST analyses, the DQOs were modified to also include identification of the contributing sources of fecal coliform bacteria under dry-weather conditions.

This project involved several steps:

- A sanitary survey of the watershed and a review of available data and literature to identify potential contributing sources of fecal bacteria to be considered.
- Development of watershed-specific libraries of ribotypes and antibiotic resistance profiles of *E. coli* isolated from fecal matter collected from known sources.
- Collection and culturing of a representative set of *E. coli* isolates from the waterbodies of concern under dry and wet-weather conditions.
- Determination of the ribotypes and antibiotic resistance profiles of these waterborne *E. coli* isolates, followed by matching to those from the known source library to identify the sources of each *E. coli* isolate.
- Quantification of the accuracy and precision of the ribotyping and ARA source determinations.
- Estimation of the relative source contributions of *E. coli* in the MRG watersheds, and the confidence of these estimates, based on the above measurements.

It is important to note that fecal coliform was cultured and enumerated from water samples, but *E. coli* was subjected to MST procedures. Fecal coliform are a method-defined group of coliform bacteria that can grow at the elevated temperatures inside the mammalian intestinal tract, whereas *E. coli* is a particularly abundant species of the fecal coliform group of bacteria. However, the fecal coliform group also includes many other species of bacteria, some of which are not exclusively fecal in origin. Fecal coliform, rather than *E. coli*, was enumerated in water samples because current New Mexico WQSs are based on fecal coliform concentrations. However, the ribotyping method must be applied to a single species of bacteria because inter-specific DNA differences likely outweigh the small differences between intra-specific strains of bacteria on which the MST method is based. *E. coli* were selected for MST in this project because they are exclusively endemic to the mammalian intestinal tract in temperate zones and they comprise a substantial fraction of the fecal coliform group. The fraction of fecal coliform verified to be *E. coli* in water samples were noted.

3.2 Library Development

On the basis of the watershed sanitary survey, review of existing data and information, and communications with stakeholders, the Parsons team identified the major potential sources of fecal pollution to the MRG and devised a plan to develop libraries of *E. coli* isolated from fecal samples from these potential sources. The goal was to develop a local library of 1,000 *E. coli* isolates from approximately 500 samples with, on average, two *E. coli* collected from each known source sample. The sources sampled included sewage, wildlife, avian, pet, livestock, and exotic species from the Albuquerque Zoo (Table 3.1). These locally collected isolates supplemented an extensive library of over 65,000 isolates ribotyped by Dr. Samadpour at MEI from hundreds of different species and sources, including many of the domestic and wild species found in the MRG watershed. Because the abundance of particular *E. coli* strains in animals is expected to vary with space and time, reliance on the MEI library alone would likely result in an unacceptably large percentage of *E. coli* with ribotypes for which sources could not be identified. In numerous studies, Dr. Samadpour found that sources can be identified for approximately 60-70 percent of *E. coli* isolates based on his nationwide library alone, but that with a local library, the identified percentage can be increased by 15 to 25 percent.

Known source samples were collected directly from the source when possible. An exception was human samples, which were collected from septage haulers, sanitary sewers, and WWTFs. Because sewage can also contain fecal matter from other sources, it will not be referred to as human in origin. However, it was expected that the bulk of fecal matter in most septage and sewage was human in origin. In addition to sewage and septage, the MEI library also contains samples directly from human feces. In many cases, wildlife and pet samples were collected indirectly, from “found” fecal samples. However, only samples for which the field biologists were confident of the source were collected. After a few reported cases of hanta virus in the area, it was decided for safety reasons not to collect additional fecal samples from mice and rats for the library, but to rely on the existing rodent *E. coli* library at MEI.

Fecal samples were collected aseptically into sterile test tubes, capped, and sealed. All sample containers were labeled with the following information: sample type, host species, sample date and time, sample location, and sampler’s initials. The sample information was recorded in a field log. Samples were then shipped on ice via overnight courier to MEI. Only a single sample was collected from each individual animal.

