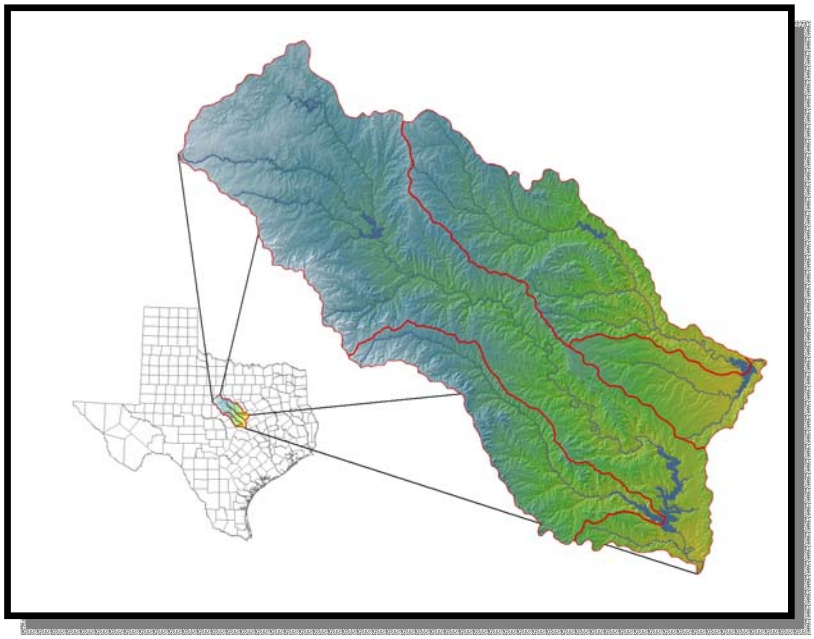
ASSESSMENT OF BACTERIAL SOURCES IMPACTING LAKE WACO AND BELTON LAKE



Prepared for:

TEXAS FARM BUREAU

Prepared by:

PARSONS

TEXAS A&M UNIVERSITY AGRICULTURAL RESEARCH & EXTENSION CENTER AT EL PASO, TEXAS AGRICULTURAL EXPERIMENT STATION

TEXAS A&M UNIVERSITY – CORPUS CHRISTI

TEXAS A&M UNIVERSITY

Funding for this Project was provided by the Texas State Soil and Water Conservation Board through a Clean Water Act §319(h) grant from the U.S. Environmental Protection Agency, and funds from the Texas Farm Bureau, Brazos River Authority, and the City of Waco.

FEBRUARY 2006

ASSESSMENT OF BACTERIAL SOURCES IMPACTING LAKE WACO AND BELTON LAKE

Prepared for:

TEXAS FARM BUREAU

Prepared by:

PARSONS

TEXAS A&M UNIVERSITY AGRICULTURAL RESEARCH & EXTENSION CENTER AT EL PASO, TEXAS AGRICULTURAL EXPERIMENT STATION

TEXAS A&M UNIVERSITY – CORPUS CHRISTI

TEXAS A&M UNIVERSITY

FEBRUARY 2006

Funding for this Project was provided by the Texas State Soil and Water Conservation Board through a Clean Water Act §319(h) grant from the U.S. Environmental Protection Agency, and funds from the Texas Farm Bureau, Brazos River Authority, and the City of Waco.

EXECUTIVE SUMMARY

This assessment report describes the investigation and categorization of the sources of fecal contamination in Lake Waco, Belton Lake, and portions of major tributaries to those lakes. While both Lake Waco and Belton Lake meet Texas Surface Water Quality Standards for contact recreation and public water supply, concerns have been expressed over potential fecal contamination from animal agriculture operations in their watersheds. This assessment project was initiated and coordinated by the Texas Farm Bureau (TFB), and funded by the Texas State Soil and Water Conservation Board (TSSWCB) through a Clean Water Act §319(h) grant from the U.S. Environmental Protection Agency (USEPA). Other financial, technical, and in-kind support for the project was provided by the TFB, the City of Waco, and the Brazos River Authority (BRA). Scientific and technical collaborators on the project included Parsons Water and Infrastructure Inc. (Parsons), the Texas A&M University Agricultural Research Extension Center at El Paso (EP AREC), Texas A&M University-Corpus Christi (TAMU-CC), and Texas A&M University (TAMU).

The two primary objectives of the project were to use bacterial source tracking (BST) methods to (1) assess the relative contributions of fecal bacteria from livestock (particularly cattle), humans, wildlife, and other animals in Lake Waco and Belton Lake; and (2) develop libraries of known bacteria sources, genetic and biochemical, which can be used in determining the origin (animal or human) of fecal contamination in surface water. As this was one of the first studies of its kind in the state, a secondary objective was to evaluate and compare several analytical methods for BST to identify the optimal method, or combination of methods, for future application in Texas. Given these objectives it should be noted that throughout this report data results are discussed and presented in relation to two separate but inter-dependent themes: 1) data results that support conclusions associated with different BST analytical methods; and 2) data results that support conclusions specific to bacteria sources identified in each watershed.

Water Sampling

The collection of all water quality samples was governed by approved sampling methods summarized in a TSSWCB- and USEPA-approved Quality Assurance Project Plan (QAPP). To obtain representative results, including wet and dry conditions and seasonal variation, the ambient water sampling was performed on a routine monthly schedule from September 2003 through June 2004, to capture dry and runoff-influenced events at their natural frequency. Four stations in Lake Waco, two stations on the North Bosque River upstream of Lake Waco, four stations in Belton Lake, and one station on the Leon River upstream of Belton Lake were sampled for *E. coli* on 10 dates. One of the 10 sampling events (November) appeared to be strongly influenced by rainfall runoff, based on streamflow increases as well as rainfall in the days just before the sampling event, while two other events (January and March) exhibited minor runoff influences at some sites.

Following collection, water samples were delivered to the City of Waco Laboratory for *E. coli* culturing and enumeration via the membrane filter modified mTEC method. A 6-hour holding time for sample delivery to the laboratory and initiation of analysis was maintained. Following incubation and enumeration using USEPA Method 1603, the

modified mTEC plates with *E. coli* colonies were shipped on ice overnight to the BST laboratory at EP AREC for *E. coli* isolation, confirmation, and archival. A total of 650 water samples were collected, of which 412 were positive for *E. coli*. One to twelve *E. coli* were isolated from each positive water sample for a total of 631 isolates. Of these, 555 were analyzed by all four BST methods.

BST Library Development

The BST techniques used in this study are library-dependent. Therefore, it was necessary to construct reference libraries of genetic and phenotypic fingerprints for *E. coli* isolated from known sources (*e.g.,* domestic sewage, livestock, and wildlife) derived from each watershed. By matching the fingerprints of *E. coli* isolated from water samples with the fingerprints in the known source libraries, the likely animal or human (domestic sewage) fecal sources of surface water contamination can be determined.

The primary sampling design consideration for this BST study was to obtain as many unique *E. coli* library isolates as possible from individual animals or samples to represent the diversity and abundance of fecal contamination sources occurring in Lake Waco and Belton Lake watersheds. Potential fecal sources were identified through a sanitary survey conducted by Parsons. Between October 2003 and October 2004, a total of 994 fecal samples were collected from known sources. Municipal wastewater treatment plant influent/effluent and septage samples (collectively referred to as "domestic sewage"), livestock, wildlife, and pet fecal samples, were obtained from a variety of sources throughout the Lake Waco and Belton Lake watersheds based on the sanitary survey. An additional 100 *E. coli* isolates from south Texas non-avian wildlife from a previous BST study were also included (Mott and Lehman 2001).

Typically one to five *E. coli* isolates were archived for each source sample, while one to 12 isolates per water sample were archived. These frozen cultures can remain viable for years and can be easily shared with other researchers.

E. coli isolates from known source samples were screened using a repetitive sequence polymerase chain reaction method (ERIC-PCR). ERIC-PCR is a genetic fingerprinting method used in previous BST studies as well as many microbial ecology and epidemiological studies. In total, 2,275 *E. coli* isolates from know source samples were screened using ERIC-PCR. After determining the number of different *E. coli* ERIC-PCR types and exclusion of clones for each sample, one to three isolates per sample were selected for inclusion in the library. A total of 883 *E. coli* isolates were selected for the library and analyzed by all four BST methods.

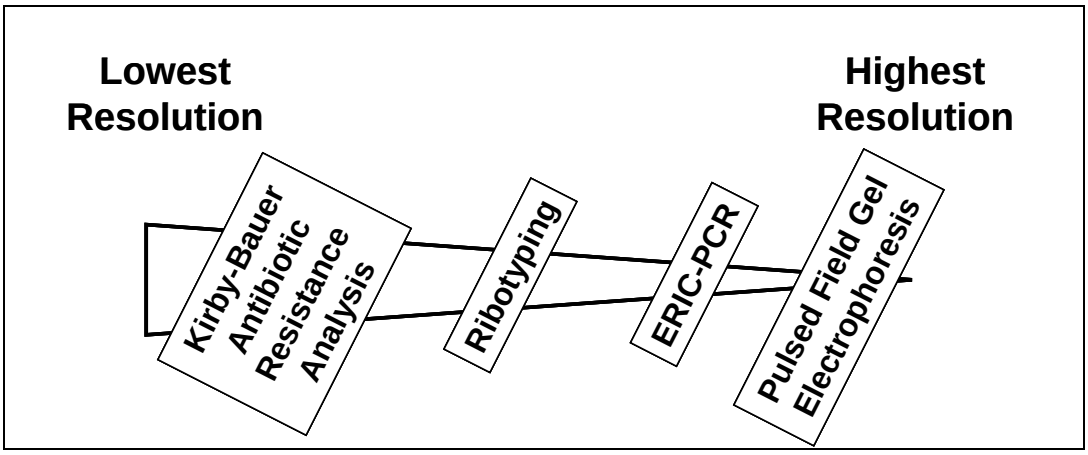
BST Methods

All methods selected for this project were used in previous BST studies, as well as other microbiological studies, and have been published in peer-reviewed journals. They span the spectrum in their ability to resolve differences in related bacterial strains, technical training and labor required, equipment cost, reagent cost, sample throughput, and ease of data analysis.

Figure ES-1 displays a conceptual sensitivity continuum of the BST methods used in this study. In this study, the combination of the DuPont RiboPrinter System, PFGE, ERIC-PCR, and KB-ARA allowed development of one of the most comprehensive

E. coli BST libraries and one of the first side-by-side comparisons of these BST methods. Recent BST studies also suggest that a combination of complementary methods may be the best approach for accurately identifying sources of contamination.

Figure ES-1 Discriminatory Capabilities of Different BST Methods



The validity of this study and the conclusions drawn from the results are strengthened through the use of the multiple techniques, and in particular, composite data sets. Further, by using standardized methods, the library can be expanded through future projects and the data shared with other BST investigators and regulatory agencies. Peer-reviewed publication of project results is also a goal, and journal manuscripts are being prepared as of the date of this report.

BioNumerics software (Applied Maths, Austin, TX) was used to analyze the BST data for this project. This software is currently used in many BST studies, especially for molecular fingerprint methods. For this project, BioNumerics was used in three ways: 1) processing gel images, 2) determining the relationships between the isolates by comparing their molecular fingerprint patterns for development of the BST libraries, and 3) for identification of water isolates.

Fingerprint patterns were compared individually for each of the molecular methods (ERIC-PCR, RiboPrinting, and PFGE) using the BioNumerics software. An equivalent comparison of the phenotypic KB-ARA profile treated the zone of inhibition measurements for each of the 20 antibiotics as character data to compare isolate profiles using the BioNumerics software. BioNumerics has the unique ability to allow the construction of composite data sets. A composite data set takes into account all four fingerprint profiles (ERIC-PCR, RiboPrinting, PFGE, and KB-ARA) or various combinations of these profiles for each isolate. Each BST method describes a different trait or aspect with different degrees of resolution or discrimination. An analogy would be the composite sketch of a suspect gathered from the descriptions of eye-witnesses to a crime, each having a different perspective. Composite data sets for only known library isolates and water isolates that had patterns for all four BST methods were created.

Host sources were divided into seven groups, 1) domestic sewage; 2) pet; 3) cattle; 4) other livestock, avian; 5) other livestock, non-avian; 6) wildlife, avian; and 7) wildlife, non-avian. The division of host sources into these particular classes was based on

discussion with project participants and anticipated usefulness for the development of Best Management Practices (BMP).

Quality Control

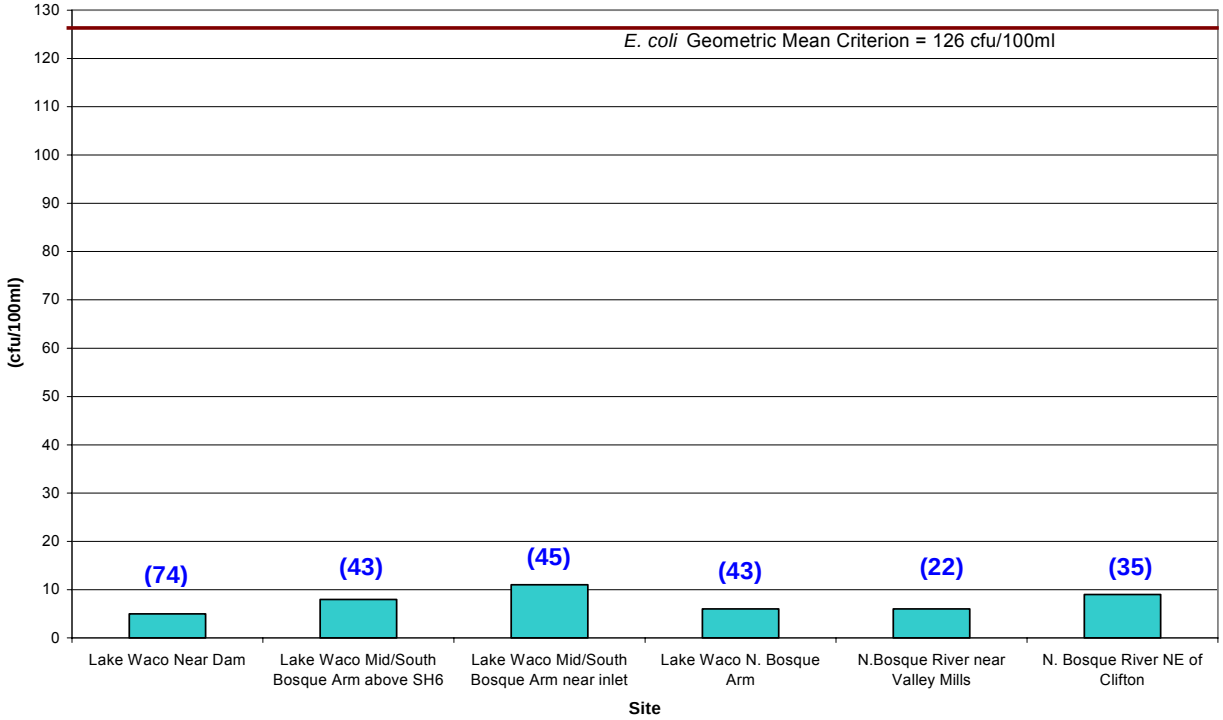
Standard laboratory practices were followed to assure quality data. Each laboratory method included additional required Quality Control (QC) measures. BST does not lend itself easily to the same QC methods as chemical quantification because each measurement is essentially qualitative, not quantitative. Therefore, in this study laboratory method accuracy and precision were quantified through a special QC study with blinded safeguards. The ERIC-PCR, RiboPrinting, and PFGE techniques performed equally well, with 100 percent identification of replicate isolates (precision) and 70 to 90 percent accuracy in identification of replicate isolates to specific QC library isolate (method accuracy) and correct source class (source identification accuracy). On the other hand, the KB-ARA scored only 40 percent for identification of replicate isolates (precision) and 50 percent for method and source identification accuracy. Most important, however, are the four-method composite data set results. The composite data set results correctly identified 100% of the replicate QC cultures (precision), and had 100% accuracy for *E. coli* strain and source class identification of the isolates. Therefore, the composite four-methods performed better than any single method.

The use of four different BST methods in this study provided valuable insights to BST methods. This is believed to be one of the first studies to directly compare BST methods by using the same collection of *E. coli* isolates. The methods used covered the spectrum in cost, ease of use, and discriminatory ability. As hypothesized, data from the combined methods (composite data set) were more useful than any individual method. Further, congruence measurements suggest that an ERIC-PCR and RiboPrinting composite data may be as useful as the four methods combined. In future studies, ERIC-PCR and RiboPrinting may be the methods of choice, and the other methods used to further characterize specific groups of isolates as needed. This would provide for cost, labor, and time savings while not compromising integrity of the BST result.

Levels of E. coli

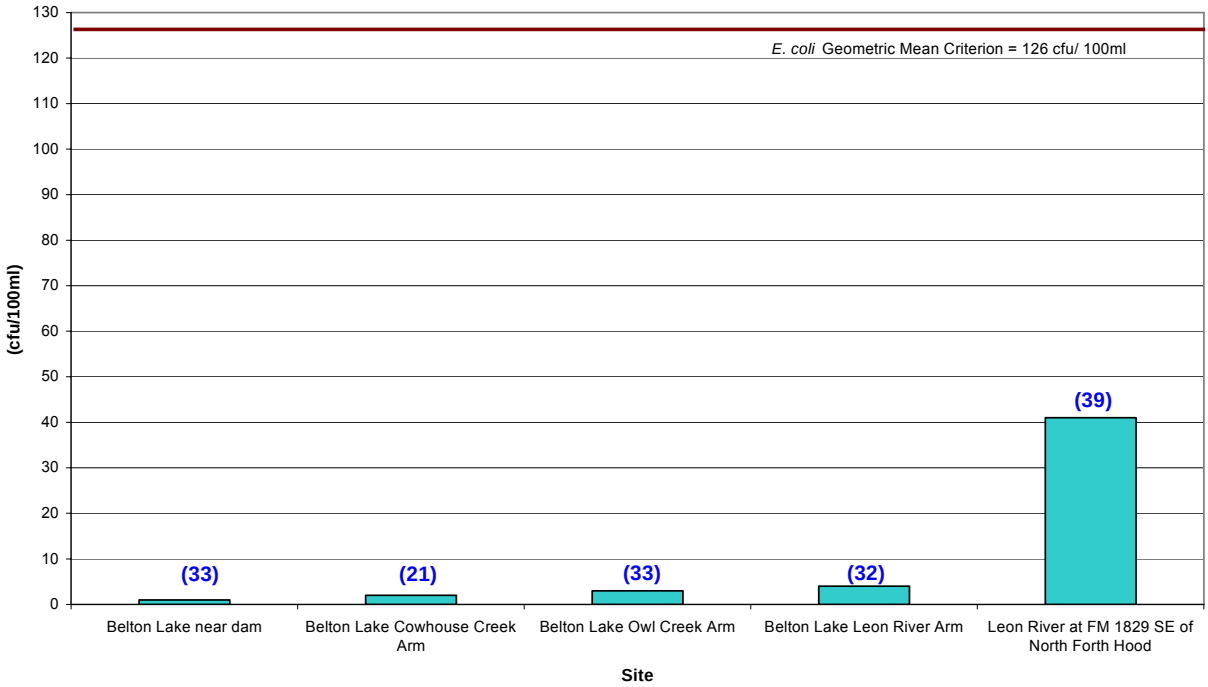
The *E. coli* samples collected during the study were assessed in accordance with the Texas Commission on Environmental Quality's Guidance for Assessing Texas Surface and Finished Drinking Water Quality Data, 2004. Overall, the levels of *E. coli* observed in the samples collected from Lake Waco and North Bosque River were far below the contact recreation geometric mean water quality criterion of 126 *E. coli* per 100 ml. The geometric mean concentrations ranged from 5 cfu/100 ml at the Lake Waco site near the dam to 11 cfu/100 ml near the inlet of the Middle and South Bosque Rivers (Figure ES-2).

Figure ES-2 Geometric Mean of *E. coli* Concentrations - Lake Waco Sampling Sites, September 2003 through June 2004



In Belton Lake, *E. coli* levels were very low throughout the lake. *E. coli* were absent from 131 (52%) of the 250 samples collected. Geometric mean *E. coli* concentrations ranged from 1 cfu/100 ml near the dam to 4 cfu/100 ml in the Leon River arm. The highest *E. coli* levels of any site sampled in this study were observed in the Leon River upstream of Belton Lake. Concentrations ranged from 0 to more than 20,000 cfu/100 ml, and the geometric mean *E. coli* concentration was 41 cfu/100 ml (Figure ES-3).

Figure ES-3 Geometric Mean of *E. coli* Concentrations - Belton Lake Sampling Sites, September 2003 through June 2004



BST Results: Source Identification of Water Isolates

To facilitate the interpretation of results, it is helpful to have some understanding of the uncertainties involved in the source estimates. As the number of isolates identified increases, so does the confidence in the estimate of the source contribution. At select individual sampling sites a lower confidence exists in the precision of the source category percentages since only a few dozen *E. coli* isolates were identified.

There was a wide variety of contributing sources to each lake and at each sampling station, and no single source category dominated. Considering all Lake Waco and North Bosque River sites and dates combined (Figure ES-4), wild birds were identified as the contributors of 23 percent of the *E. coli* to water, followed by non-avian wildlife at 17 percent of the *E. coli*. Thus, wildlife contributed approximately 40 percent of the *E. coli* to Lake Waco. Livestock contributed 29 percent of the *E. coli*, with cattle contributing 16 percent, other livestock - avian at 3 percent, and other non-avian livestock at 10 percent. Sewage was identified as the source of an estimated 17 percent of the *E. coli*. Pets (dogs and cats) were estimated to contribute approximately 3 percent of the *E. coli*. The source of 11 percent of the *E. coli* could not be identified with acceptable confidence.

Considering all Belton Lake and Leon River sites and all dates combined (Figure ES-5), wild birds were identified as the contributors of 28 percent of the *E. coli* to water, followed by non-avian wildlife at 21 percent of the *E. coli*. Thus, wildlife contributed approximately half of the *E. coli* to Belton Lake. Livestock contributed 32 percent of the *E. coli*, with cattle contributing 17 percent, other livestock - avian at 2 percent, and other non-avian livestock at 13 percent. Sewage was identified as the source of an estimated

11 percent of the *E. coli*. Pets (dogs and cats) were estimated to contribute approximately 3 percent of the *E. coli*. The source of 5 percent of the *E. coli* could not be identified with acceptable confidence.

Figure ES-6 presents the same data in a different format to provide a side-by-side comparison which demonstrates that the sources of *E. coli* in Lake Waco and Belton Lake were similar. Sewage contributed a larger percentage of *E. coli* in the Lake Waco watershed, and wildlife represented a larger contribution in the Belton Lake watershed, but the differences are not statistically significant at the 95 percent confidence level.

Although source contribution estimates for individual sites are too imprecise for statistically significant comparisons, the highest sewage contributions were seen in Lake Waco near the dam, a site near heavily populated areas of the City of Waco. The highest contributions from cattle were observed in the Leon River arm of Belton Lake and the Middle/South Bosque River arm of Lake Waco. The largest contributions from other livestock - avian were observed in Belton Lake near the dam, and the highest contributions from other non-avian livestock were observed in the Leon River near north Fort Hood. The highest avian wildlife contributions were observed in the North Bosque River near Valley Mills, perhaps due to a large swallow population observed to be nesting under the bridge. Non-avian wildlife contributions were most substantial in the Owl Creek arm of Belton Lake and near the Belton Lake dam. The highest pet contributions were observed in the North Bosque River near Valley Mills and in the Cowhouse Creek arm of Belton Lake. At four of the 11 stations more than 10 percent of the isolates were not identifiable with Lake Waco Middle/South Bosque River arm at SH 6 having the highest percentage at 26.

Figure ES-4 Source Contribution Estimates using Four-Method Composite Data for all Lake Waco and North Bosque River Sites and all Dates Combined (348 *E. coli* Isolates)

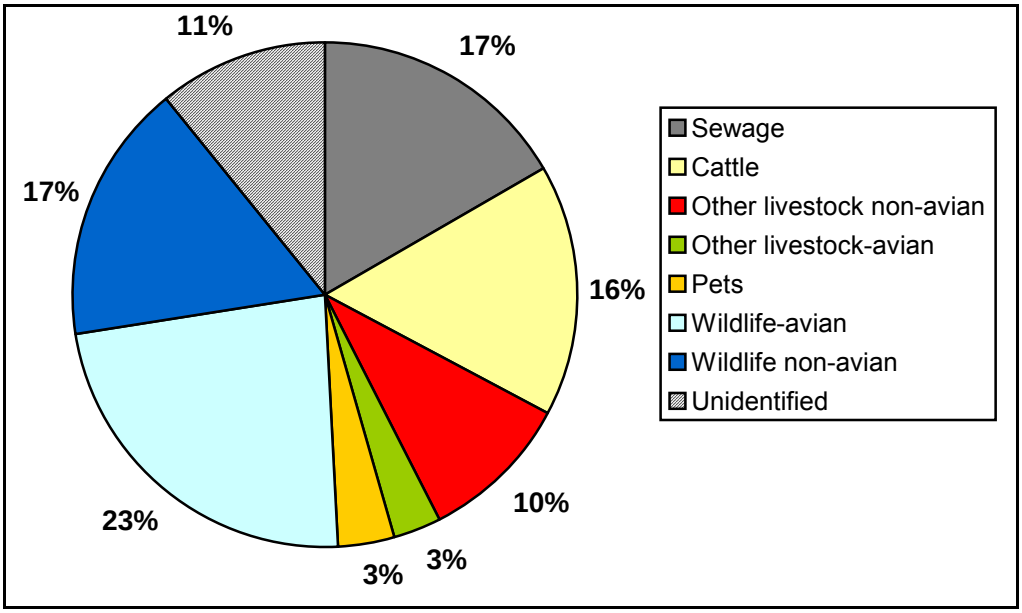


Figure ES-5 Source Contribution Estimates using Four-Method Composite Data for all Belton Lake and Leon River Sites and all Dates Combined (207 *E. coli* Isolates)

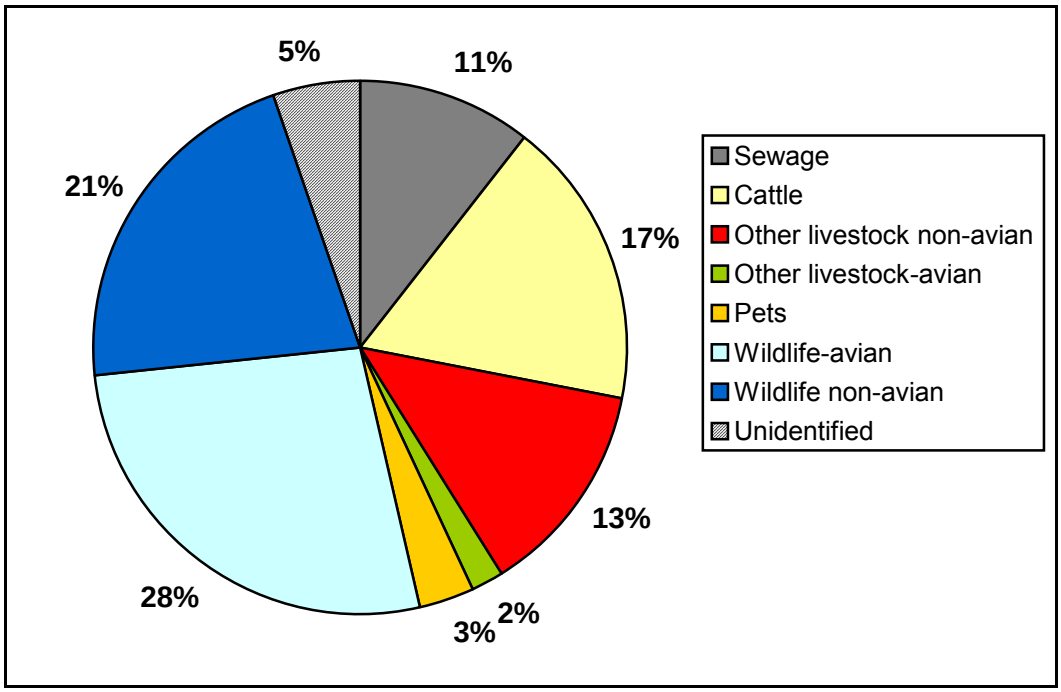
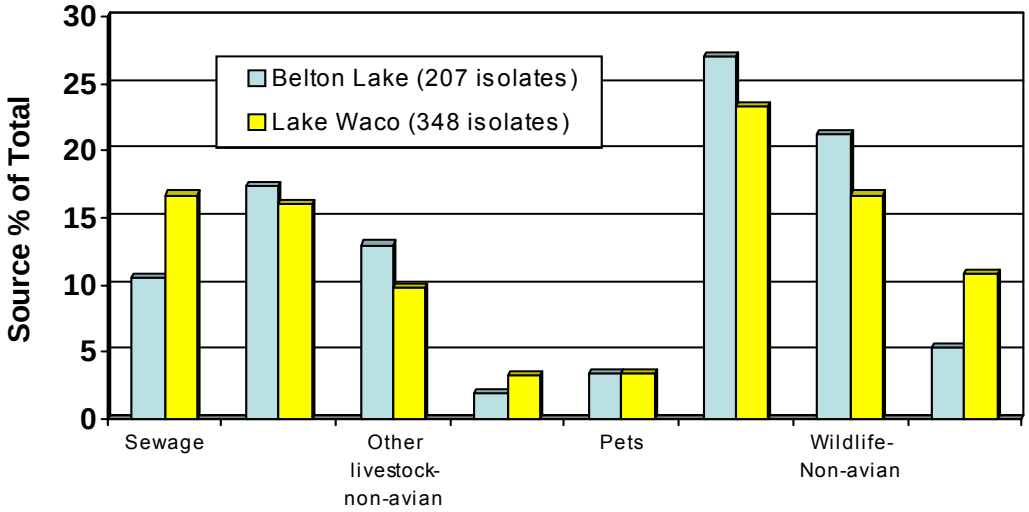


Figure ES-6 Comparison of Source Contribution Estimates for Belton Lake vs. Lake Waco using Four-Method Composite Data



Recommendations for Future BST Studies

Some of the key issues which affect data quality objectives (DQOs) and the sampling design of BST projects that must be given special attention are:

- budget,
- discriminatory capability desired of BST methods,
- size and diversity of known source library, and
- long-term maintenance of library isolates.

Library size and the representativeness of *E coli* strains in a known source library are two major considerations that need to be carefully assessed before embarking on any BST study (USEPA 2005). The level of discriminatory capability desired will also drive DQOs. BST projects are typically designed to differentiate:

1) human vs. all other sources,
2) species specific results (human vs. cows vs. horses vs. deer, *etc*.),
3) group comparisons (human vs. livestock vs. wildlife), and
4) specific individual hosts (cows from a certain farm vs. other farms vs. other livestock on farms vs. human etc. (USEPA 2005).

Key Recommendations

- Special attention should also be given to ensuring that future BST projects in Texas adhere to a TCEQ- or TSSWCB-approved QAPP.
- BST projects should adequately invest in conducting a thorough sanitary survey.
- When considering the use of library dependent methods, special attention should be given to the long-term maintenance of the library isolates for future use and application.
- The combination of ERIC-PCR and RiboPrinting appear to be the most suitable as well as accurate methods for future library-based BST studies.

It may not be feasible in terms of both cost and time considerations to perform all four BST methods used here for future BST studies. Through this project the ERIC-PCR RiboPrinting composite data set was found to be the closest two-method combination (90.7 percent similar) to the four-method composite data set. These results suggest that a combination of just ERIC-PCR and RiboPrinting may be most suitable for future library-based BST studies.

Conclusions from BST Results

The data now available from a BST study such as this validate the wide array of bacteria sources contributing loading to surface water. It should be recognized that data from BST studies such as this have limitations to the spatial and temporal questions that can be answered. This study has accomplished the project objectives and adhered to the data quality objectives set forth in the QAPP. The key conclusions derived from this BST project are:

- BST results indicate that wildlife (avian and non-avian) is the major contributor to fecal pollution for both Lake Waco and Belton Lake. Wildlife

was identified as the source of 40 percent and 49 percent of the *E. coli* isolated from Lake Waco and Belton Lake water samples, respectively.

- It had been previously speculated that livestock, in particular cattle, and other agricultural activities were responsible for the majority of fecal pollution in these watersheds. Contributions to fecal pollution by cattle as determined by BST are similar for both Lake Waco and Belton Lake watersheds, with 16 percent and 17 percent of the *E. coli* isolates from water identified as cattle source, respectively. Looking at the contributions from cattle at each of the 11 monitoring stations at both lakes, the percentage is less than or equal to 25 percent. Thus, these results demonstrate that cattle are not the major contributing source of bacteria to Lake Waco and Belton Lake. If the cattle, other livestock - avian, and non-avian livestock source classes are combined into a "livestock" class, livestock appear to be the second leading source of fecal pollution, with 29 percent and 32 percent of the water isolates from Lake Waco and Belton Lake, respectively, identified to this source.
- Unexpectedly, domestic sewage is identified as the third leading contributor to fecal pollution in the watersheds, with 17 percent of the Lake Waco and 11 percent of the Belton Lake *E. coli* isolates identified to this source. Of particular concern is the finding that the sampling site with the highest occurrence (27%) of *E. coli* isolates identified as sewage is at Lake Waco near the dam (Station 11942), which is also near the drinking water treatment plant intake.
- The BST results provide valuable information that will assist water resource managers in targeting future management strategies to address bacteria contributions from specific source categories in each watershed.

TABLE OF CONTENTS

EXECUTIVE SUMMARY ES-1
SECTION 1 INTRODUCTION..... 1-1
1.1 Project Objectives and Description..... 1-1
1.2 Introduction to Bacterial source tracking..... 1-5
1.3 Project Description..... 1-6
1.4 Project Team and Collaboration 1-7
SECTION 2 POTENTIAL SOURCES OF FECAL CONTAMINATION IN WATERSHEDS OF LAKE WACO AND BELTON LAKE..... 2-1
2.1 Water Body Characteristics 2-1
2.1.1 Lake Waco 2-1
2.1.2 Lake Waco Tributaries..... 2-2
2.1.3 Belton Lake..... 2-2
2.1.4 Belton Lake Tributaries 2-3
2.2 Watershed Characteristics..... 2-3
2.3 Watershed Land Use 2-4
2.4 Human and Pet populations, and Sewage Disposal 2-10
2.5 Watershed Livestock Populations..... 2-11
2.6 Watershed Wildlife Populations 2-12
2.7 Point sources of Potential Fecal Contamination 2-12
2.7.1 Permitted Wastewater Discharges 2-12
SECTION 3 AMBIENT WATER SAMPLING..... 3-1
3.1 Sampling Plan and Procedures..... 3-1
3.2 Isolation and ConfIrmation of *E. Coli* from Water..... 3-7
SECTION 4 *E. COLI* KNOWN SOURCE LIBRARY DEVELOPMENT..... 4-1
4.1 Collection, Isolation, and Confirmation of *E. coli* from Known Sources..... 4-1
4.2 Archival of Isolates 4-3
4.3 Screening of known Source *E. coli* Using ERIC-PCR for Library Construction..... 4-3
4.4 Exclusion of Identical (Clonal) *E. coli* Isolates from the Same Known Source Sample..... 4-5
4.5 Selection of Isolates for the Library..... 4-5
SECTION 5 DESCRIPTION OF BST METHODS 5-1
5.1 RiboTyping 5-2
5.2 Pulsed Field Gel Electrophoresis 5-2

5.3 Enterobacterial Repetitive Intergenic Consensus Polymerase Chain Reaction (ERIC-PCR) 5-3
5.4 Kirby-Bauer Antibiotic Resistance Analysis 5-3
SECTION 6 BST DATA PROCESSING AND ANALYSIS USING BIONUMERICS 6-1
6.1 Gel Electrophoresis Image Processing 6-1
6.2 Correlation Coefficients, Matching Criteria 6-1
6.3 Composite Data Sets 6-2
6.4 Discriminant Analysis 6-4
SECTION 7 EVALUATION OF BST TECHNIQUES 7-1
7.1 Precision and Accuracy of Techniques 7-1
7.2 Laboratory Controls 7-1
7.3 Blind QC Study 7-1
7.4 Library Evaluation 7-3
7.4.1 Jackknife Analyses of Each BST Method 7-3
7.4.2 Interpretation of Source Class Cross-Identification of *E. coli* Library Isolates 7-5
7.4.3 Library Quality Measures 7-5
7.5 Discriminant Analysis of KB-ARA Data 7-8
SECTION 8 OBSERVED *E. COLI* LEVELS IN LAKE WACO, BELTON LAKE, AND THEIR MAJOR TRIBUTARIES 8-1
SECTION 9 BST RESULTS 9-1
9.1 Source Identification of Water Isolates 9-1
9.2 Interpretation of Composite BST Results and Discussion 9-2
9.2.1 Aggregate Results Comparison between Lake Waco and Belton Lake 9-7
9.2.2 BST Classification of Sources for Individual Monitoring Stations 9-10
SECTION 10 STAKEHOLDER INVOLVEMENT 10-1
10.1 Workshops 10-1
10.2 Summary of Questions and Responses from Workshop 2 10-1
SECTION 11 RECOMMENDATIONS FOR FUTURE BST STUDIES 11-1
11.1 BST Project design 11-1
11.2 Correspondence Between Methods 11-2
SECTION 12 CONCLUSIONS FROM BST RESULTS 12-1
SECTION 13 REFERENCES 13-1

APPENDICES

Appendix A Ambient Water Quality Data
Appendix B Laboratory Protocol for Isolation and Confirmation of *Escherichia coli* from Water Samples
Appendix C Summary of Known Fecal Samples Collected from Lake Waco and Belton Lake Watersheds
Appendix D Laboratory Protocol for Isolation and Confirmation of *Escherichia coli* from Fecal Specimens
Appendix E Archival of *Escherichia coli* Isolates
Appendix F Laboratory Protocol for ERIC-PCR Fingerprinting of *Escherichia Coli*
Appendix G Selection of Isolates for Library Inclusion Based on ERIC-PCR Fingerprints
Appendix H Laboratory Protocol for Automated Ribotyping of *Escherichia Coli* using the DuPont Qualicon RiboPrinter
Appendix I Laboratory Protocol for Molecular Subtyping of *Escherichia Coli* by Pulse Field Gel Electrophoresis (PFGE)
Appendix J Laboratory Protocol for Antibiotic Resistance Analysis by the Kirby-Bauer Disk Diffusion Method
Appendix K BST Workshops Stakeholders List

LIST OF FIGURES

Figure ES-1 Discriminatory Capabilities of Different BST Methods 3

Figure ES-2 Geometric Mean of *E. coli* Concentrations - Lake Waco Sampling Sites, September 2003 through June 2004 5

Figure ES-3 Geometric Mean of *E. coli* Concentrations - Belton Lake Sampling Sites, September 2003 through June 2004 6

Figure ES-4 Source Contribution Estimates using Four-Method Composite Data for all Lake Waco and North Bosque River Sites and all Dates Combined (348 *E. coli* Isolates) 7

Figure ES-5 Source Contribution Estimates using Four-Method Composite Data for all Belton Lake and Leon River Sites and all Dates Combined (207 *E. coli* Isolates) 8

Figure ES-6 Comparison of Source Contribution Estimates for Belton Lake vs. Lake Waco using Four-Method Composite Data 8

Figure 1-1 Lake Waco and Belton Lake Watersheds in Central Texas 1-3

Figure 1-2 Project Organization Chart 1-8

Figure 2-1 Land Use/Land Cover for Lake Waco Watershed 2-5

Figure 2-2 Land Use/Land Cover for Belton Lake Watershed 2-7

Figure 3-1 Water Quality Monitoring Stations on Lake Waco and the North Bosque River 3-3