At the MEI laboratory in Seattle, Washington, fecal samples were plated on MacConkey agar and incubated at 35°C overnight. The next day three to five lactose fermenting, non-mucoid colonies that exhibited *E. coli* morphology were picked and re-plated on MacConkey agar for purification. Well-isolated non-mucoid colonies were picked from these plates and plated on tryptic soy agar. After overnight incubation at 35°C, each colony was tested by a Spot indol test using appropriate positive and negative controls. Indol positive cultures were further tested for the ability to utilize citrate using Simon citrate media. Indol positive, citrate negative colonies were identified as *E. coli* and were given isolate numbers and incorporated into the library. A portion of each E. coli isolated from the samples was stored at –80°C, in nutrient broth plus 15 percent glycerol.

In total, 579 fecal and sewage samples from known sources were collected from the Albuquerque vicinity for the library. These samples are summarized in Table 3.1. From these samples, a total of 1,733 colonies were picked and, ultimately 861 individual *E. coli* isolates were added to the known source library.

Table 3.1 Summary of Fecal Source Sampling for Library Development

Major Category	Category 2	Category 3	Total Samples Collected	Sample Target	Sample % of Target
Sewage	Septage haulers		26		
	Sanitary sewers		22		
	Wastewater treatment plant influent		28		
Sewage Total			**76**	**80**	**95%**
Livestock	Bovine	Beef cattle	36		
		Dairy cattle	14		
		Bison	5		
		Subtotal	55	55	100%
	Equine	Horse	50		
		Pony	2		
		Donkey	2		
		Mule	2		
		Subtotal	56	50	112%
	Poultry	Chicken	11		
		Turkey	5		
		Duck	4		
		Emu	5		
		Geese	2		
		Subtotal	27	15	180%
	Goats		14	4	350%
	Sheep		14	6	233%
	Hogs & pigs		4	5	80%
	Rabbits, domestic		4	5	80%
	Alpacas, llamas, & guanicos		3		
Livestock Total			**177**	**140**	**126%**
Pets	Cat		43	40	108%
	Dog		51	50	102%
	Other mammals		13	10	
	Reptile		3		
Pets Total			**110**	**100**	**110%**
Wildlife	Large herbivores	Deer	10		
		Elk	1		
		Pronghorn	1		
		Subtotal	12	15	80%
	Large carnivores	Bobcat	2		
		Cougar	3		

Major Category	Category 2	Category 3	Total Samples Collected	Sample Target	Sample % of Target
		Coyote	1		
		Wolf	4		
		Subtotal	10		
	Rodents	Prairie dog	12	3	
		Rabbit	18	10	
		Squirrel	1		
		Subtotal	31	53 (13)*	58% (238%)*
	Other mammals	Bats	2		
		Javelina	1		
		Raccoon	3		
		Subtotal	6	10	60%
	Water birds	Ducks/Geese	43		
		Cranes	13		
		Herons	2		
		Subtotal	58	50	116%
	Upland birds	Pigeons	16	10	
		Crows / ravens	7	5	
		Grackles	7	5	
		Hawks/Falcons	6		
		Owls	7		
		Roadrunner	3		
		Other	7		
		Subtotal	53	20	265%
	Reptiles		2		
Wildlife Total			**172**	**170 (130)***	**101% (132%)***
Zoo	Exotic species	mammals	36		
		birds	2		
	Native species	mammals	3		
	Mixed species		3		
Zoo Total			**44**	**20**	**220%**
Grand Total			579	520 (480)*	111% (121%)

* *Number in parenthesis is the revised target after decision to exclude small rodent trapping due to concern over hantavirus.*

3.3 Ambient Water Sampling Plan

The stakeholders and sponsors of this project identified 30 water quality monitoring sites in the MRG watershed to be sampled in this project, based on their needs for information to assist in the design of water quality management measures. Eight stations were selected to represent the ambient conditions along the MRG, from Angostura Diversion Dam at the northernmost extent of the watershed, to Isleta Diversion Dam at the

southernmost extent. Other stations were selected on major arroyos or drains flowing to the MRG, or at pump stations that drain watersheds where elevated fecal coliform levels have been observed. The stations and the properties of their subwatersheds were described previously under Section 2.3. The subwatersheds of the tributaries represent a wide variety of land use types and potential fecal sources, and drain a large percentage of the MRG watershed within the study area.

To identify fecal coliform sources associated with rain events, water samples were initially collected only during, or within 24 hours after, a rainfall event. Because drought conditions prevented obtaining the required number of samples, the monitoring plan was revised to include sample collection during February and March 2004 to quantify fecal coliform sources under dry weather conditions. The samples were collected under base flow conditions before water releases for irrigation use began, at stations where water is often present under base flow conditions**.** These stations included the eight Rio Grande stations, as well as the North Diversion Channel and South Diversion Channel.