Figure 3-2 Water Quality Monitoring Stations on Belton Lake and Leon River 3-5

Figure 3-3 Isolation of *E. coli* from Water Samples 3-7

Figure 5-1 Discriminatory Capabilities of Different BST Methods 5-1

Figure 7-1 Blind Quality Control Study Results 7-3

Figure 7-2 Cross-Validation Jackknife Analysis of the Four-Method Composite Data Library 7-6

Figure 8-1 Geometric Mean of *E. coli* Concentrations - Lake Waco Sampling Sites, September 2003 through June 2004 8-3

Figure 8-2 Geometric Mean of *E. coli* Concentrations - Belton Lake Sampling Sites, September 2003 through June 2004 8-3

Figure 9-1 Identification of Water Isolates from the Lakes Waco and Belton Watersheds using the Individual Methods and Four-Method Composite (555 water isolates vs. 883 known source isolates) 9-2

Figure 9-2 Source Contribution Estimates using Four-Method Composite Data for all Sites and Dates Combined (555 *E. coli* Isolates) 9-6

Figure 9-3 Comparison of Source Contribution Estimates for Runoff Samples (November) from all Sites vs. all Other Dates Combined Using Four-Method Composite Data 9-6

Figure 9-4 Source Contribution Estimates using Four-Method Composite Data for all Lake Waco and North Bosque River Sites and all Dates Combined (348 *E. coli* Isolates) 9-7

Figure 9-5 Source Contribution Estimates using Four-Method Composite Data for all Belton Lake and Leon River Sites and all Dates Combined (207 *E. coli* Isolates) 9-8

Figure 9-6 Comparison of Source Contribution Estimates for Belton Lake vs. Lake Waco using Four-Method Composite Data 9-8

Figure 9-7 Comparison of Source Contribution Estimates by Month using Four-Method Composite Data for all Lake Waco / North Bosque River Sites Combined (348 *E. coli* Isolates) 9-9

Figure 9-8 Comparison of Source Contribution Estimates by Month using Four-Method Composite Data for all Belton Lake / Leon River Sites Combined (207 *E. coli* Isolates) 9-9

Figure 9-9 Source Contribution Estimates using Four-Method Composite Data for Belton Lake near Dam (11921) - all Dates Combined (27 *E. coli* Isolates) 9-10

Figure 9-10 Source Contribution Estimates using Four-Method Composite Data for Belton Lake Cowhouse Creek Arm (11922) - all Dates Combined (24 *E. coli* Isolates) 9-11

Figure 9-11 Source Contribution Estimates using Four-Method Composite Data for Belton Lake Owl Creek Arm (TBD4) - all Dates Combined (42 *E. coli* Isolates) 9-11

Figure 9-12 Source Contribution Estimates using Four-Method Composite Data for Belton Lake Leon River Arm near Headwater (11923) - all Dates Combined (44 *E. coli* Isolates) 9-12

Figure 9-13 Source Contribution Estimates using Four-Method Composite Data for Leon River at FM 1829 southeast of North Fort Hood (11925) - all Dates Combined (70 *E. coli* Isolates) 9-12

Figure 9-14 Source Contribution Estimates using Four-Method Composite Data for Lake Waco near Dam (11942) - all Dates Combined (67 *E. coli* Isolates) 9-13

Figure 9-15 Source Contribution Estimates using Four-Method Composite Data for Lake Waco Middle/South Bosque River Arm above SH 6 (11948) - all Dates Combined (45 *E. coli* Isolates) 9-13

Figure 9-16 Source Contribution Estimates using Four-Method Composite Data for Lake Waco Middle/South Bosque River Arm near Inlet of Middle/South Bosque Rivers (TBD3) - all Dates Combined (54 *E. coli* Isolates) 9-14

Figure 9-17 Source Contribution Estimates using Four-Method Composite Data for Lake Waco North Bosque River Arm (11945) - all Dates Combined (57 *E. coli* Isolates) 9-14

Figure 9-18 Source Contribution Estimates using Four-Method Composite Data for North Bosque River at SH 56 near Valley Mills (11953) - all Dates Combined (53 *E. coli* Isolates) 9-15

Figure 9-19 Source Contribution Estimates using Four-Method Composite Data for North Bosque River at FM 219 northeast of Clifton (11956) - all Dates Combined (72 *E. coli* Isolates) 9-15

Figure 11-1 Congruence of Individual BST Methods and Composite Data Sets..... 11-4

LIST OF TABLES

Table 2-1 Land Use/Land Cover Summary for Lake Waco Watershed 2-9

Table 2-2 Land Use/Land Cover Summary for Belton Lake Watershed..... 2-9

Table 2-3 Estimated Populations of Humans and Pets..... 2-10

Table 2-4 Estimated Livestock Populations by Watershed..... 2-11

Table 2-5 Permitted Wastewater Discharges to Lake Waco and its Tributaries..... 2-13

Table 2-6 Permitted Wastewater Discharges to Belton Lake and its Tributaries 2-14

Table 3-1 Sampling Sites 3-2

Table 3-2 Summary of Unknown Isolates from Water Quality Samples..... 3-8

Table 4-1 Summary of Known Source *E. coli* Isolates Used for Library Construction..... 4-4

Table 6-1 Comparison of Similarity Cutoffs for the Four-Method Composite Data Set and Effect on Library RCC and Percentage of Source and Water Isolates Left Unidentified..... 6-3

Table 7-1 Jackknife Analysis Rates of Correct Classification (%) for Individual and Composite BST Methods for the 883 Isolate BST Library 7-4

Table 7-2 Quality Measures for the Cross-Validated (Jackknifed) Composite Data Set Library and Seven-Way Split of Source Classes..... 7-7

Table 7-3 Quality Measures for the Cross-Validated (Jackknifed) Composite Data Set Library and Two-Way Split of Source Classes 7-7

Table 8-1 Observed *E. coli* Levels from Water Quality Samples Collected..... 8-2

Table 9-1 Example 95% Confidence Intervals for Various Contributing Source Magnitudes and Number of *E. coli* Isolates..... 9-3

Table 9-2 Sources of *E. coli* Isolates Based on Four-Method Composite Identification..... 9-4

Table 9-3 Sources of 74 *E. coli* Isolates Based on a Three-Method Composite Identification (No PFGE Data Available) 9-5

ACRONYMS AND ABBREVIATIONS

°C	degrees Celsius
ARA	Antibiotic resistance analysis
ARP	Antibiotic resistance profile
BHI	Brain heart infusion
BMP	Best management practice
BRA	Brazos River Authority
BST	Bacterial source tracking
CDC	Centers for Disease Control and Prevention
cfu	Colony forming unit (of bacteria)
DNA	Deoxyribonucleic acid
DQO	Data quality objective
E. coli	*Escherichia coli*
EP AREC	Texas A&M Agricultural Research and Extension Center at El Paso
ERIC-PCR	Enterobacterial repetitive intergenic consensus sequence polymerase chain reaction
FM	Farm to Market Road
ID	Identification
KB-ARA	Kirby-Bauer antibiotic resistance analysis
Magenta Gluc	5-bromo-6-chloro-3-indolyl-β-D-glucuronide
MGD	Million gallons per day
ml	Milliliter
MS4	Municipal Separate Storm Sewer System
MUG	4-methylumbelliferyl-β-D-glucuronide
NLCD	National Land Cover Data Set
NPDES	National Pollutant Discharge Elimination System
Parsons	Parsons Water and Infrastructure Inc.
PCR	Polymerase chain reaction
PFGE	Pulsed-field gel electrophoresis
QA	Quality assurance
QAPP	Quality assurance project plan
QC	Quality control
RCC	Rate of correct classification
RNA	ribonucleic acid
SH	state highway
SWMP	Stormwater Management Program
TAC	Texas administrative code
TAMU	Texas A&M University (College Station campus implied)
TAMU-CC	Texas A&M University – Corpus Christi
TBD	To be determined
TCEQ	Texas Commission on Environmental Quality
TFB	Texas Farm Bureau
TMDL	Total maximum daily load
TPDES	Texas Pollution Discharge Elimination System

ACRONYMS AND ABBREVIATIONS (*continued*)

TSB	Tryptone soy broth
TSSWCB	Texas State Soil and Water Conservation Board
USDA	United States Department of Agriculture
USEPA	United States Environmental Protection Agency

SECTION 1
INTRODUCTION

1.1 PROJECT OBJECTIVES AND DESCRIPTION

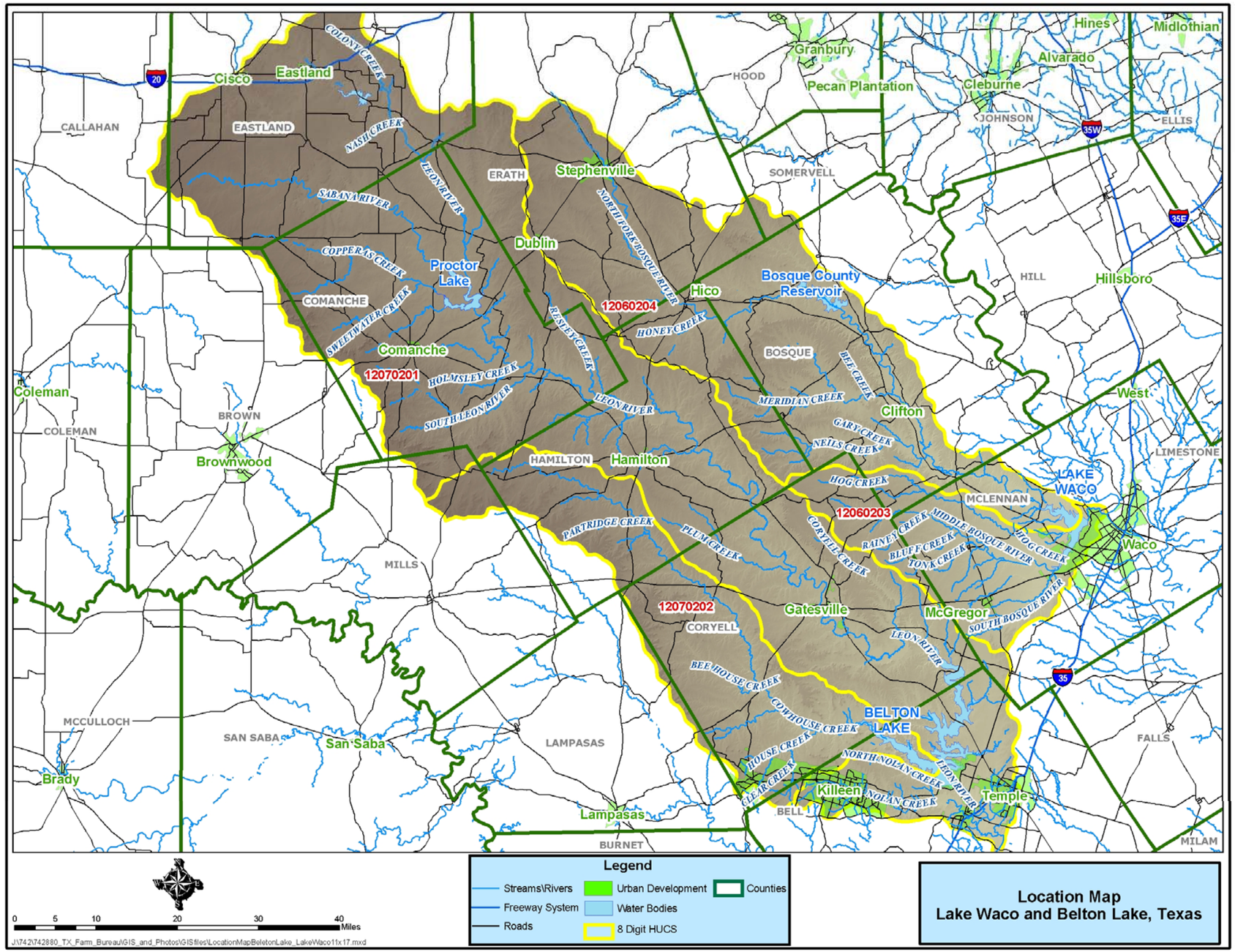
Protection of our water resources is one of the most significant environmental challenges of the new millennium. With over 260 surface water segments named to the Draft 2004 Texas §303(d) List because of nonsupport of contact recreation (TCEQ 2005a), quantifying the sources, fate, and transport of bacteria loading has become a priority to water resource managers and citizens around the state. Fecal contamination in surface water originates from a wide variety of point and nonpoint sources. Thus, a key need by water quality managers is identification and assessment of specific bacteria sources of microbial contamination in surface water. Proper evaluation of point and nonpoint sources is needed to identify appropriate control measures and best management practices (BMP) to control microbial pollution.

This assessment report describes the investigation and categorization of the sources of fecal contamination in Lake Waco, Belton Lake, and portions of major tributaries to those lakes. Figure 1-1 depicts the location and geographic extent of the Lake Waco and Belton Lake watersheds. While both Lake Waco and Belton Lake meet Texas Surface Water Quality Standards for contact recreation and public water supply, concerns have been expressed over potential fecal contamination from animal agriculture operations in their watersheds. This assessment project was initiated and coordinated by the Texas Farm Bureau (TFB), and funded in part by the Texas State Soil and Water Conservation Board (TSSWCB) through a Clean Water Act § 319(h) grant from the U.S. Environmental Protection Agency (USEPA). Other financial, technical, and in-kind support for the project was provided by the TFB, the City of Waco, and the Brazos River Authority (BRA). Scientific and technical collaborators on the project included Parsons Water and Infrastructure Inc. (Parsons), the Texas A&M University Agricultural Research Extension Center at El Paso (EP AREC), Texas A&M University-Corpus Christi (TAMU-CC), and Texas A&M University (TAMU).

The two primary objectives of the project were to use bacterial source tracking (BST) methods to (1) assess the relative contributions of fecal bacteria from livestock (particularly cattle), humans, wildlife, and other animals in Lake Waco and Belton Lake; and (2) develop libraries of known bacteria sources, genetic and biochemical, which can be used in determining the origin (animal or human) of fecal contamination in surface water. As this was one of the first studies of its kind in the state, a secondary objective was to evaluate and compare several analytical methods for BST to identify the optimal method, or combination of methods, for future application in Texas. Given these objectives it should be noted that throughout this report data results are discussed and presented in relation to two separate but inter-dependent themes: 1) data results that support conclusions associated with different BST analytical methods; and 2) data results that support conclusions specific to bacteria sources identified in each watershed.

THIS PAGE INTENTIONALLY LEFT BLANK

Figure 1-1 Lake Waco and Belton Lake Watersheds in Central Texas



THIS PAGE INTENTIONALLY LEFT BLANK

1.2 INTRODUCTION TO BACTERIAL SOURCE TRACKING

Fecal coliform bacteria are commonly used as an indicator of fecal pollution and the potential presence of other pathogenic microorganisms in water. Bacteria in ambient water may emanate from a wide variety of sources, including domestic sewage from wastewater collection systems, septic tanks, domestic pets, livestock, and wildlife. It has been established that the fecal coliform bacterium *Escherichia coli* (*E. coli*) is more closely associated with fecal pollution than other fecal coliform bacteria which may normally reside and multiply in the environment. *E. coli* is a common inhabitant of animal and human intestines; and recent studies show that isolates from humans and various host animals (*e.g.* cattle, chickens, and pigs) may differ genetically and phenotypically (*i.e.,* physical characteristics).

It is believed that different strains of *E. coli* have adapted to environmental conditions (temperature, pH, nutrient content, antibiotic exposure, *etc*.) in the intestines of specific warm-blooded animal hosts. These adaptations presumably allow certain *E. coli* strains to survive and compete favorably within the host's intestines, which gives rise to a level of host specificity among *E. coli*. Genetic and biochemical tests can identify specific strains of *E. coli*, which in the case of a host specific strain, may allow the original host animal to be identified, a process referred to as BST.

Molecular (genetic) tools appear to hold the greatest promise for BST, providing the most conclusive characterization and level of discrimination for isolates. Of the molecular tools available, ribosomal ribonucleic acid (RNA) genetic fingerprinting (ribotyping), enterobacterial repetitive intergenic consensus sequence polymerase chain reaction (ERIC-PCR) and pulsed-field gel electrophoresis (PFGE) are emerging as versatile and feasible BST techniques. Antibiotic resistance profiling, a phenotypic characterization method, also has the potential to identify the human or animal origin of isolates.

Each of these methods has its strengths and weaknesses. Ribotyping has a moderate ability to resolve different strains of the same species of bacteria. While an automated ribotyping system saves labor costs and requires little training, the initial investment and the consumable cost per isolate are expensive. PFGE has very high resolution and can discriminate between closely related strains. While this allows higher confidence in the matches made, fewer unknown isolates are typically identified compared to other BST techniques. ERIC-PCR has moderately high ability to resolve different strains of the same species of bacteria. Consumable costs for ERIC-PCR are inexpensive, but labor costs for sample processing and data analyses are moderate. Of the four methods applied in this study, antibiotic resistance analysis using the Kirby Bauer method has the lowest ability to discriminate closely related bacterial strains. It also has the lowest initial and per sample cost and takes the least time and training, but the statistical analysis of data can be complex and time-consuming. A disadvantage of all these techniques is that reference "libraries" or known sources of bacterial genetic fingerprints and antibiotic resistance profiles are needed to correctly identify the source of bacteria isolated from environmental water samples.

1.3 PROJECT DESCRIPTION

This assessment project began with the review and evaluation of existing data and information pertaining to bacterial contributions and sources to Lake Waco and Belton Lake. The primary task however, was the collection and analysis of new data, of known and specified quality, to differentiate and quantify the relative contributions of bovine livestock and other animal and human bacteria sources into Lake Waco and Belton Lake. The assessment and differentiation between bacteria sources were integrated with the development of the BST library generated by TAMU, EP AREC, and TAMU-CC. The tasks set out below were designed to provide sufficient documentation of the data and technical analyses conducted to aid the TFB in communicating the assessment results to watershed stakeholders, the TSSWCB, Texas Commission on Environmental Quality (TCEQ), and the USEPA, as necessary. The project was also structured to obtain support from stakeholders, TCEQ, TSSWCB, and USEPA Region 6 for the technical sufficiency of the project for application in other watersheds.

This project involved several steps:

- A sanitary survey of the watershed to identify potential contributing sources of fecal contamination based on existing literature and data as well as watershed reconnaissance.
- Coordination among all team members throughout the project and two stakeholder meetings, one at the beginning of the project and one to present results of the draft assessment report.
- Development of a sampling plan and quality assurance project plan (QAPP) to ensure that appropriate types and amount of data would be collected, and that the data would be reliable.
- Collection of fecal matter from known sources in the watershed to develop libraries of genetic signatures and antibiotic resistance profiles of *E. coli*.
- Collection and culturing of a representative set of *E. coli* isolates from Lake Waco, Belton Lake, and their major tributaries.
- Determination of the genetic signatures and antibiotic resistance profiles of these *E. coli* isolates from Lake Waco, Belton Lake, and their major tributaries.
- Application of statistical analysis to match the *E. coli* fingerprints from ambient water samples to those from known sources.
- Comparison and quantification of the accuracy and precision of four different BST methods and source determinations – antibiotic resistance analysis (ARA), enterobacterial repetitive intergenic consensus sequence polymerase chain reaction (ERIC-PCR), PFGE, and ribotyping. The BST techniques are described in more detail in Section 4 of this report.
- Estimation of the relative source contributions of *E. coli* in Lake Waco, Belton Lake, and their major tributaries, and the confidence of these estimates, based on the above measurements.

1.4 PROJECT TEAM AND COLLABORATION

Performance of this investigation involved collaboration of the TFB, TSSWCB, BRA, the City of Waco, Parsons, EP AREC, TAMU-CC, and TAMU. The following list identifies the organizations participating in the project with their specific roles and responsibilities. Figure 1-2 depicts the project organization chart.

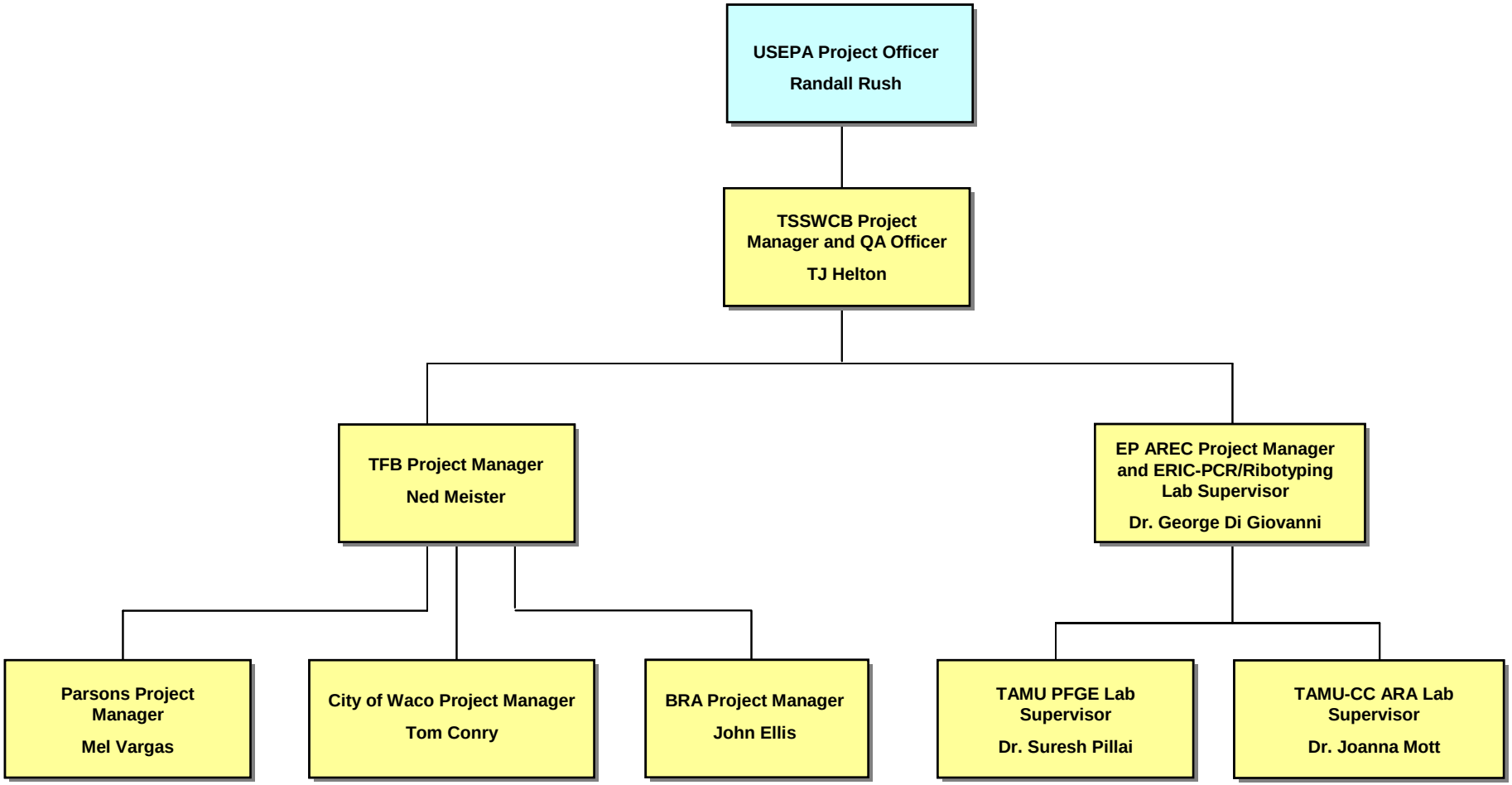
Team Members providing Project Management, Technical Direction, or Data Collection and Analysis:

- Texas Farm Bureau – Responsible for project coordination, review, and delivery of quarterly reports and the final project report to the TSSWCB.
- TSSWCB - Provided overall project review and approval. Obtained primary project funding with a Clean Water Act § 319(h) grant from the USEPA.
- City of Waco – Performed water sampling of Lake Waco. Responsible for the culturing and enumeration of *E. coli* in all ambient water samples and providing those cultures to the EP AREC laboratory.
- Brazos River Authority – Performed water sampling of Belton Lake.
- Parsons – Responsible for performing a sanitary survey of the watershed to identify potential fecal sources, collection of fecal samples from known sources for library development, collection of ambient water samples from the North Bosque River and Leon River, development of data quality objectives (DQO) and a QAPP, data analysis and interpretation (with EP AREC), and developing a final report for submittal to TFB and TSSWCB.

Team Members providing Laboratory Analysis and Data Analysis:

- Texas A&M Agricultural Research and Extension Center at El Paso – Responsible for isolation and purification of *E. coli* from fecal specimens and water samples, archival of *E. coli* cultures, ERIC-PCR screening and ribotyping of *E. coli* isolates, data analysis, and reporting tasks for the project, including (together with Parsons) development of data quality objectives (DQOs) and a QAPP. EP AREC was responsible (with the TFB and Parsons) for coordination, development, and delivery of quarterly reports as well as the final project report.
- Texas A&M University – Responsible for PFGE analysis of *E. coli* isolates provided by EP AREC. Responsible for PFGE data analysis and quality assurance (QA) procedures for that task.
- Texas A&M University – Corpus Christi – Responsible for ARA of *E. coli* isolates provided by EP AREC. Responsible for ARA data analysis and QA procedures for that task.

Figure 1-2 Project Organization Chart



SECTION 2
POTENTIAL SOURCES OF FECAL CONTAMINATION IN WATERSHEDS OF LAKE WACO AND BELTON LAKE

A critical step in any BST project is conducting a sanitary survey. A sanitary survey involves compiling existing information and conducting field reconnaissance regarding land use, population density, wastewater and stormwater infrastructure, agricultural practices and species census, and wildlife census to assist in identifying the potential sources of fecal contamination within the contributing watersheds of Lake Waco and Belton Lake. Conducting the sanitary survey provided information valuable in designing the approach for collecting ambient water samples and the known fecal samples for the library. This section summarizes information from the sanitary survey.

2.1 WATER BODY CHARACTERISTICS

2.1.1 Lake Waco

Lake Waco is a 7,178-acre reservoir in McLennan County, Texas, located entirely within the city limits of Waco. Lake Waco, segment 1225, serves as the water supply for over 110,000 central Texans (TCEQ 2005b). The dam at the northern end of the lake, completed in 1964 by the U.S. Army Corps of Engineers, impounds the North Bosque River to a normal pool elevation of 462 feet (USACE 2005). Other major tributaries include the Middle Bosque River, South Bosque River, and Hog Creek. The lake's primary purposes include flood control and water conservation.

In Texas Surface Water Quality Standards (Texas Administrative Code [TAC] §§307.1-307.10), the designated uses of Lake Waco include aquatic life use, contact recreation use, general use, fish consumption use, and public water supply use. All of these uses, except fish consumption which was not assessed, were determined to be supported by TCEQ in the Draft 2004 Texas Water Quality Inventory and §303(d) List (TCEQ 2005a). This assessment listed concerns over nutrient enrichment by nitrate and nitrite nitrogen, and excessive algal growth. Historically the TCEQ has relied on fecal coliform data collected from streams and lakes to assess support or nonsupport of the contact recreation use. Fecal coliform data was compiled and assessed in accordance with the TCEQ Guidance for Assessing Texas Surface and Finished Drinking Water Quality Data, 2004. The geometric mean of 53 fecal coliform measurements in the lake near the dam between March 1996 and February 2001 was 9.5 colony forming units (cfu)

Contact Recreation Use Assessment

In its *Guidance for Assessing Texas Surface and Finished Drinking Water Quality Data, 2004*, the TCEQ states "For routinely monitored bacteria data, the following long-term geometric averages have been established as criteria: fecal coliform, 200 cfu/100 ml; *E.coli*, 126 cfu/100 ml; and Enterococci, 35 cfu/100 ml. A fecal coliform criterion of 400 cfu/100 ml, an *E.coli* criterion of 394 cfu/100 ml, and an Enterococci criterion of 89 cfu/100 ml also apply to individual grab samples. The contact recreation use is not supported if the geometric average of the samples collected exceeds the mean criterion or if the criteria for individual samples are exceeded greater than 25 percent of the time... (TCEQ 2004)

per 100 milliliter (ml) of water, far less than the geometric mean water quality criterion of 200 cfu/100 ml. Two of the 53 samples near the dam exceeded the single sample water quality criterion of 400 cfu/100 ml. In the North Bosque River Arm of the lake, the geometric mean of 68 fecal coliform measurements during the same period was 11.17 cfu/100 ml. None of these 68 measurements exceeded the single sample criterion. In the Middle/South Bosque River Arm of the lake, the geometric mean of 67 fecal coliform measurements during the same period was 2 cfu/100 ml. Four of the 67 measurements exceeded the single sample water quality criterion of 400 cfu/100 ml.

2.1.2 Lake Waco Tributaries

The North Bosque River supports contact recreation use. The geometric mean of 30 fecal coliform samples collected from the portion of the North Bosque just upstream of Lake Waco was 88 cfu/100 ml, and five of the 30 samples, or 17 percent, exceeded the single sample criterion.

Hog Creek supports contact recreation use. The geometric mean of 106 fecal coliform samples was 115 cfu/100 ml, and 16 of the 106 samples, or 15 percent, exceeded the single sample criterion.

The Middle Bosque River supports contact recreation use. The geometric mean of 82 fecal coliform samples was 105 cfu/100 ml, and 13 of the 82 samples, or 16 percent, exceeded the single sample criterion.

The South Bosque River supports contact recreation use. The geometric mean of 101 fecal coliform samples was 145 cfu/100 ml, and 21 of the 101 samples, or 21 percent, exceeded the single sample criterion.

2.1.3 Belton Lake

Belton Lake is a 12,300-acre reservoir in Bell and Coryell Counties, roughly 3 miles north of the City of Belton and approximately 8 miles west of the City of Temple. Belton Lake, segment 1220, serves as the water supply for over 280,000 central Texans (TCEQ 2005b). The dam near Belton, completed in 1954 by the U.S. Army Corps of Engineers, impounds the Leon River to a normal pool elevation of 594 feet. In addition to the Leon River, other major tributaries include Cowhouse Creek and Owl Creek. In addition to providing flood control, numerous parks and recreational facilities and wildlife management areas rim the lake.

In Texas Surface Water Quality Standards (30 TAC §§307.1-7), designated uses of the lake include aquatic life use, contact recreation use, general use, fish consumption use, and public water supply use. All these uses except fish consumption, which was not assessed, were determined to be supported by TCEQ in the Draft 2004 Texas Water Quality Inventory and §303(d) List (TCEQ 2005a). Nutrient enrichment of the lake due to nitrate and nitrite nitrogen levels was identified as a concern.

The geometric mean of 66 fecal coliform measurements in the lower reaches of the lake between March 1996 and February 2001 was 3 cfu/100 ml of water. In the upper reaches of the lake, the geometric mean of 29 fecal coliform measurements during the same period was 4 cfu/100 ml. These concentrations were well below the contact

recreation water quality criterion of 200 cfu/100 ml. No single measurements exceeded the fecal coliform single sample criterion of 400 cfu/100 ml.

2.1.4 Belton Lake Tributaries

Among the tributaries to Belton Lake, some portions of the Leon River below Lake Proctor (segment 1221) did not support contact recreation because of elevated levels of fecal coliform bacteria, while other portions did support contact recreation. Contact recreation was supported in the portion of the Leon River just upstream of Lake Belton. The geometric mean of 19 fecal coliform measurements between March 1996 and February 2001 was 176 cfu/100 ml, which meets the water quality criterion of 200 cfu/100 ml. Four of the 19 samples, or 21 percent, exceeded the single sample criterion of 400 cfu/100 ml, which is below the threshold for nonsupport of contact recreation use. The portions of Leon River not supporting contact recreation were in northern Hamilton and southern Comanche Counties.

Cowhouse Creek fully supports contact recreation use. The geometric mean of 22 samples was 38 cfu/100 ml, and three of the 22 samples, or 14 percent, exceeded the single sample criterion.

2.2 WATERSHED CHARACTERISTICS

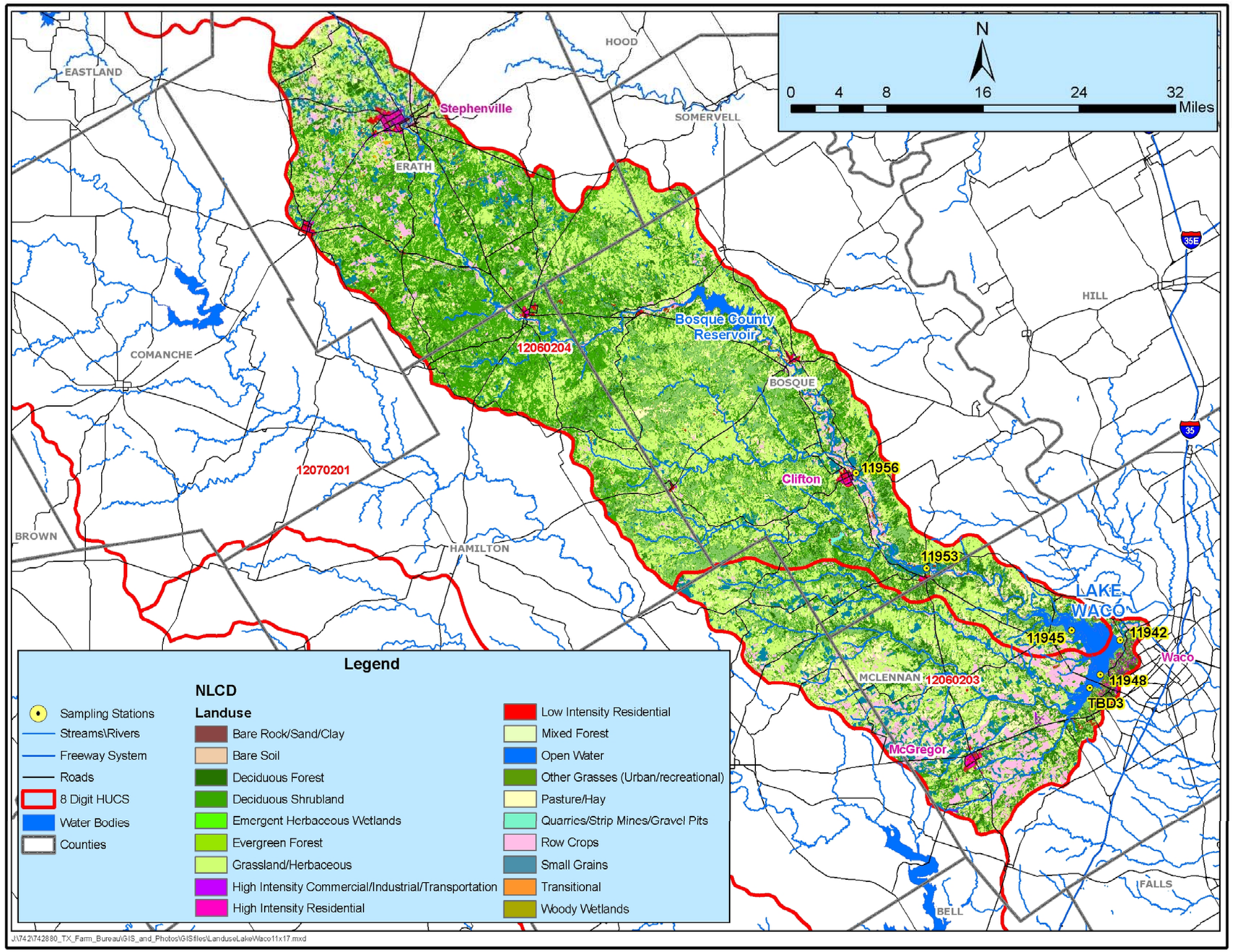
The Lake Waco watershed covers approximately 1,655 square miles. It extends westward from Waco to McGregor, then progresses northwest upstream along the North Bosque River. Approximately 74 percent of the watershed area is in the North Bosque River drainage, with the balance divided between the Middle Bosque River (12%), the South Bosque River (5%), Hog Creek (5%), and direct drainage to the lake from its east (1%), north (1%), and west (2%) shores. The watershed covers parts of six Texas Counties: Bosque, Coryell, Erath, Hamilton, McLennan, and Somervell. Cities and towns partially or wholly within the Lake Waco watershed include Stephenville, Dublin, Hico, Meridian, Iredell, Cranfills Gap, Clifton, Valley Mills, Crawford, McGregor, Woodway, and Waco.

The Belton Lake watershed covers approximately 3,568 square miles, slightly more than double the size of the Lake Waco watershed. It extends westward from the western edge of Temple to the northern edge of Killeen, then progresses northwest upstream along Cowhouse Creek and Leon River. Approximately 74 percent of the watershed area is in the Leon River drainage, with the balance divided between Cowhouse Creek drainage (20%), Owl and Flint Creek drainages (2%), and direct drainage to the lake from its immediate vicinity (4%). The watershed covers parts of 12 Texas Counties: Bell, Brown, Callahan, Comanche, Coryell, Eastland, Erath, Hamilton, Lampasas, McLennan, Mills, and Stephens. Cities and towns in the Belton Lake watershed include Ranger, Eastland, Dublin, Cisco, Carbon, Gorman, DeLeon, Rising Star, Comanche, Gustine, Hamilton, Gatesville, Evant, South Mountain, Morgan's Point Resort, McGregor, Copperas Cove, and small portions of Belton, Temple, and Killeen.

2.3 WATERSHED LAND USE

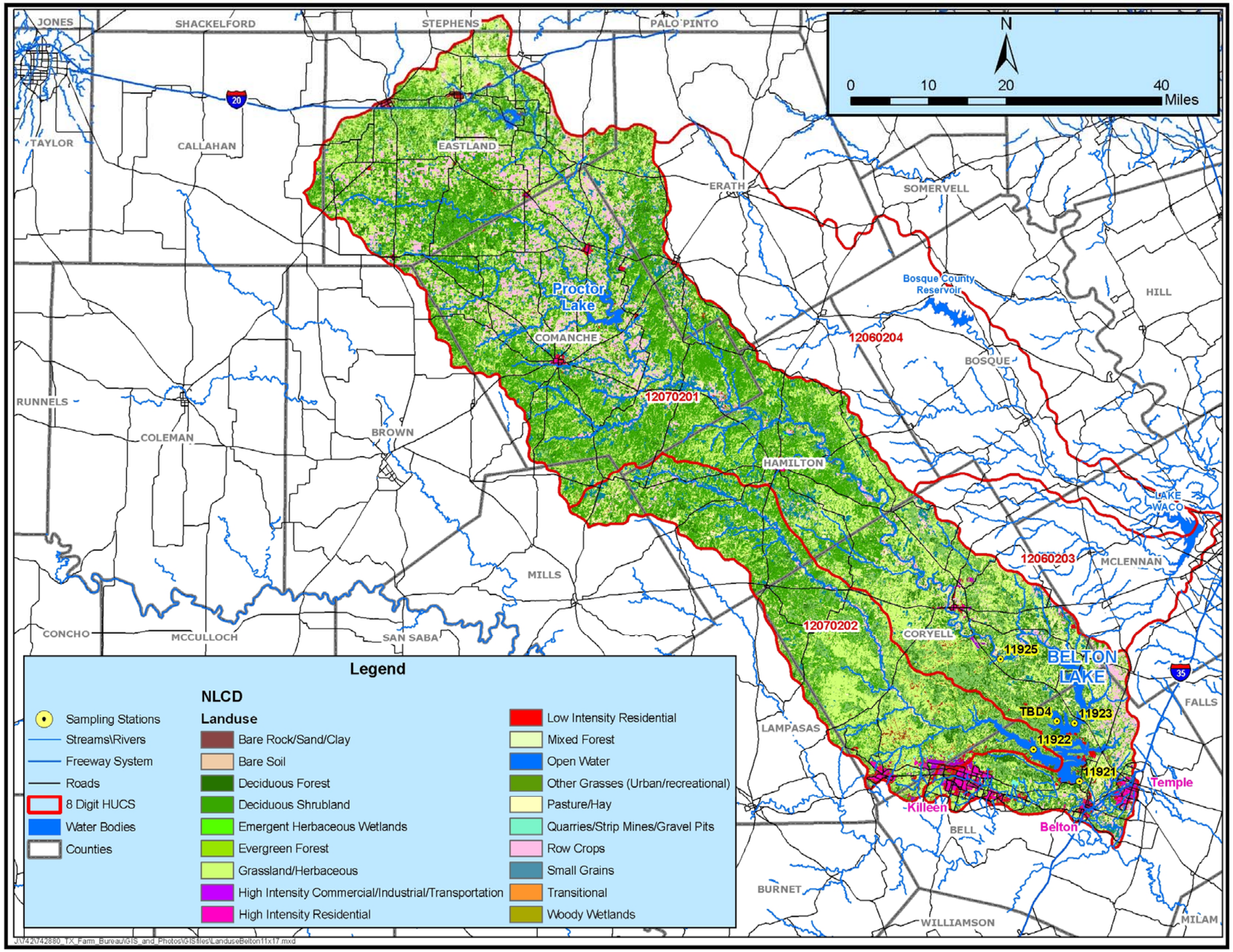
Figure 2-1 displays land use/land cover data for the Lake Waco watershed from the National Land Cover Data Set (NLCD) (U.S. Geological Survey [USGS] 1999), which was developed from satellite photographs taken in the early 1990s. The most dominant land use/land cover categories, detailed in Table 2-1, are grasslands, shrubland, forest, and planted/cultivated agricultural lands. Developed land comprises less than 2 percent of the watershed. Figure 2-2 displays the land use/land cover data for the Belton Lake watershed which, overall, has similar percentages to that of the Lake Waco watershed, as summarized in Table 2-2.