The sampling design called for each location to be sampled on five dates (events). In addition, two "integrator" sites were to be sampled on 10 dates to provide more precise source contribution estimates. These integrator sites were located on the Rio Grande, at the upper and lower ends of the study area. The upper integrator site, at the Angostura Diversion Dam on the Rio Grande, provided an estimate of background fecal source contributions in the Rio Grande watershed upstream of the study area. The lower integrator site, at the Isleta Diversion Dam on the Rio Grande, quantified fecal source contributions from the entire study area.

Because fecal coliform populations have been found to vary on fine spatial and temporal scales, Parsons increased sampling representativeness by collecting composite water samples from six sub-samples collected several minutes and/or several meters apart. Parsons' field staff followed the field sampling procedures for field and conventional chemical parameters documented in the 2002 QAPP for Water Quality Management Programs of the NMED SWQB and the Scientific Laboratory Division of the New Mexico Health Department. Sampling personnel wore clean, disposable, powder-free gloves while collecting samples. At many sites, water samples were collected directly from the stream (approximately 1 foot below the surface) into sterile wide-mouthed polypropylene bottles supplied by the culturing laboratory. Containers were rinsed in sample water before the grab samples were collected. In cases where, for safety reasons, it was inadvisable to enter the stream bed, staff used a clean bucket and rope from a bridge to collect the samples from the stream, and poured the water into the sample bottles. Care was taken to avoid contaminating the sample and ensure that the bucket and rope did not come into contact with the bridge. The bucket was thoroughly rinsed between stations, and sanitized with a bleach- or isopropyl alcohol-soaked wipe. The first bucketful of water collected from a bridge was used to rinse the bucket and the sampler's gloved hands. Samples were collected from subsequent buckets of water.

Following collection, samples were delivered within 24 hours to the Assaigai Analytical Laboratory in Albuquerque for fecal coliform culturing and enumeration via the membrane filter-based Standard Method 9222D (APHA 1995). Following their incubation and enumeration, Assaigai shipped the fecal coliform culture plates on ice overnight to MEI for ribotyping and ARA.

Water samples were received by MEI in the form of mFC agar plates, with (usually) at least 12 colonies per plate. Twelve or more randomly-selected blue colonies were picked from mFC plates corresponding to each sample, and each was plated onto MacConkey agar (Difco Laboratories, Detroit, Michigan, USA) for purification as individual *E. coli* colonies. At this stage, each of the colonies picked from a given sample bore the Sample ID number and an accession number. A single well-isolated non-mucoid colony that exhibited *E. coli* morphology was picked from each MacConkey plate and plated on Triptic soy agar. After overnight incubation at 35°C, each culture was tested by the Spot indol test using appropriate positive and negative controls. Indol positive cultures were further tested for the ability to utilize citrate using the Simon citrate media. Indol positive, citrate negative colonies were identified as *E. coli* and given isolate numbers.

As noted, for each water sample, MEI performed ribotyping and ARA on approximately 12 *E. coli* isolated from the culture plates. Given five sampling events per site, this provided 60 isolates for each site. For the two integrator sites, 120 isolates per site were typed. The number of isolates typed impacts the confidence level around the estimate of the fecal source contributions at each site. Table 3.2 shows the predicted confidence levels around the source contribution estimate for 30, 60, and 120 isolates per site for major sources comprising 20 to 50 percent of the total fecal loading. Typing 60 isolates permitted Parsons to quantify the relative contribution of a source comprising 20 percent of the total *E. coli* loading with a precision of plus or minus 10 percent at the 95 percent confidence level.