Figure 2-1 Land Use/Land Cover for Lake Waco Watershed



THIS PAGE INTENTIONALLY LEFT BLANK

Figure 2-2 Land Use/Land Cover for Belton Lake Watershed



THIS PAGE INTENTIONALLY LEFT BLANK

Table 2-1 Land Use/Land Cover Summary for Lake Waco Watershed

Land Use	Acres	% of Total
Grassland/Herbaceous	346,662	32.69
Deciduous Shrubland	267,375	25.21
Evergreen Forest	105,182	9.92
Deciduous Forest	101,375	9.56
Pasture/Hay	87,291	8.23
Row Crops	62,788	5.92
Small Grains	44,568	4.20
Open Water	21,876	2.06
Commercial/Industrial/Transportation	7,664	0.72
Bare Rock/Sand/Clay	6,575	0.62
Low Intensity Residential	3,365	0.32
High Intensity Residential	3,188	0.30
Other Grasses (Urban/recreational)	1,203	0.11
Quarries/Strip Mines/Gravel Pits	625	0.06
Emergent Herbaceous Wetlands	506	0.05
Planted/Cultivated (orchards, vineyards, groves)	142	0.01
Mixed Forest	17	0.00
Bare Soil	0	0.00
Woody Wetlands	0	0.00
	1,060,403	**100.00**

Source: National Land Cover Dataset (USGS 1999)

Table 2-2 Land Use/Land Cover Summary for Belton Lake Watershed

Land Use	Acres	% of Total
Deciduous Shrubland	758,719	31.74
Grassland/Herbaceous	743,511	31.10
Deciduous Forest	209,949	8.78
Evergreen Forest	189,156	7.91
Row Crops	162,529	6.80
Pasture/Hay	112,507	4.71
Small Grains	88,775	3.71
Open Water	46,126	1.93
Commercial/Industrial/Transportation	21,283	0.89
Low Intensity Residential	19,318	0.81
Bare Rock/Sand/Clay	19,057	0.80
High Intensity Residential	12,641	0.53
Other Grasses (Urban/recreational)	3,259	0.14
Quarries/Strip Mines/Gravel Pits	1,181	0.05
Emergent Herbaceous Wetlands	922	0.04
Mixed Forest	212	0.01
Bare Soil	842	0.04
Woody Wetlands	745	0.03
Planted/Cultivated (orchards, vineyards, groves)	23	0.00
	2,390,756	**100.00**

Source: National Land Cover Dataset (USGS 1999)

2.4 HUMAN AND PET POPULATIONS, AND SEWAGE DISPOSAL

The total human population of the Lake Waco watershed was estimated from the 2000 federal census (U.S. Census Bureau 2003) to be approximately 78,000 with an overall average population density of 47 persons per square mile. In the Belton Lake watershed, the population was estimated to be approximately 113,000, with an average population density of 32 persons per square mile.

Based on the number of households in the watersheds, and national averages of pet ownership from the American Veterinary Medical Association (AVMA 2004), approximately 20,000 domestic dogs and 23,000 cats reside in the Lake Waco watershed, and 25,000 dogs and 29,000 cats reside in the Belton Lake watershed. Pet fecal waste can represent a major source of fecal contamination to water bodies, particularly during rainfall events when runoff washes fecal material into streams. Table 2-3 summarizes the population estimates of humans and pets by watershed.

Table 2-3 Estimated Populations of Humans and Pets

Category		Lake Waco Watershed	Belton Lake Watershed
Human	Humans#	78,000	113,000
Pets	Cats*	23,000	29,000
Pets	Dogs*	20,000	25,000

 # Based on 2000 U.S. Census

**Based on 2000 U.S. Census and 2002 U.S. Pet Ownership & Demographics Sourcebook, American Veterinary Medical Association*

Sewer line breaks, bypasses, and overflows can contribute substantial quantities of fecal contamination to rivers and streams. Residents of approximately 71 percent of the housing units in the Lake Waco watershed, and 62 percent of the units in the Belton Lake watershed, reported in the 1990 federal census that their sewage was disposed via public sanitary sewer to a wastewater treatment facility[1]. However, approximately 8,600 homes in the Lake Waco watershed, and 14,700 homes in the Belton Lake watershed reported in this census that they utilized septic tanks or other unspecified non-sewer means for sewage disposal. Malfunctioning septic tanks can represent a significant source of fecal contamination to surface and ground waters, especially in areas near lakes and streams, and areas with shallow groundwater tables or impermeable clay soil. The rate of septic tank malfunction was estimated at 12 percent in the counties comprising the lower portions of the Lake Waco and Belton Lake watersheds, and 8 percent in the counties in the upper portions of the watersheds, based on a survey of officials with the responsibility for permitting septic tanks (Reed, Stowe, and Yanke 2001). The two primary factors listed as causes of malfunction were inappropriate soil type and age of the systems.

[1] This question was not asked in the 2000 federal census. The percentage of households connected to a public sanitary sewer may have increased somewhat since 1990 due to new housing outside sewered areas. For example, since 2001, the City of Waco reported extending sanitary sewers to serve 400 homes formerly utilizing septic tanks near Lake Waco.

Septic tanks are generally inappropriate for dense developments. New septic tanks are not allowed on lot sizes smaller than one-half acre (30 TAC §285.4), and problems may occur at lower densities. With the low spatial resolution of available data, it was not possible to identify particular geographic areas where the density of septic tanks exceeds recommended limits.

2.5 WATERSHED LIVESTOCK POPULATIONS

Table 2-4 summarizes livestock populations estimated from the 2002 Census of Agriculture, conducted by the National Agricultural Statistics Service of the U.S. Department of Agriculture (USDA) (USDA 2002), or from more recent (January 1, 2005) estimates of the Texas Agricultural Statistics Service, when available. These data are available only at the county level, and since the watersheds generally include only a portion of the counties, these estimates were calculated from the sum of the county populations multiplied by the fraction of the county within the watershed. Thus, these should be considered rough approximations.

Cattle were the most abundant species in both watersheds, and beef cattle and cows outnumbered dairy cows by a wide margin. Dairy cows were most abundant in Erath and Comanche Counties, relatively far upstream from Lake Waco and Belton Lake. Goats and sheep were also abundant in both watersheds, particularly in Mills County. Chickens were abundant in the Lake Waco watershed.

Table 2-4 Estimated Livestock Populations by Watershed

Category		Lake Waco Watershed	Belton Lake Watershed
Livestock	Cattle & Calves-All†	149,000	313,000
Livestock	Beef cows†	52,000	131,000
Livestock	Dairy cows	37,000[a]	57,000[a]
Livestock	Horses, mules, burros, & donkeys‡	3,100	4,400
Livestock	Hogs & Pigs‡	1,200	9,500
Livestock	Goats†	18,000	49,000
Livestock	Sheep†	7,600	30,000
Livestock	Rabbits‡	150	130
Livestock	Chickens‡	77,000	5,100
Livestock	Other poultry‡	1,600	1,000

† *As of January 1, 2005 Texas Agricultural Statistics Service*

‡ *2002 Census of Agriculture, USDA*

[a] *Based on TCEQ inspection reports of permitted facilities for fiscal year 2004-2005, with estimates for smaller facilities (max 200 head each) under the jurisdiction of the TSSWCB. Source: Texas Institute for Applied Environmental Research, Tarleton State University*

Livestock manure applied to land as fertilizer can serve as a source of fecal contamination to surface water if it is incompletely composted, not adequately incorporated into the soil, or if heavy rains fall on freshly applied manure. Based on county level reports from the 2002 Census of Agriculture, it is estimated that manure was applied to approximately 10,000 acres, or approximately 0.9% of the land area of the Lake Waco watershed, and to approximately 17,000 acres, or approximately 0.7% of the area of the Belton Lake watershed. The Implementation Plan for Soluble Reactive Phosphorus in the North Bosque River Watershed (TCEQ 2002) established a target for composting and exporting 50 percent of the dairy manure from the watershed. Between 2001 and August 2004, 495,368 tons of dairy manure were collected and composted by the TCEQ/TSSWCB Composted Manure Incentive Program, and 156,141 tons were exported from the watershed (TCEQ 2005c). While the goal of the Composted Manure Incentive Program is to reduce loadings of phosphorus in the North Bosque River, it should also reduce fecal contamination.

2.6 WATERSHED WILDLIFE POPULATIONS

Although wildlife can clearly contribute fecal contamination to surface water, it is difficult to develop population estimates to gauge the potential impact. White-tailed deer are very common throughout the region. Waterfowl are also seasonally abundant – the Texas Parks and Wildlife Department (TPWD) estimates that the mid-winter duck population of Texas averaged 3.6 million from 1997 to 2005, with perhaps 15 percent of those observed in central Texas (TPWD 2005). Fecal contamination from waterfowl is expected to be enhanced because much of their feces are deposited directly into or adjacent to water bodies. It is a general expectation that wildlife species that tend to occur near streams will cause more fecal contamination of water than their more upland counterparts. For example, fecal contamination from a raccoon might be expected to exceed that from an armadillo.

2.7 POINT SOURCES OF POTENTIAL FECAL CONTAMINATION

2.7.1 Permitted Wastewater Discharges

Under the Texas Pollution Discharge Elimination System (TPDES), 15 active facilities hold permits to discharge wastewater into Lake Waco and its tributaries (Table 2-5), and 28 facilities hold permits to discharge wastewater into Belton Lake and its tributaries (Table 2-6). The current average total wastewater discharges to Lake Waco tributaries is approximately 3.4 million gallons per day (MGD). Current average wastewater discharges to Belton Lake and its tributaries are estimated to total 6.4 MGD. Most of the wastewater permits do not include specific limits and monitoring requirements for fecal coliform concentrations in their effluents, but most do require monitoring of disinfection of wastewaters. Several additional facilities and agricultural operations hold irrigation/ agricultural (no discharge) permits.

Table 2-5 Permitted Wastewater Discharges to Lake Waco and its Tributaries

Permit Identification (ID)	Facility	Dates Monitored	Reported Average Flow (MGD)	Permitted Average Flow (MGD)	Receiving Water
WQ0010043	City of Clifton	12/1/2003 – 7/31/2005	0.35	0.65	North Bosque River
WQ0010113	City of Meridian	7/1/2004 – 7/31/2005	0.31	0.45	North Bosque River
WQ0010188	City of Hico	8/1/2004 – 7/31/2004	0.41	0.20	North Bosque River
WQ0010219	City of McGregor	11/30/2001 – 6/30/2005	0.70	0.99	South Bosque River
WQ0010290	City of Stephenville	4/1/2004 - 7/31/2005	1.5	3.5	Upper North Bosque River
WQ0010307	City of Valley Mills	12/1/2004 – 6/30/2005	0.055	0.36	North Bosque River
WQ0010656	City of Crawford	2/1/2004 – 7/31/2005	0.014	0.09	Middle Bosque River
WQ0011565	City of Iredell	9/1/2004 – 7/31/2005	0.041	0.049	North Bosque River
WQ0013436	City of Walnut Springs	NR	NR	0.065	North Bosque River
WQ0013966	Stephenville Mobile Home Park	NR	NR	0.007	Upper North Bosque River
WQ0013971	City of Waco Mt. Carmel Plant	5/1/2004 – 7/31/2005	0.04	0.24	Lake Waco
WQ0014169	City of Cranfills Gap	9/1/2004 – 6/30/2005	0.027	0.04	North Bosque River
WQ0003041	Chemical Lime	NR	NR	NR	North Bosque River
WQ0003074	Schreiber Foods/Dairy Farmers of America	NR	NR	0.20	Upper North Bosque River
WQ0004573	Ronald Eugene Schaefer	NR	NR	NR	Upper North Bosque River
Total			**3.4**	**6.8**	

Notes:
NR = Not reported.
MGD = millions of gallons per day.

Table 2-6 Permitted Wastewater Discharges to Belton Lake and its Tributaries

Permit ID	Facility	Dates Monitored	Reported Average Flow (MGD)	Permitted Average Flow (MGD)	Receiving Water
WQ0010045-4	City of Copperas Cove Northeast Plant	6/1/2005 – 7/31/2005	0.65	2.5	Cowhouse Creek
WQ0010045-5	City of Copperas Cove Northwest Plant	12/1/2004 – 8/31/2005	1.87	4.0	Cowhouse Creek
WQ0010078	City of De Leon	11/1/2004 – 2/28/2005	0.160	0.295	Leon River above Lake Proctor
WQ0010091	City of Gorman	6/1/2004 – 7/31/2005	0.04	0.12	Leon River above Lake Proctor
WQ0010176-1 WQ0010176-4	City of Gatesville Leon WWTP	9/1/2002 – 7/31/2005	0.54	1.0	Leon River below Lake Proctor
WQ0010176-2	City of Gatesville WWTP # 2	11/1/2004 – 8/31/2005	1.33	2.2	Leon River below Lake Proctor
WQ0010176-3	City of Gatesville WWTP	7/1/2004 – 7/31/2005	0.24	0.25	Leon River below Lake Proctor
WQ0010225	City of Moody	6/1/2004 – 7/31/2005	0.12	0.20	Belton Lake
WQ0010351-1	Bell Co. WCID #1		NR	NR	Belton Lake
WQ0010405	City of Dublin	12/1/2004 – 8/31/2005	0.25	0.45	Leon River below Lake Proctor
WQ0010492-2	City of Hamilton	11/1/2003 – 7/31/2005	0.34	0.44	Leon River below Lake Proctor
WQ0010637	City of Eastland	8/1/2004 – 7/31/2005	0.37	0.9	Leon River above Lake Proctor
WQ0010719 WQ0014445	City of Comanche	2/1/2004 – 7/31/2005	0.289	0.595	Leon River below Lake Proctor
WQ0010841	City of Gustine	11/1/2004 – 5/31/2005	0.020	0.082	Leon River below Lake Proctor
WQ0010888	NJB & Sons Greenbrier Golf Club	9/1/2004 – 6/30/2005	0.001	0.005	Belton Lake
WQ0010914	City of Oglesby	4/1/2004 – 7/31/2005	0.032	0.05	Leon River below Lake Proctor
WQ0010918-2	City of Morgan's Point Resort	NR	NR	NR	Belton Lake
WQ0011011	City of Evant	NR	NR	0.06	Cowhouse Creek
WQ0002233 WQ0002230	U.S. Army Fort Hood	NR	NR	NR	Cowhouse Creek
WQ0012096	U.S. Army North Fort Hood	NR	NR	0.25	Leon River below Lake Proctor

Table 2-6 Permitted Wastewater Discharges to Belton Lake and its Tributaries (*continued*)

Permit ID	Facility	Dates Monitored	Reported Average Flow (MGD)	Permitted Average Flow (MGD)	Receiving Water
WQ0013726	Eastland County WSD	NR	NR	0.1	Leon River above Lake Proctor
WQ0013965 WQ0014515	City of Rising Star	NR	NR	0.14	Leon River above Lake Proctor
WQ0014206	Upper Leon River WMD	2/1/2005 – 7/31/2005	0.096	0.249	Leon River below Lake Proctor
WQ0011764	Upper Leon River WMD	NR	NR	0.06	Leon River above Lake Proctor
WQ0014481	Sobrante Management	NR	NR	NR	Belton Lake
WQ0002335	U.S. Navy Weapons Plant	NR	NR	0.20	Leon River below Lake Proctor
WQ0004464	City of Gatesville	NR	NR	NR	Leon River below Lake Proctor
WQ0004749	Comanche Pottery	NR	NR	0.00035	Leon River above Lake Proctor
Total			**6.4**	**14.1**	

Notes:
NR = Not reported.
MGD = millions of gallons per day.

The National Pollution Discharge Elimination System (NPDES) Phase II rule, promulgated in 1999, requires municipalities in urban areas to obtain permits for their stormwater systems. These permits, known as Municipal Separate Storm Sewer System (MS4) permits, require cities to reduce discharges of pollutants in stormwater to the "maximum extent practicable" by developing and implementing a Stormwater Management Program (SWMP). The SWMPs require specification of BMPs for six minimum control measures:

- Public education and outreach;
- Public participation/involvement;
- Illicit discharge detection and elimination;
- Construction site runoff control;
- Post-construction runoff control; and
- Pollution prevention/good housekeeping.

The City of Waco and Texas Department of Transportation jointly hold an MS4 permit for the Waco urban area. Of the municipalities in the Belton Lake watershed, only Temple is designated as an urban area that must comply with the requirements of the NPDES Phase II program.

THIS PAGE INTENTIONALLY LEFT BLANK

SECTION 3
AMBIENT WATER SAMPLING

3.1 SAMPLING PLAN AND PROCEDURES

A primary objective of this BST project was identification and quantification of the relative contributing sources of *E. coli* to Lake Waco and Belton Lake. The collection of all water quality samples was governed by approved sampling methods summarized in a TSSWCB- and USEPA-approved QAPP. Four stations in Lake Waco, two stations on the North Bosque River upstream of Lake Waco, four stations in Belton Lake, and one station on the Leon River upstream of Belton Lake were sampled for *E. coli* on 10 dates. The location descriptions of water quality monitoring stations are listed in Table 3-1 and displayed in Figures 3-1 and 3-2. To obtain results representative of the normal occurrence of wet and dry conditions and seasonal variation, the ambient water sampling was performed on a routine monthly schedule from September 2003 through June 2004, to capture dry and runoff-influenced events at their natural frequency. It was not a goal of the project to quantify sources under wet or dry conditions separately, due to budget limitations.

While the sampling approach was intended to include some samples under rainfall runoff conditions, it should be recognized that it is difficult to gauge the extent of impact of rainfall runoff in reservoirs when there is often no measurable change in flow. The occurrence of rainfall across each watershed and the magnitude and intensity of rain required to generate runoff is quite variable. Thus it was difficult to define a runoff-influenced sampling event based on rainfall data alone. For this project a runoff-influenced sampling event is based on elevated flow in the Leon River at Gatesville and the North Bosque River at Valley Mills, and as such only one of the 10 sampling events (November) was strongly influenced by runoff on the day samples were collected. During the March sampling event, the rivers exhibited a minor increase in flow in response to rainfall, and elevated suspended solids and *E. coli* concentrations imply the likely influence of runoff, but it is not clear that a strong runoff influence was present. Finally, a major rainfall event occurred two days before the January sampling event, and river flow had returned to near pre-storm levels during sampling. Water quality in the rivers also appeared to have returned to pre-runoff conditions on this sampling date, but in the lakes elevated suspended solids and *E. coli* concentrations implied that some influence of runoff persisted there.