Table 3.2 Confidence Intervals of Source Contribution Estimates

True Source % Contribution to Fecal Loading	Number of Isolates Ribotyped	Predicted 95% Confidence Interval of Estimated Source Contribution
20%	30	5.7 - 34.3%
	60	9.9 – 30.1%
	120	12.8 - 27.2%
30%	30	13.6 - 46.4 %
	60	18.4 - 41.6%
	120	21.8 - 38.2%
40%	30	22.5 - 57.5%
	60	27.6 - 52.3%
	120	31.2 - 48.8%
50%	30	32.1 - 67.9%
	60	37.3 - 62.7%
	120	41.1 - 58.9%

3.4 MST Methods

As mentioned previously, MEI characterized the isolated *E. coli* strains using both ribotyping and ARA methods. Results from both methods will be described individually. In addition to the usefulness of the direct head-to-head comparison between ribotyping and ARA, application of two independent methods can validate the results, increasing stakeholder confidence in the outcome. Second, because any one method may not perform completely successfully in all samples of a given study, a second method provides back-up to ensure the study will generate useful results. For instance, if an *E. coli* ribotype from a water sample does not match a ribotype from a known source

species, ARA may be able to at least indicate whether the source was wildlife, livestock, or human. A third benefit of using two methods is that the results may be more directly compared to other studies.

3.4.1 MST Ribotyping Procedure

MST was performed at the Seattle, Washington laboratories of MEI, Inc.

Genomic DNA Isolation and Restriction Endonuclease Digestion

Confluent growth was scraped with a sterile flat-head toothpick and suspended in 200 microliters (μL) 50mM Tris, 50mM EDTA (pH 8.0). Six hundred μL more of 50mM Tris, 50mM EDTA was then added and the suspension was thoroughly mixed by pipetting up and down. Then 45 μL 20% sodium dodecyl sulfate (SDS) and then 10 μL proteinase K (20 μg/mL; Pharmacia, Piscataway, N.J.) were added. This solution was then incubated at 40°C for 1 hour. After an equal volume of phenol was added to each tube, samples were vortexed, and centrifuged for 5 minutes. The top layer was extracted, and an equal volume of chloroform was added. The preparation was vortexed again, centrifuged, and extracted. Two and a half volumes of absolute ethanol were added and the DNA was precipitated out and spooled onto a glass capillary pipette. The DNA was washed with a few drops of absolute ethanol, dried, and re-suspended in 50μL distilled water (dH_2O).

Separate restriction endonuclease digestion reactions were set up using *Eco*R1 and *Pvu*II, 10 units/μL (Boehringer Mannheim, GmbH, Germany) as instructed by the manufacturer using 2 μL DNA. They were incubated at 37°C overnight. The samples were then centrifuged and 0.5 μl of enzyme was added. The samples were re-incubated at 37°C for a minimum of 3 hours. They were centrifuged again and 3 μL stop dye was added. Every batch of restriction enzyme reaction contained two reactions with a positive control strain that was included on two lanes on each gel.

Gel Electrophoresis and Southern Blot Hybridization

Samples were run on a 0.8% agarose gel in 1X Tris-borate-EDTA at 22 volts and 17 milliamps, for 17 hours. λ HindIII was used as a size standard along with a known *E. coli* isolate designated as 3915. Each agarose gel was assigned a number, and when more than one gel was run, the position of the first standard reference strain was changed in each gel (1st lane on the first gel, to the Nth lane on the Nth gel). Electrophoresis gels were stained in ethidium bromide. If two gels were stained in a single container, one corner of the gel with the higher number was clipped. The label for each gel was also transferred to the staining container. Each gel was then photographed and a hard copy of the print was labeled with the gel sheet (containing the isolate numbers loaded on each lane, and the enzyme used to cut the DNA, plus date, gel number, voltage, current, gel strength, buffer strength, and electrophoresis time information). After photography each gel was returned to the same staining container.

The DNA fragments were then transferred to a Nitran filter (Schleicher & Schuell, Keene, N.H.), baked at 80°C for 1 hour, and probed with ^{32}P-labeled copies of *E. coli* rRNA, which were made by extension of random hexanucleotide primers using Avian Myeloblastosis Virus reverse transcriptase (Stratagene, La Jolla, California) under conditions specified by the supplier. Each membrane filter was labeled with the gel

number, restriction enzyme designation, date, and technician's initials. Hybridization was 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) were used to block non-specific binding. Three washes were done with a solution of 2X SSC and 0.1% SDS, once at 25°C for 20 minutes and twice at 65°C for 20 minutes to wash off low-homology, non-specific binding. Blots were then exposed with an intensifying screen to x-ray film (Kodak, Rochester, New York) for 24 hours at -70°C. Two to three exposures were done to ensure all possible bands would show up.

All reagents and buffers were made according to formulas in the MEI standard operating procedures. Reagents and buffers were tested for sterility.