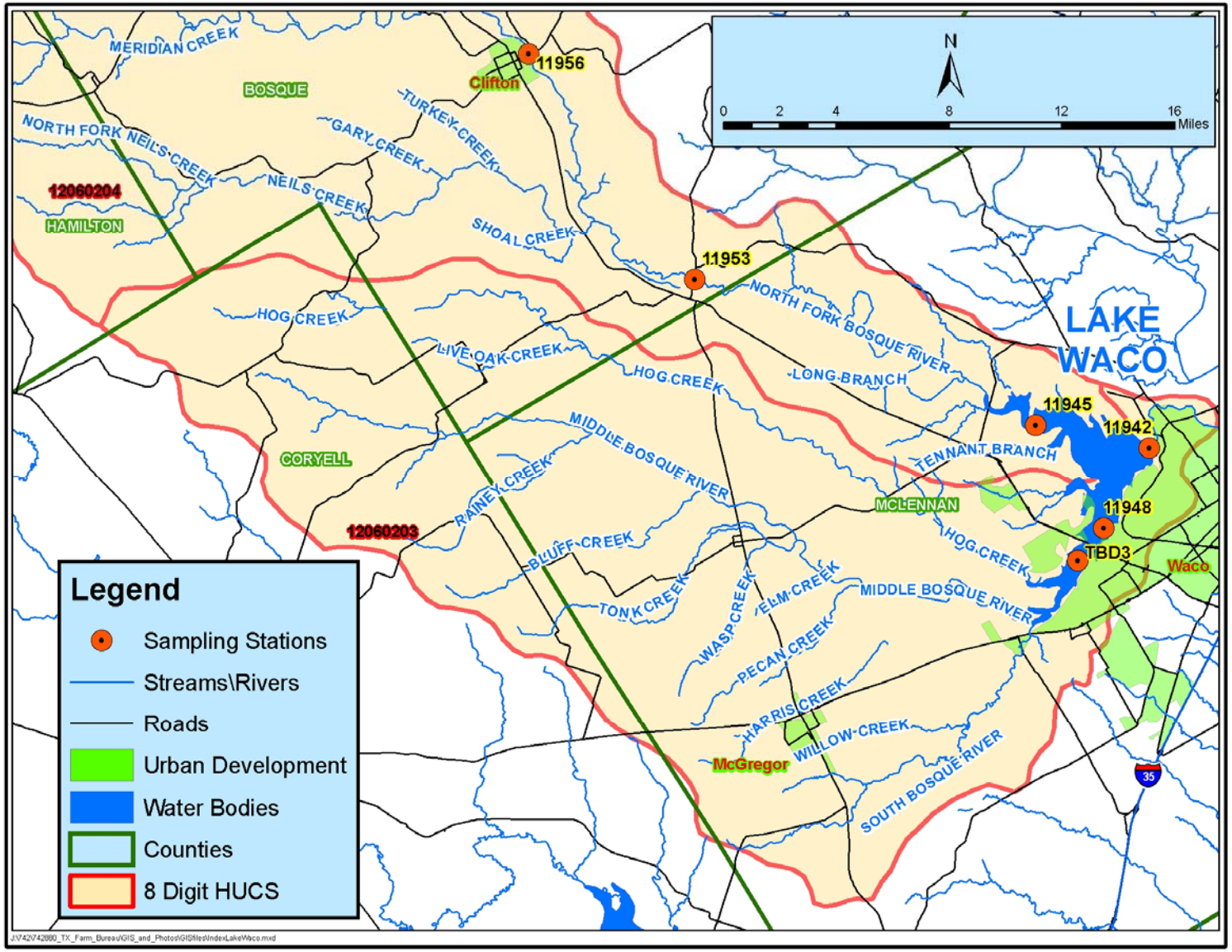
Because *E. coli* populations have been found to vary on fine spatial and temporal scales, five independent water samples were collected at each station and event, 1-2 minutes and 3-10 feet apart, to increase the representativeness of natural conditions by the sampling. Typically, this was done by sampling five points evenly spaced around the perimeter of a boat. At the two stations near the dams (11942 in Lake Waco and 11921 in Belton Lake) and in drinking water intakes of Lake Waco and Belton Lake, the samples were collected in duplicate for a total of 10 samples per event.

Table 3-1 Sampling Sites

Station ID	Station Description	Sampling Events, Frequency
11942	Lake Waco near dam	10, monthly
11945	Lake Waco North Bosque Arm	10, monthly
11948	Lake Waco Middle/South Bosque Arm above state highway (SH) 6	10, monthly
TBD3	Lake Waco Middle/South Bosque Arm near inlet of Middle/South Bosque River	10, monthly
11953	North Bosque River at SH 56 near Valley Mills	10, monthly
11956	North Bosque River at farm to market (FM) 219 northeast of Clifton	10, monthly
11921	Belton Lake near dam	10, monthly
11922	Belton Lake Cowhouse Creek Arm	10, monthly
11923	Belton Lake Leon River Arm near headwater	10, monthly
TBD4	Belton Lake Owl Creek Arm	10, monthly
11925	Leon River at FM 1829 southeast of North Fort Hood	10, monthly

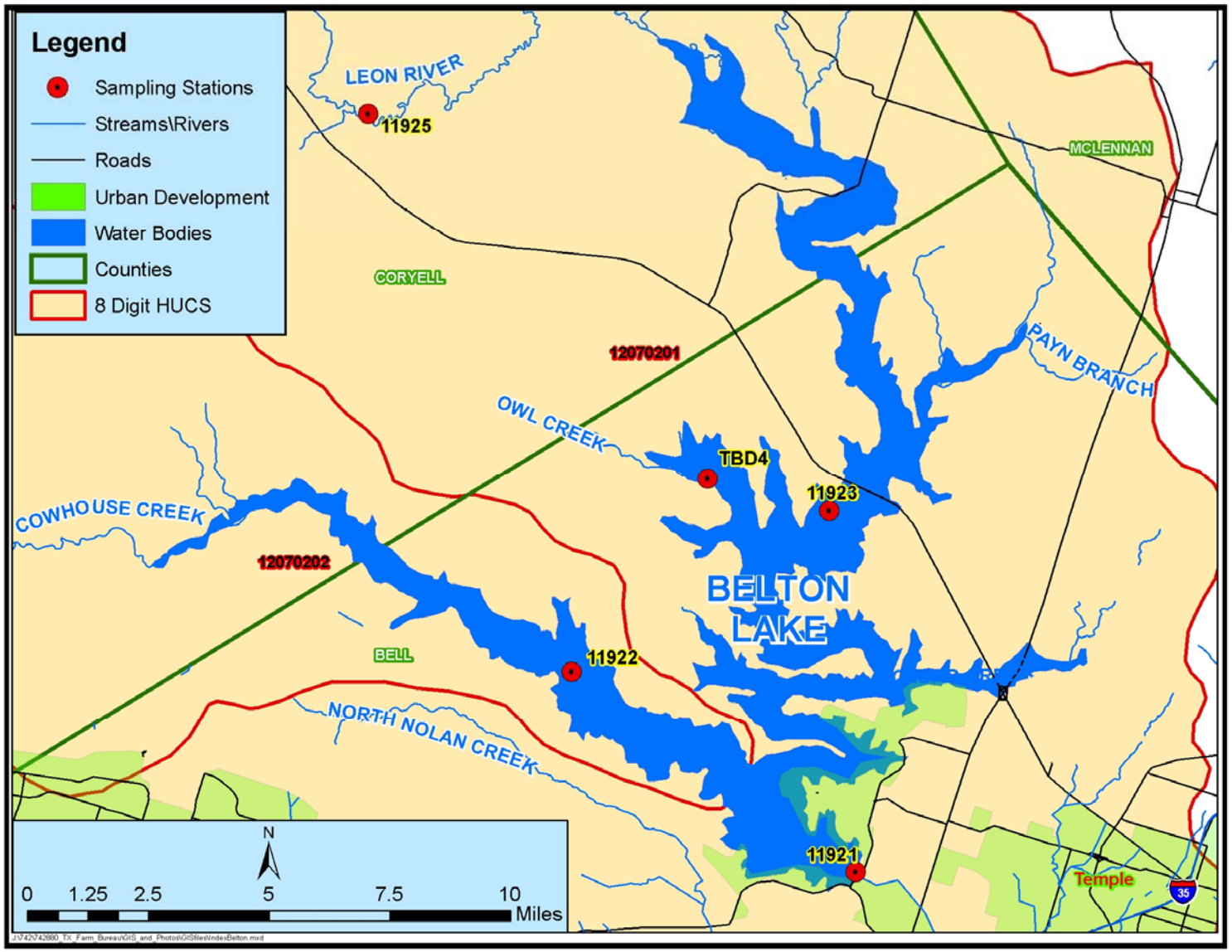
Water samples were collected directly from the lake or stream (approximately 1 foot below the surface) into sterile wide-mouthed polypropylene bottles supplied by the City of Waco laboratory. Care was exercised to avoid the surface microlayer of water, which can be enriched with bacteria and not representative of the water column. Field staff wore clean, disposable, powder-free gloves while collecting all samples. Blank samples were collected at a rate of one per 10 ambient samples to verify that no contamination took place. Upon collection, all water samples were placed in an iced container and transported to the City of Waco laboratory for analysis within 6 hours. A table summarizing the sampling results for each water quality monitoring station is provided in Appendix A.

Figure 3-1 Water Quality Monitoring Stations on Lake Waco and the North Bosque River



THIS PAGE INTENTIONALLY LEFT BLANK

Figure 3-2 Water Quality Monitoring Stations on Belton Lake and Leon River

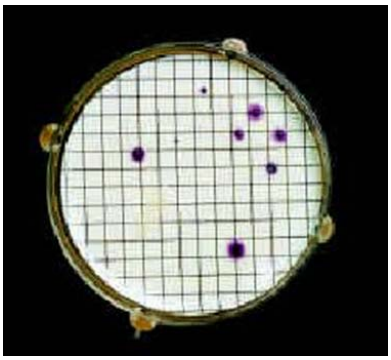


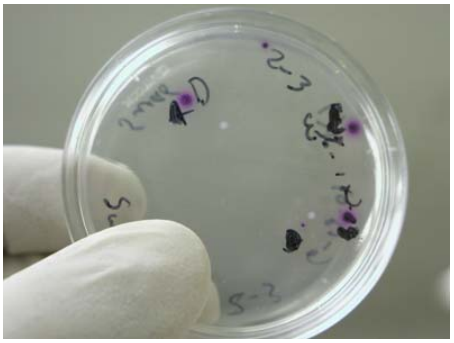
THIS PAGE INTENTIONALLY LEFT BLANK

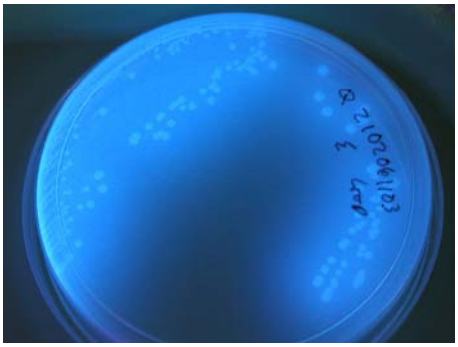
3.2 ISOLATION AND CONFIRMATION OF *E. COLI* FROM WATER

Following collection, water samples were delivered to the City of Waco Laboratory for *E. coli* culturing and enumeration via the membrane filter modified mTEC method. A 6-hour holding time for sample delivery to the laboratory and initiation of analysis was maintained. Following incubation and enumeration using USEPA Method 1603, the modified mTEC plates with *E. coli* colonies were shipped on ice overnight to the BST laboratory at EP AREC for *E. coli* isolation, confirmation, and archival. Figure 3-3 displays examples of the filters and plates used for isolation of *E. coli* colonies. *E. coli* colonies from the modified mTEC plates were picked and streaked for purity on nutrient agar with MUG (NA-MUG) to confirm glucuronidase activity and culture purity.

Figure 3-3 Isolation of *E. coli* from Water Samples







A sample processing protocol for isolation and confirmation of *E. coli* isolates from water samples is included as Appendix B. Table 3-2 provides a summary of the number of unknown isolates successfully archived from the 11 different water quality monitoring stations.

Table 3-2 Summary of Unknown Isolates from Water Quality Samples

Station ID	Station Description	Number of Water Samples Collected	Number of *E. coli*-positive Samples Sent to EP AREC	Number of *E. coli* Isolated From Samples	Number of *E. coli* Analyzed by All Four Methods
11942	Lake Waco near dam	100	74	75	67
11945	Lake Waco North Bosque Arm	50	43	63	57
11948	Lake Waco Middle/South Bosque Arm above SH 6	50	43	52	45
TBD3	Lake Waco Middle/South Bosque Arm near inlet of Middle/South Bosque River	50	45	64	54
11953	North Bosque River at SH 56 near Valley Mills	50	22	65	53
11956	North Bosque River at FM 219 northeast of Clifton	50	35	84	72
11921	Belton Lake near dam	100	31	29	27
11922	Belton Lake Cowhouse Creek Arm	50	17	24	24
11923	Belton Lake Leon River Arm near headwater	50	30	49	44
TBD4	Belton Lake Owl Creek Arm	50	32	50	42
11925	Leon River at FM 1829 southeast of North Fort Hood	50	40	76	70
	TOTALS	**650**	**412**	**631**	**555**

SECTION 4
E. COLI KNOWN SOURCE LIBRARY DEVELOPMENT

4.1 COLLECTION, ISOLATION, AND CONFIRMATION OF *E. COLI* FROM KNOWN SOURCES

The BST techniques used in this study are library-dependent. Therefore, it was necessary to construct reference libraries of genetic and phenotypic fingerprints for *E. coli* isolated from known sources (*e.g.,* domestic sewage, livestock, and wildlife) derived from each watershed. By matching the fingerprints of *E. coli* isolated from water samples with the fingerprints in the known source libraries, the likely animal or human (domestic sewage) fecal sources of surface water contamination can be determined.

BST investigators recognize two major considerations when developing source libraries: representativeness and library size. A source library should be representative of the different potential human and animal sources of fecal contamination for the watershed, as well as represent the genetic diversity of the target organism population (in this case *E. coli*) from these different sources. Library size can have an effect on the accuracy of source tracking results, although there is currently no consensus among BST investigators on how to determine the appropriate size of a library. Many genetic-based studies have used relatively small libraries of approximately 35 to 500 source isolates, while many phenotypic-based studies have used larger libraries of approximately 1,000 to 6,000 isolates (Johnson, *et al*. 2004). Genetic-based (genotype) approaches rely on molecular methods used to analyze deoxyribonucleic acid (DNA). Phenotypic-based approaches rely on biochemical methods to identify different reactions by bacteria from different sources. Small libraries may have difficulty identifying environmental isolates, while some of the larger libraries used in prior studies included numerous isolates from the same source sample (many likely identical strains, referred to as "clones"), which may have introduced statistical bias in cross-validation evaluation of the library sensitivity and specificity. Also, depending on the statistical algorithms used for identifying unknowns, large libraries may confound identification.

The primary sampling design consideration for this BST study was to obtain as many unique *E. coli* library isolates as possible from individual animals or samples to represent the diversity and abundance of fecal contamination sources occurring in Lake Waco and Belton Lake watersheds. Budget and time constraints limited library size to approximately 1,000 unique *E. coli* to be isolated from approximately 1,000 different source samples, a moderately large library and high number of source samples. Potential fecal sources were identified through a sanitary survey conducted by Parsons. Municipal wastewater treatment plant influent/effluent and septage samples (collectively referred to as "domestic sewage"), livestock, wildlife, and pet fecal samples, were obtained from a variety of sources throughout the Lake Waco and Belton Lake watersheds based on the sanitary survey. The known source samples collected from the two watersheds and used to establish the local library are summarized in Appendix C.

To the extent possible, known source samples were collected directly from the source feces. An exception was the domestic sewage samples collected from wastewater treatment plants and septic tanks, as opposed to individual human samples. In some

cases, wildlife samples had to be collected indirectly from "found" fecal samples. The sources of these "found" wildlife fecal samples were identified to the lowest practical taxonomic level by experienced field biologists. No samples of uncertain sources were used for library development. Only a single sample was collected from an individual animal. Fresh animal fecal samples were collected aseptically, using a sterile spatula or swab, into sterile, screw-cap polypropylene specimen tubes. Samples were kept on ice and shipped overnight to the EP AREC laboratory for isolation of *E. coli*. Source samples were collected over a 13-month period, from October 2003 to October 2004.

Another important consideration for source tracking studies is the method used to isolate target organisms from water and source samples (USEPA 2005). Since *E. coli* was the target organism in this study and the basis of a regulatory water quality standard, a USEPA-approved method for monitoring this organism in water was used (USEPA 2002). This was important from a regulatory perspective as well as from a scientific one. Historically, microbiologists have known that the culture medium used to isolate bacteria has a significant effect on the types and diversity of the organisms recovered. For example, the types of *E. coli* isolated from source fecal specimens using clinical media may be different from the types of *E. coli* isolated from water using regulatory testing media. In this study, this potential problem was minimized by using the same medium, modified membrane thermotolerant *E. coli* agar (modified mTEC), for the isolation of *E. coli* from source and water samples.

Fecal specimens or domestic sewage samples were streaked (resuspended in buffer if necessary) onto modified mTEC medium, a selective and differential medium for *E. coli*, which is the basis for USEPA Method 1603 detection of *E. coli* in water (USEPA 2002). Inoculated plates were incubated at 35±0.5°C for 2 hours to resuscitate stressed bacteria, then incubated at 44.5±0.2°C for approximately 20-24 hours. The modified mTEC method is a single-step method that uses one medium and does not require testing using any other substrate. The modified medium contains the chromogen 5-bromo-6-chloro-3-indolyl-β-D-glucuronide (Magenta Gluc), catabolized to glucuronic acid, and a red/magenta-colored compound by *E. coli* that produces the enzyme β-D-glucuronidase. This enzyme is the same enzyme tested for using other substrates such as the fluorogenic reaction with 4-methylumbelliferyl-β-D-glucuronide (MUG) observed using UV fluorescence in other *E. coli* assays (*e.g.* IDEXX Colilert and QuantiTray). At least two attempts to isolate *E. coli* were made before considering the sample negative for *E. coli*. *E. coli* colonies from the modified mTEC medium were picked and streaked for purity on nutrient agar with MUG (NA-MUG) to confirm glucuronidase activity and culture purity. A sample processing protocol for isolation and confirmation of *E. coli* from known fecal samples is included as Appendix D.

Between October 2003 and October 2004, a total of 994 fecal samples were collected from known sources. A significant number of these samples did not yield suitable *E. coli* isolates, either because of shipping problems, because no *E. coli* could be recovered from the samples, or because of non-specific or non-*E. coli* bacterial growth on modified mTEC plates from sewage samples. An additional 100 *E. coli* isolates from south Texas non-avian wildlife from a previous BST study were also included (Mott and Lehman 2001).

A total of 813 *E. coli*-positive samples were obtained, with one to three *E. coli* isolates from each samples screened by ERIC-PCR. Some isolates could not be analyzed using PFGE which resulted in the exclusion of 68 samples. In all, 2,275 *E. coli* isolates from source samples were screened by ERIC-PCR. After excluding clonal isolates, one to three isolates per sample (745 samples) for a total of 883 isolates were selected for the library and analyzed by all four BST methods. Table 4-1 provides a summary of samples, *E. coli* isolated from those samples, and total numbers of samples and isolates included in the source library.

4.2 ARCHIVAL OF ISOLATES

Typically one to five *E. coli* isolates were archived for each source sample, while one to 12 isolates per water sample were archived. All additional isolates were archived in the event some isolates did not freeze well or additional isolates for analysis might be required at a later date. Due to the relatively low number of *E. coli* positive samples obtained from water, all the archived isolates were subsequently analyzed. Well-isolated colonies of *E. coli* cultured overnight on either NA MUG or brain heart infusion agar (BHI, the medium used for the DuPont Qualicon RiboPrinter Microbial Characterization System) were resuspended in tryptone soy broth (TSB) with 20 percent glycerol in cryovials and stored at -70 to -80°C. Frozen cultures of selected isolates were provided to the TAMU and TAMU-CC laboratories. A sample processing protocol is included as Appendix E. These frozen cultures can remain viable for years and can be easily shared with other researchers.

4.3 SCREENING OF KNOWN SOURCE *E. COLI* USING ERIC-PCR FOR LIBRARY CONSTRUCTION

As mentioned previously, if identical strains (clones) of *E. coli* from the same sample are included in a BST library, statistical bias may occur, leading to inflated estimates of method performance. A common method of evaluating BST libraries is jackknife analysis. The jackknife analysis used in this study involved pulling each library isolate one-at-a-time from the library and treating each as an unknown to determine the percentage of isolates correctly identified to the true host source. This is referred to as the rate of correct classification (RCC). An example of the statistical bias which may occur for a library containing clones is that isolates will frequently match back to a clone from the same sample, inflating the RCCs. In this study, diversity of the *E. coli* isolates in the library was maximized to increase the likelihood of identifying *E. coli* isolated from water samples. This was accomplished by collecting and analyzing high numbers of known source samples collected from individual animals and selectively including only one to three *E. coli* isolates per sample in the library.

Table 4-1 Summary of Known Source *E. coli* Isolates Used for Library Construction

Source Samples	Desired Number of Samples	Number of Samples Collected	Number of *E. coli*-positive Samples	Number of *E. coli* Isolated and Archived From Samples	Number of *E. coli* Screened by ERIC-PCR	Number of *E. coli* Identified for Library by ERIC-PCR	Number of *E. coli*-positive Samples Used for Library	Number of *E. coli* Isolates in Library (Analyzed by All Four Methods)
Domestic sewage	240	294	186	803	624	229	184	226
Pet	85	56	35	140	95	44	33	42
Cattle	150	173	150	657	440	170	130	147
Other livestock avian	20	23	21	92	59	28	19	25
Other livestock non-avian	108	115	97	413	284	112	79	89
Wildlife avian	191	195	121	559	371	163	111	145
Wildlife non-avian†	306	238	203	567	402	234	189	209
TOTALS	**1100**	**1094**	**813**	**3231**	**2275**	**980**	**745**	**883**

† *Includes 100 South Texas wildlife isolates from a previous BST study (Mott and Lehman 2001).*

4.4 EXCLUSION OF IDENTICAL (CLONAL) *E. COLI* ISOLATES FROM THE SAME KNOWN SOURCE SAMPLE

E. coli isolates from known source samples were screened using a repetitive sequence polymerase chain reaction method (ERIC-PCR). ERIC-PCR is a genetic fingerprinting method used in previous BST studies as well as many microbial ecology and epidemiological studies. ERIC elements are repeat DNA sequences found in varying numbers and locations in the genomes of different bacteria such as *E. coli*. The PCR is used to amplify the DNA regions between adjacent ERIC elements. This generates a DNA banding pattern or fingerprint which looks similar to a barcode pattern. Different strains of *E. coli* bacteria have different numbers and locations of ERIC elements in their bacterial genomes, and therefore, have different ERIC-PCR fingerprints. ERIC-PCR was chosen as the screening technique because of its moderate cost and moderately high ability to resolve different strains of the same species of bacteria. A sample processing protocol for ERIC-PCR is included in Appendix F.

In general, three *E. coli* isolates from each known source sample were fingerprinted using ERIC-PCR. BioNumerics software (Applied Maths, Austin, TX) was used to analyze the ERIC-PCR fingerprints, as well as the other BST data for this project. First, the ERIC-PCR fingerprints obtained from each isolate from a sample were compared to each other using the densitometric curve-based Pearson-product similarity coefficient. An 80 percent similarity cutoff was used, and isolates having more than an 80 percent similarity were considered identical (clones). This similarity cutoff was based on preliminary analysis of replicate ERIC-PCR fingerprints for laboratory QC isolates which were found to be reproducible with approximately 85% similarity. The number of different ERIC-PCR types represented by the isolates from a single sample was determined. A total of 2,275 *E. coli* isolates were screened using ERIC-PCR. A sample processing protocol is included as Appendix G.

4.5 SELECTION OF ISOLATES FOR THE LIBRARY

After determining the number of different *E. coli* ERIC-PCR types and clones for a sample, one to three isolates per sample were selected for inclusion in the library. Isolates representing the different ERIC-PCR types identified for each sample were compared to the *E. coli* isolates selected from other samples for the source library at the time. The same similarity cutoff of 80 percent was used, and isolates having a more than an 80 percent similarity to an existing library isolate were considered already represented in the library, without regard to source. Each isolate from a single known source sample that was novel (<80% similarity) compared to the library isolates was then selected for inclusion in the library. Also, if an isolate had more than an 80 percent similarity to only a single library isolate, then it was also selected for the library, regardless of the selection of other isolates from the same sample. Therefore, clusters were composed of isolates from different samples.

At least one *E. coli* isolate from each known source sample was included in the library. If all ERIC-PCR types represented by the isolates from a single sample were already present in the library, then an isolate representing the most abundant (and as such

being the most representative) ERIC-PCR type for that sample was selected for the library. Therefore, abundant/common strains of *E. coli* isolates from different samples and animals were represented in the library. If isolates were equally abundant in a sample, then the isolate that best filled in the tree was selected. Building the library was a dynamic process; isolates were added to the library as their ERIC-PCR patterns were processed.

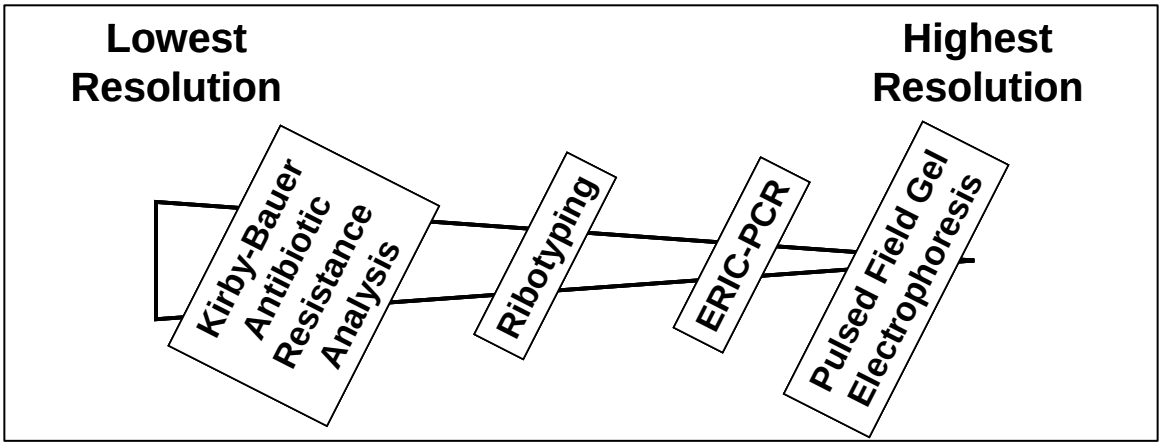
In addition, 100 *E. coli* isolates from known wildlife sources obtained in a previous BST study in south Texas were characterized and included in the library (Mott and Lehman 2001). This not only helped to increase the diversity of wildlife *E. coli* isolates included in the library, which often can be difficult to obtain, but also helped to make progress toward development of a statewide library. A total of 980 *E. coli* isolates were identified for inclusion in the library using ERIC-PCR. Of these 980, 883 isolates from 745 source samples were successfully analyzed by the four BST methods described in Section 5 and were used to construct the source library (Table 4-1).

SECTION 5
DESCRIPTION OF BST METHODS

Of the molecular tools available, ribotyping and PFGE DNA fingerprinting are recognized as promising BST techniques. A phenotypic (non-genetic) characterization method, Kirby-Bauer antibiotic resistance analysis (KB-ARA), also has the potential to identify the human or animal origin of *E. coli* isolates. Currently, there is still no consensus among BST researchers as to the single most useful BST method.

All methods selected for this project were used in previous BST studies, as well as other microbiological studies, and have been published in peer-reviewed journals. They span the spectrum in their ability to resolve differences in related bacterial strains, technical training and labor required, equipment cost, reagent cost, sample throughput, and ease of data analysis. Figure 5-1 displays a conceptual sensitivity continuum of the BST methods used in this study. Phenotypic based methods are at the less sensitive domain of the continuum, while genotypic based methods constitute the more sensitive end of the spectrum.

Figure 5-1 Discriminatory Capabilities of Different BST Methods



In this study, the combination of the DuPont RiboPrinter System, PFGE, ERIC-PCR, and KB-ARA allowed evaluation and development of one of the most comprehensive *E. coli* BST libraries and one of the first side-by-side comparisons of these BST methods. Recent BST studies also suggest that a combination of complementary methods may be the best approach for accurately identifying sources of contamination.

The validity of this study and the conclusions drawn from the results are strengthened through the use of the multiple techniques, and in particular, composite data sets. Further, by using standardized methods, the library can be expanded through future projects and the data shared with other BST investigators and regulatory agencies. Peer-reviewed publication of project results is also a goal, and journal manuscripts are being prepared as of the date of this report.

5.1 RIBOTYPING

Ribotyping is a genetic fingerprinting method used in previous BST studies as well as many microbial ecology and epidemiological studies, although there is not a consensus as to the best protocol. In general, an endonuclease enzyme (*Hin*d III) selectively cuts *E. coli* DNA wherever it recognizes a specific DNA sequence. The resulting DNA fragments are separated by size and probed for fragments containing particular conserved ribosomal RNA gene sequences, which results in DNA banding patterns or fingerprints that look similar to barcode patterns. Different strains of *E. coli* bacteria have differences in their DNA sequences and different numbers and locations of enzyme cutting sites, and therefore have different ribotyping fingerprints.

By automating the process, the DuPont Qualicon RiboPrinter Microbial Characterization System can analyze up to 32 samples per day, whereas manual ribotyping methods may require up to several days to complete. All bacterial isolate sample processing is automated using standardized reagents and a robotic workstation, providing a high level of reproducibility. The RiboPrinter was originally developed for use in identification and BST of microbial isolates for the food industry. Since the system employs standardized methods and reagents, results obtained from other laboratories using the system are directly comparable. RiboPrinting has a moderate ability to resolve different strains of the same species of bacteria. Although the automated system saves time and requires little training, the initial investment and the processing cost per isolate are expensive. A sample processing protocol for RiboPrinting is included in Appendix H.

5.2 PULSED FIELD GEL ELECTROPHORESIS

Pulsed-field gel electrophoresis (PFGE) is another leading genetic fingerprinting method used in BST. The entire bacterial genome is fragmented using an infrequent cutting restriction endonuclease enzyme (e.g. *Xba* I) which cuts DNA wherever it recognizes a specific rare sequence. All the DNA fragments are separated by size and visualized resulting in a genetic fingerprint that resembles a barcode. Different strains of *E. coli* bacteria have differences in their DNA sequences and different numbers and locations of enzyme cutting sites and therefore, have different PFGE fingerprints.

PFGE is currently being used by the Centers for Disease Control and Prevention (CDC) to track foodborne *E. coli* O157:H7 and *Salmonella* isolates. TAMU Food and Environmental Microbiology Laboratory followed the standardized CDC protocol using *Xba* I for PFGE analysis of *E. coli* in this BST study. CDC currently uses this standardized protocol as the basis of their "PulseNet" outbreak surveillance network which allows public health laboratories nationwide to quickly compare their PFGE fingerprints to the CDC central reference library. A sample processing protocol is included in Appendix I.

Although it requires more training and cost, PFGE has very high resolution and can discriminate between closely related strains. While this allows higher confidence in the matches made, fewer identifications can be made, even with lower similarity cutoffs. In addition, some bacterial strains have genomic DNA in configurations that do not permit

effective restriction endonuclease digestions. The inability to obtain patterns from replicates implies there is an issue at the organism/genome level rather than the procedure. It should be noted that PFGE patterns could not be generated for approximately 10 percent of the known library isolates and water isolates, limiting the number of isolates that could be analyzed by all four methods from 980 to 883 library isolates, and from 631 to 555 unknown water isolates.

5.3 ENTEROBACTERIAL REPETITIVE INTERGENIC CONSENSUS POLYMERASE CHAIN REACTION (ERIC-PCR)

ERIC-PCR was described in Section 4. A sample processing protocol for ERIC-PCR is included in Appendix F.

5.4 KIRBY-BAUER ANTIBIOTIC RESISTANCE ANALYSIS

The KB-ARA method was performed by the TAMU–CC Environmental Microbiology Laboratory. This technique followed methods used in the clinical laboratory for evaluating the antibiotic resistance of bacterial isolates. Commonly, the disk diffusion method is used which involves measuring the diameter of the zone of inhibition of bacterial growth around a filter disk impregnated with a specific antibiotic. By comparison to resistant and susceptible control strains, the response of the *E. coli* isolates can be determined. To further standardize and automate the assay, an image analysis system was used to measure the zones of inhibition and provide electronic archival of data. The KB-ARA profile for an isolate consists of the measurements of the zones of inhibition in response to 20 antibiotics, each at a standard single concentration. Discriminant analysis is the standard statistical analysis tool for KB-ARA results. A sample processing protocol is included as Appendix J.

Of the four methods applied in this study, KB-ARA has the lowest ability to discriminate closely related bacterial strains. However, it also has the lowest initial and per sample cost and takes the least time and training, although the statistical analysis can be complex.

THIS PAGE INTENTIONALLY LEFT BLANK

SECTION 6
BST DATA PROCESSING AND ANALYSIS USING BIONUMERICS

BioNumerics software (Applied Maths, Austin, TX) was used to analyze the BST data for this project. This software is currently used in many BST studies, especially for molecular fingerprint methods. For this project, BioNumerics was used in three ways: 1) processing gel images, 2) determining the relationships between the isolates by comparing their molecular fingerprint patterns for development of the BST libraries, and 3) for identification of water isolates.

6.1 GEL ELECTROPHORESIS IMAGE PROCESSING

Once the ERIC-PCR and PFGE gels were completed, digital photos were taken of the gels containing fingerprints of the isolates. The digital images were then processed using BioNumerics software in a four-step procedure. First, the area of the gel image and the individual lanes were defined. Second, spectral analyses of the densitometric curves were optimized. Third, the marker lanes were defined and normalized, and fourth, the fingerprint was linked to the isolate-designated key in the BioNumerics database. RiboPrints were partially processed by the Qualicon software and were imported into BioNumerics using the Load Samples Amplified script from DuPont Qualicon.

6.2 CORRELATION COEFFICIENTS, MATCHING CRITERIA

The curve-based Pearson-product similarity coefficient allowed for both the position and the relative intensity of the fingerprint bands to be considered in the comparisons, not just the presence or absence of a band (*i.e.* band-matching). From these comparisons, dendrograms (or family trees) were constructed using the unweighted pair group method with arithmetic means (UPGMA) to depict the relationships between the isolates.

An epidemiological-like approach (*i.e.* one-to-one matching) was used for each of the molecular techniques to find the best match of an unknown isolate's fingerprint to the fingerprint of a single library isolate. To accomplish this, the BioNumerics Best Matches script provided by Applied Maths was used. A fingerprint was not considered a match unless it was at least 85 percent similar to another single entry for the ERIC-PCR and RiboPrinting techniques, or 70 percent for the PFGE fingerprint. This similarity cutoff was based on reproducibility of the fingerprint obtained with the *E. coli* QC101 strain (RiboPrinter System QC strain). Repeated analyses of the *E. coli* QC101 strain were run at approximately once per ERIC-PCR or PFGE batch/gel or daily for the RiboPrinter. Densitometric Pearson-product coefficient matching is so stringent that it is essentially not possible for two different molecular fingerprints to be 100 percent identical. It should be noted that fingerprints, for example RiboPrint patterns, greater than 85 percent similar as determined by Pearson-product coefficient matching, often appear identical to the naked eye.

6.3 COMPOSITE DATA SETS

Fingerprint patterns were compared individually for each of the molecular methods (ERIC-PCR, RiboPrinting, and PFGE) using the BioNumerics software as described above. An equivalent comparison of the KB-ARA profile treated the zone of inhibition measurements for each of the 20 antibiotics as character data to compare isolate profiles using the BioNumerics software. While KB-ARA usually uses discriminant analysis to identify sources of isolates in a population biology approach (see below), KB-ARA data were also analyzed using BioNumerics and the Pearson-product coefficient.

BioNumerics has the unique ability to allow the construction of composite data sets. A composite data set takes into account all four fingerprint profiles (ERIC-PCR, RiboPrint, PFGE, and KB-ARA), or various combinations of profiles for each isolate. Each BST method describes a different trait or aspect with different degrees of resolution or discrimination. An analogy would be the composite sketch of a suspect gathered from the descriptions of eye-witnesses to a crime, each having a different perspective.

To incorporate all four BST methods, BioNumerics was used to calculate a composite data set using the unweighted averages of the individual method similarity matrices, resulting in a new, single similarity matrix which incorporates attributes of each individual method. Composite data sets for only known library isolates and water isolates that had patterns for all four BST methods were created. Since some isolates were not amenable to PFGE and did not generate a fingerprint with this method, this reduced the number of composite data set library source isolates from 980 to 883, and the number of water isolates from 631 to 555. The similarity cutoff used for comparing composite set data was 70 percent to allow for variation of the individual methods. This similarity cutoff was confirmed by calculating the rates of correct classification for the library source isolates, and discerning where there were fewest cross-identifications balanced with the number of water isolates that would be left unidentified (Table 6-1).

Although fingerprint profiles are considered a match to a single entry, identification is to the host source class, and not to the individual animal represented by the best match. Host sources were divided into seven groups, 1) domestic sewage; 2) pet; 3) cattle; 4) other livestock, avian; 5) other livestock, non-avian; 6) wildlife, avian; and 7) wildlife, non-avian. The division of host sources into these particular classes was based on discussion with project participants and anticipated usefulness for the development of BMPs.

Table 6-1 Comparison of Similarity Cutoffs for the Four-Method Composite Data Set and Effect on Library RCC and Percentage of Source and Water Isolates Left Unidentified

4-Method composite data set	Percent unidentified water (555)*	Domestic sewage (226)		Pet (42)		Cattle (147)		Other livestock, avian (25)		Other livestock, non-avian (89)		Wildlife, avian (145)		Wildlife, non-avian (209)	
Similarity cutoff		% RCC	% no ID	% RCC	% no ID	% RCC	% no ID	% RCC	% no ID	% RCC	% no ID	% RCC	% no ID	% RCC	% no ID
65%	3	83	6	30	5	60	1	21	4	39	1	49	6	65	7
70%	9	83	15	33	14	61	3	22	8	40	8	48	11	66	11
75%	24	84	30	36	33	64	15	29	44	43	22	52	29	70	22
80%	56	91	48	44	57	66	46	11	64	56	54	62	48	73	41

**Number of isolates in parentheses*

*** % no ID = percentage of isolates in the source class that were below the similarity cutoff and were therefore left unidentified*

6.4 DISCRIMINANT ANALYSIS

Discriminant analysis is a commonly used approach for analysis of antibiotic resistance profiles. Discriminant analysis is a multivariate technique that can be used to classify items into categories based on a set of test variables. The RCC for each source can be used to evaluate the predictive capabilities of the database. Additional analyses can use step-wise methods such as Wilks' lambda and Mahalanobis' distance as alternatives to remove variables which do not contribute or contribute little to the classification, in order to improve the ability to correctly classify sources. As classifications based upon the cases used to create the model tend to be too "optimistic" in the sense that their classification rate is inflated, cross-validation is performed by classifying each case while leaving it out from the model calculations (leave-one-out method); however, this method is generally still more "optimistic" than subset validation. Subset validation can be performed using the holdout method.