Restriction Fragment Length Polymorphism Analysis

Each ribotype was then analyzed by assigning an alphanumeric pattern based on the distance between the bands. Bands more than 3 mm apart were counted as singles while bands that were within 3 mm of each other were counted as doubles or triples. For example, two bands that were closer than 3 mm were designated "2" and a group of three bands with 3 mm or less between each band were designated "3." A "1" designated a single band more than 3 mm distant from another band. Each unique banding pattern was called a ribotype and assigned an alphanumeric pattern.

Two isolates that had the same numeric value but different banding patterns were assigned letters to differentiate the two ribotypes. For example, two isolates with an identical numerical pattern of 2122111, but with the bands shifted so the two isolates did not have identical banding patterns, were labeled 2122111A and 2122111B.

The ribotypes were then entered into a Microsoft® Access™ database and compared to the other ribotypes of known source in the library database. Ribotype patterns that numerically appeared to be similar were compared side-by-side visually to judge matching.

Isolates with the same *Pvu*II and *Eco*R1 ribotypes are deemed to be members of the same ribogroup. Using this approach, only isolates with two identical ribotypes were grouped together.

3.4.2 ARA Procedure

Antibiotic resistance and susceptibilities were determined by disk diffusion assay, following the standard protocol of the National Committee for Clinical Laboratory Standards (NCCLS 1999), included in Appendix A. The specific antibiotics and concentrations used, in antibiotic disks obtained from Remel, Inc. (Lenexa, KS), were amikacin (30 µg), ampicillin (10 µg), ampicillin/sublactarn (10/10 µg), ceftriaxone (30 µg), chloramphenicol (30 µg), ciprofloxacin (5 µg), gentamicin (10 µg), kanamycin (30 µg), nalidixic acid (5 µg), streptomycin (10 µg), tetracycline (30 µg), and trimethoprin (5 µg). These antibiotics are commonly used for human and veterinary purposes, and/or in animal feed. The control organism ATCC 25923 was used with each susceptibility assay. Using this method, each *E. coli* is tested for growth or inhibition in the presence of each of the 12 antibiotics. Growth of *E. coli* indicates it has developed some resistance to the antibiotic, while inhibition of growth indicates susceptibility. A

12-character ARP code is assigned to each *E. coli* indicating its resistance or susceptibility to the 12 antibiotics, in the order listed above. A "0" indicates inhibition, while a "2" indicates growth (resistance). For example, the antibiotic resistance profile 200000000020 would indicate an *E. coli* with resistance to amikacin and tetracycline, but susceptible to all other antibiotics tested. Unlike the ARA protocols utilized by some other investigators, this method does not utilize quantification of the intermediate degree of growth and inhibition based on the diameter of the zone of inhibition around an antibiotic disk, but simply classifies growth or inhibition.

3.5 Quality Assurance/Quality Control Measures

Along with an estimate of the relative contributions from various fecal sources in the watershed, it is important to understand the uncertainty of those estimates. Precision, accuracy, sensitivity, completeness, and representativeness are critical data quality issues affecting uncertainty. Representativeness was controlled by developing an environmental monitoring program characteristic of actual environmental conditions. Accuracy, precision, sensitivity, and completeness were similarly controlled through careful planning, but also were quantified via quality control (QC) measures. These QC measures included analysis of blank samples, replicate samples, and known standards (in MST, samples of known origin). A detailed QAPP was developed to ensure the quality of data produced by the project.

3.5.1 Quantification of Accuracy and Precision in Ribotyping and ARA Source Determinations

MST does not lend itself easily to the same QC methods as chemical quantification. Blank samples may be irrelevant, and replicate water samples may often yield different *E. coli* strains. Parsons quantified method accuracy and precision through a special QC study with "blind" safeguards, as practiced in epidemiological QC. MEI prepared triplicate cultures of 100 *E. coli* isolates randomly selected from various sources in the MEI library, including human (10), bovine (31), seagull (17), dog (18), sanitary sewage (13), and ambient water (12) sources. The cultures were placed in identical slant tubes, each with a removable label indicating its source and the isolate number (1-100). These tubes were mailed to the Parsons quality assurance (QA) manager. The Parsons QA manager replaced each tube label with a new label, numbered from 1 to 300 in random fashion, and recorded both the old and new label numbers for each tube in a key. After verifying there was no way to distinguish the tubes, the Parsons QA manager then sent the 300 slant tubes back to MEI as unknowns. The *E. coli* in each tube was then processed through the ribotyping and ARP procedures by MEI in a blind fashion; that is, MEI did not know the source. MEI then identified the original isolate number of each culture through ribotyping and ARA and sent the results to the Parsons QA manager, who then evaluated and reported the accuracy and precision of the two methods. Precision was evaluated as correct assignment of the three identical triplicates of each of the 100 unknown cultures to the same original isolate number. Accuracy was evaluated as correct assignment of unknown cultures to their original isolate number. The results are shown in Section 4.1.

3.5.2 Bottle and Equipment Blanks

An equipment blank is a sample of reagent water poured into a sample bottle, or poured over or pumped through a sampling or analysis device. It is collected in the same type of container as the environmental sample, preserved in the same manner, and analyzed for the same parameter. In addition to regularly collected bottle and equipment blanks, laboratory equipment blanks are prepared at the laboratory where collection materials are cleaned between uses. These blanks document that the materials provided by the laboratory are free of contamination. The QC check is performed with each new batch of equipment or bottles, as documented in the laboratory quality assurance manual. The analysis of equipment blanks should yield values less than the minimum analytical level (MAL). When target analyte concentrations are very high, blank values must be less then 20 percent of the lowest value of the batch.

3.5.3 Field Duplicates

Field duplicates utilized in this study were split replicates, collected by splitting a composite sample, immediately after compositing in the field, into two bottles. Field duplicate samples were sealed, handled, stored, shipped, and analyzed in the same manner as the primary sample. Because only a small subset of the bacterial colonies cultured from a water sample are typed, and there are often many different bacterial sources impacting a body of water, it is expected that two samples will yield different results, depending on which colonies are selected for typing. Thus, duplicate results from a single sampling event are not a meaningful measure of quality; they are expected to disagree. However, the overall source identification, based on multiple sampling events and many isolates, should be consistent from duplicate samples. If this difference is large, it may indicate problems with sampling, analysis, experimental design, or typing. In this study, field duplicate samples were collected at a single sampling station: Rio Grande at Isleta Diversion Dam. One DQO of this study was to elicit a 75 percent agreement between the sources identified, as illustrated in the example below:

Isolate #	Duplicate 1 Source	Duplicate 2 Source	Same?
1	Cattle	Cattle	Yes
2	Cattle	Cattle	Yes
3	Cattle	Cattle	Yes
4	Cattle	Human	No
5	Cattle	Dog	No
6	Dog	Dog	Yes
7	Dog	Dog	Yes
8	Dog	Dog	Yes
9	Horse	Human	No
10	Human	Human	Yes
11	Human	Human	Yes
12	Human	Human	Yes
			9/12 = 75% agreement

Actual precision and accuracy results were in 100 percent agreement and are described in Section 4.1.

3.5.4 Field Blanks

Field blanks consisted of sterile buffer water taken to the field and transferred to the appropriate container in precisely the same manner as a sample during the course of a

sampling event. They were used to assess contamination from field sources such as airborne materials, carryover from prior sampling sites, and containers. The analysis of field blanks should yield values less than the MAL. When target analyte concentrations are high, blank values should be less then 20 percent of the lowest value of the batch. Field blanks were to be collected at a frequency of 5 percent or greater through fecal coliform culturing and quantification steps, but not processed by ribotyping or ARA.

3.5.5 Completeness

Data completeness is simply a measure of the fraction of data that meets project DQOs. In water sampling and fecal coliform culturing and enumeration, completeness is affected by the number of water samples collected and analyzed that meet the project's DQOs relative to the original data collection planned. Data completeness can be affected by lack of suitable conditions for sampling (*e.g*, lack of runoff-influenced water at sampling sites), and by sampling or laboratory QC failures. A goal of the project was collection and analysis of 90 percent or more of the planned samples.

In ribotyping and ARA, data completeness is most affected by the number of ribotypes observed as unknowns in water samples that match ribotypes and ARPs in the known source library. This completeness can be reduced by 1) an abundance of transient clones among the *E. coli* isolated from water, or 2) a known source library of insufficient size that does not contain ribotypes or ARPs from known sources matching the *E. coli* isolated from water. Thus, a large library is important. Another DQO of the project was that a fecal source be identified for at least 70 percent of the planned number of *E. coli* isolates.