In this study, discriminant analysis was performed on antibiotic resistance zone of inhibition diameters of each isolate. The zone diameters were compiled into a library of known sources to be analyzed using discriminant analysis with SPSS™ Version 12.0 for Windows™, following the SPSS Base 9.0 Applications Guide, SPSS Online tutorial, and Discovering Statistics using SPSS for Windows.

SECTION 7
EVALUATION OF BST TECHNIQUES

7.1 PRECISION AND ACCURACY OF TECHNIQUES

Precision is a measure of the reproducibility of the methods using replicates. Accuracy reflects the truth of the identification of an isolate. Both precision and accuracy were tested in the routine laboratory controls, as well as a special blind quality control (QC) study performed as described below.

7.2 LABORATORY CONTROLS

Standard laboratory practices were followed to assure quality data. All data obtained from field and laboratory measurements were reviewed and verified for integrity and continuity, reasonableness, and conformance to project requirements. Each laboratory method included additional required QC measures. Quality control for the molecular techniques (ERIC-PCR, RiboPrinting, and PFGE) was achieved by using a QC strain of *E. coli* (*E. coli* QC101, ATCC #51739) provided by DuPont Qualicon for the RiboPrinter System. *E. coli* QC101 was analyzed under the same conditions as the library and water isolates on every ERIC-PCR and PFGE gel over the course of the study. It was also included in one of every four RiboPrinting batches, or each day the RiboPrinter was run. Gels or batches which did not yield typical *E. coli* QC101 fingerprints were not used and prompted a review of procedures, equipment, and supplies to identify and correct any QC issues. Blanks were also included in every batch of ERIC-PCR (and gel) to ensure there was no contaminating DNA in any of the reagents that could be amplified by the PCR. Normal variations in the precision of the fingerprint patterns generated for *E. coli* QC101 over time using the different BST methods were the basis of the matching criteria (similarity cutoffs) for the different methods. For the KB-ARA biochemical method, every tenth isolate was run in duplicate to assure reproducible profiles.

7.3 BLIND QC STUDY

BST does not lend itself easily to the same QC methods as chemical quantification because each measurement is essentially qualitative, not quantitative. Blank samples are less relevant, and replicate water samples may often yield different *E. coli* strains. Therefore, in this study laboratory method accuracy and precision were quantified through a special QC study with blinded safeguards. The TSSWCB project manager used a random number generator to select 60 isolates from a list of 1,024 different known source isolates (980 *E. coli* and 44 confounding isolates) collected in the Lake Waco and Belton Lake watersheds as part of this study, and previously studied south Texas wildlife isolates. From the list of 60, the TSSWCB project manager narrowed the list down to 30 by eliminating many duplicate species and ensuring that the percentage of domestic sewage isolates was similar to the percentage of this source class in the total library. The list of the 30 selected isolates was provided to EP AREC, and triplicate cultures of each isolate were prepared and sent to the Parsons project manager. The Parsons project manager selected 10 of the 30 isolates, blind labeled them, and sent the triplicate cultures of each to the BST labs. The samples were processed through the PFGE, RiboPrinting,

ERIC-PCR, and KB-ARA procedures in a blind fashion; that is, the laboratories did not know the sources. Each lab attempted to identify the 10 sets of triplicates, identify the triplicates to the correct library isolate (of the 30 possible), and to the correct source class. After each lab reported its individual results to the TSSWCB project manager, the labs shared their raw data with EP AREC for composite data set analysis. After EP AREC reported the composite data set results to the TSSWCB project manager, the key of the blind isolates was provided to all participants. Identifications were based on best matching and dendrograms for the three molecular methods and composite data sets, while the KB-ARA data were analyzed using discriminant analysis.

It should be noted that the ERIC-PCR, RiboPrinting, PFGE, and composite data set results were based on best matching to the QC library of 30 possible isolates, as well as visual inspection of the similarity dendrograms and subjective judgment. Fingerprints for the unknown QC isolates were grouped in their own dendrogram to identify the triplicates. Identification of the replicates was relatively straightforward for the ERIC-PCR, PFGE, and the composite data set, but was more complex for RiboPrinting data. The dendrogram was used, as well as a process of elimination. To identify the unknown isolates, their patterns were compared to the 30 QC library isolates used in the study.

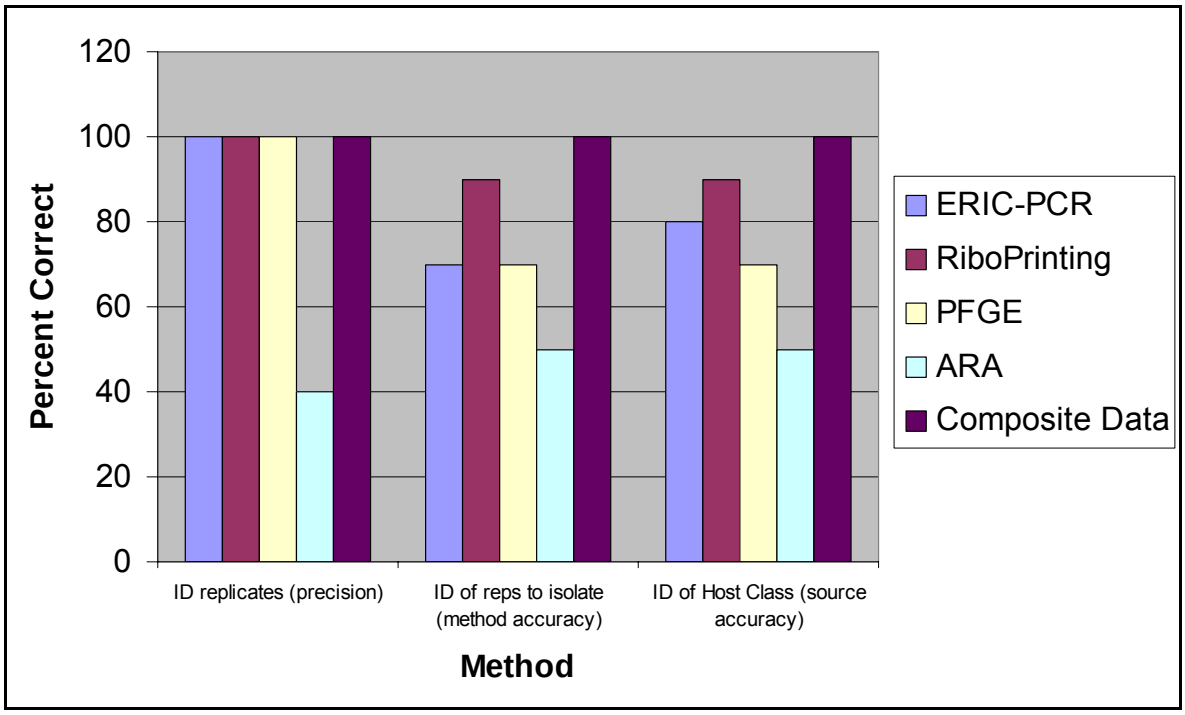
Note that PFGE fingerprints could not be generated for four of the 30 possible QC library isolates, and by chance two of these were used as blind challenge isolates. Since they could not be included in the PFGE QC library and identification was not possible with this method, PFGE was not penalized for its inability to identify these isolates.

It should also be noted that the QC study design was more suitable for the genetic analyses, where one-to-one matching of fingerprints is performed, than for the KB-ARA method using discriminant analysis. The KB-ARA profiles for the isolates selected for the QC study were not distinct enough to easily distinguish or group the isolates into the three replicates. KB-ARA is generally used to classify isolates into categories rather than individual species sources. The identifications and groupings were based on discriminant analysis and visual comparisons of the zone diameter data and subjective judgment. However, using single isolates makes discriminant analysis statistically inappropriate.

Overall, the ERIC-PCR, RiboPrinting, and PFGE techniques performed equally well, with 100 percent identification of replicate isolates (precision) and 70 to 90 percent accuracy in identification of replicate isolates to specific QC library isolate (method accuracy) and correct source class (source identification accuracy) (Figure 7-1). On the other hand, the KB-ARA scored only 40 percent for identification of replicate isolates (precision) and 50 percent for method and source identification accuracy, likely due to the reasons given above.

Most important, however, are the four-method composite data set results. The composite data set results correctly identified 100% of the replicate QC cultures (precision), and had 100% accuracy for *E. coli* strain and source class identification of the isolates. Therefore, the composite four-methods performed better than any single method.

Figure 7-1 Blind Quality Control Study Results



7.4 LIBRARY EVALUATION

7.4.1 Jackknife Analyses of Each BST Method

Jackknife analysis is treating each known library isolate as if it were an unknown by trying to identify it against the rest of the library. By comparing the true identities to the given identities, the RCCs can be calculated. The RCC is the percentage of library isolates correctly identified back to their source out of the number of attempts made to identify isolates from that source. Isolates which are left unidentified are not directly reflected in the RCC calculation. While there is no agreement as to the minimum acceptable RCC, the rates should at least reflect a better than random chance of source class identification. Some factors are known to confound the calculation. Inclusion of bacterial clones from the same sample can artificially inflate the RCC, while large libraries with increased isolate diversity can decrease the RCC and increase the number of isolates left unidentified.

To determine RCCs, matches to host source class were used, as opposed to matches with specific animal species. For example, an isolate from a wild goose matching with an isolate from a wild duck was considered a correct match for the avian wildlife source class. In rare instances (<1%), there were ties for the best match (same percentage of similarity). In these cases, the benefit of the doubt was given and the isolate most similar in host source class was selected as the match.

The RCCs for the individual and four-method composite BST Library of 883 isolates (Table 7-1) were lower than often reported in the literature for BST studies. However, these RCCs are higher than average compared to a recently published extensive and more

unbiased comparison of seven different BST protocols (Stoeckel, *et al*. 2004). While perhaps detrimental to the conventional statistical evaluation of BST libraries, moderately large libraries containing diverse isolates, as developed in this study, are more likely to reflect the potential diversity of the unknown water isolates they are meant to identify.

PFGE tended to have the highest RCCs (Table 7-1). However, almost half of the library isolates were left unidentified (*i.e.,* did not match any other library isolate) by the jackknife analyses. Interestingly, RCCs for the KB-ARA data using either best matching or discriminant analysis were fairly similar. The four-method composite data library performed well and had some of the highest RCCs. In particular, the composite library had good RCC values for source classes of special interest: 83 percent for domestic sewage and 61 percent for cattle.

Table 7-1 Jackknife Analysis Rates of Correct Classification (%) for Individual and Composite BST Methods for the 883 Isolate BST Library **

	Random*	PFGE	ERIC-PCR	RiboPrinting	KB-ARA using best matching	KB-ARA using discriminant analysis	Four-method composite data set
Domestic sewage	26	95 (35)**	64 (29)	60 (2)	60 (3)	43 (0)	83 (15)
Pet	5	54 (69)	19 (38)	17 (0)	17 (0)	27 (0)	33 (14)
Cattle	17	80 (60)	46 (13)	43 (4)	41 (0)	27 (0)	61 (3)
Other Livestock avian	3	0 (60)	10 (20)	0 (0)	8 (4)	36 (0)	22 (8)
Other livestock non-avian	10	55 (55)	30 (20)	16 (3)	24 (2)	10 (0)	40 (8)
Wildlife Avian	16	74 (52)	37 (27)	40 (6)	35 (2)	41 (0)	48 (11)
Wildlife Non-avian	24	84 (49)	55 (17)	47 (5)	60 (0)	44 (0)	66 (11)

**Random is the percentage of isolates from each source class represented in the library of 883 source isolates.*

***The number in parentheses is the percentage of isolates for that source class left unidentified after jackknife analyses (<85% similarity for ERIC, RiboPrinting and KB-ARA best match, <70% similarity for PFGE and the composite data set). There is not an unidentified classification or a minimum similarity in discriminant analysis.*

7.4.2 Interpretation of Source Class Cross-Identification of *E. coli* Library Isolates

There will naturally be some *E. coli* isolates that are not host-specific and can come from several host source classes, leading to possible cross-identification to different source classes for these promiscuous isolates. There are two general ways to approach source cross-identification. One is to determine how isolates of a true host source class are identified as reflected in the RCC (as reported above). The second approach is to determine the true identifications of those isolates identified as a particular host source class.

Cross-validation jackknife analysis of the four-method composite data library revealed that, in general, there were low levels of cross-identification for the host source classes (Figure 7-2). For each of the seven source classes, the highest bar within each source class (X-axis) belongs to the correct source and corresponds to the RCC for that particular source class. Only the other livestock avian class had cross-identifications in another single source class (cattle) greater than its correct classification. This may have been due to the low numbers of isolates from this host source class in the library. Also, the attempt to divide livestock into cattle, non-avian, and avian livestock source classes may be a challenge for the number and types of samples collected in this study.

Both correct and cross-identifications are shown, with all columns of the same color adding up to 100 percent of the identification attempts for the library isolates in each source class. The black bars represent the random chance that an isolate would be identified to each source class based on library composition. Cross-identifying isolates were not excluded from the library since a best match algorithm was used.

7.4.3 Library Quality Measures

The four-method composite library was further evaluated for sensitivity and specificity as described in the USEPA *Microbial Source Tracking Guide Document* (USEPA 2005) for both seven-way and two-way splits of source classifications (Tables 7-2 and 7-3, respectively). Sensitivity reflects the percentage of isolates giving a host source-specific fingerprint, and is also referred to as the RCC. Specificity measures how well a BST method can discriminate between source categories. Specificity is determined by performing jackknife analyses, and examining the identifications of known source isolates from other than the respective source. The number of isolates that test negative (or true negatives) and are correctly identified as not belonging to the test source class are divided by the sum of this number plus isolates incorrectly identified to the respective source class (false positives). Although there is no consensus, specificity values below 80 percent are considered of questionable discriminatory power (USEPA 2005). In this study, the four-method composite library specificity values for each source class were all above 80 percent for both the seven-way and two-way split of source classifications.

Figure 7-2 Cross-Validation Jackknife Analysis of the Four-Method Composite Data Library

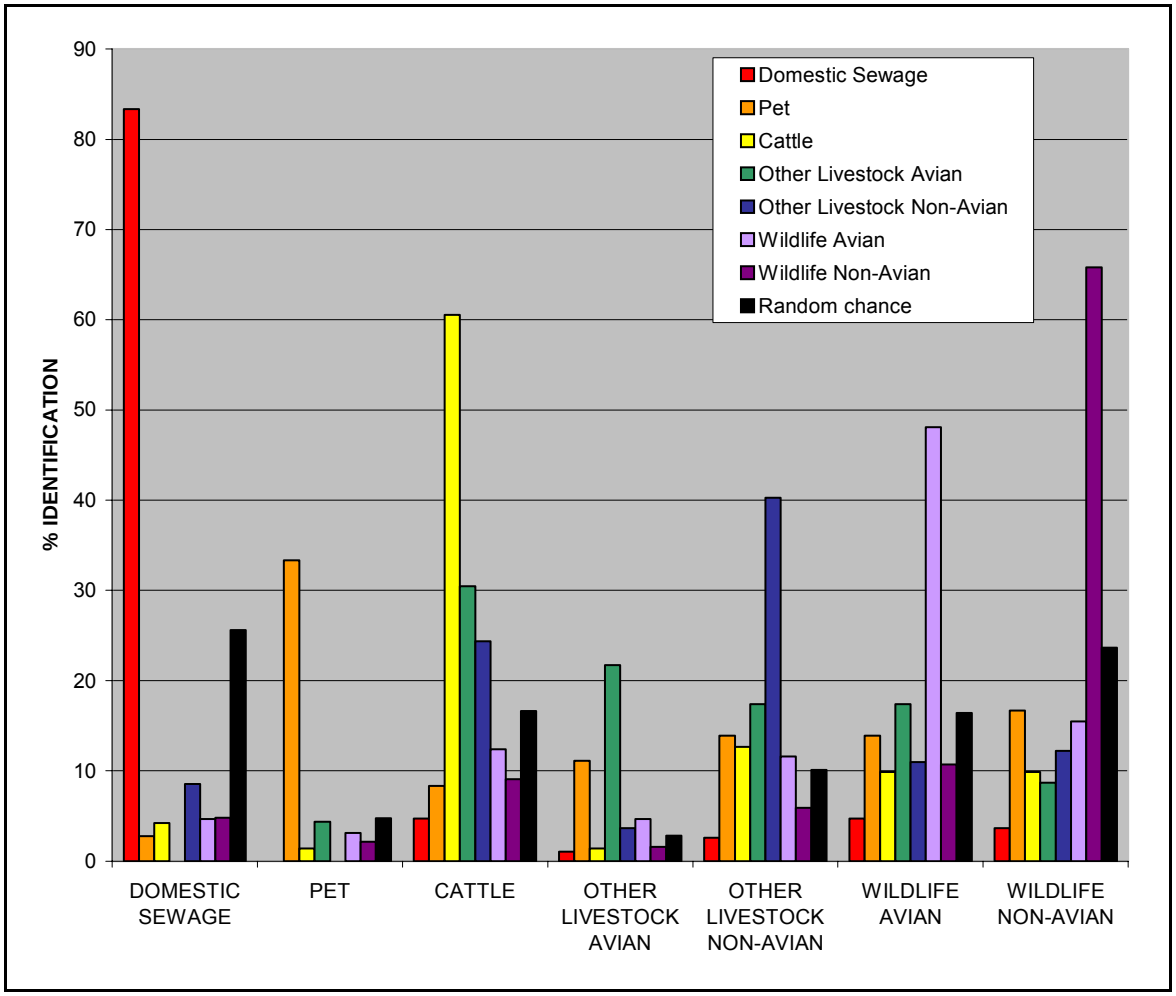


Table 7-2 Quality Measures for the Cross-Validated (Jackknifed) Composite Data Set Library and Seven-Way Split of Source Classes*

	Sensitivity (RCC)	Specificity	Positive predictive value	Random predictive value
	TP/(TP+FN) x 100	TN/(TN+FP) x 100	TP/(TP+FP) x 100	
Domestic sewage	83	95	85	26
Pet	33	99	52	5
Cattle	61	89	54	17
Other livestock, avian	22	97	20	3
Other livestock, non-avian	40	92	36	10
Wildlife, avian	48	91	50	16
Wildlife, non-avian	66	90	68	24

**RCC, rate of correct classification (also known as sensitivity). TP, test positive, isolate correctly identified to its source class; FN, false negative, isolate not identified to its true source class; TN, test negative, isolate correctly identified as not from the particular source; FP, false positive, isolate incorrectly identified as from a given source. The following numbers of isolates that were unidentified after cross-validation analyses (<70% minimum similarity) were not used in these calculations: 34 of 226 Domestic Sewage isolates; 6 of 42 Pet isolates, 4 of 147 Cattle isolates; 2 of 25 Other Livestock, Avian isolates; 7 of 89 Other Livestock, Non-avian isolates; 16 of 145 Wildlife, Avian isolates; and 22 of 209 Wildlife, Non-avian isolates.*

An attempt was also made to calculate the positive predictive values of the four-method composite library for each source class based on seven-way and two-way splits of source classifications. The positive predictive value is the percentage of library isolates identified to a given source class through jackknife analysis that are truly from that source class. Therefore, based on the library characteristics, the positive predictive value provides an estimate of the correct classification of unknown isolates (i.e., the predicted RCC). The random predictive values are based on the composition of the library by percentage of source classes. In all cases the sensitivity (RCC) and positive predictive values are much greater than random chance.

Table 7-3 Quality Measures for the Cross-Validated (Jackknifed) Composite Data Set Library and Two-Way Split of Source Classes*

	Sensitivity (RCC)	Specificity	Positive predictive value	Random predictive value
	TP/(TP+FN) x 100	TN/(TN+FP) x 100	TP/(TP+FP) x 100	
Domestic sewage	83	95	85	26
Animal	95	83	95	74

**RCC, rate of correct classification (also known as sensitivity). TP, test positive, isolate correctly identified to its source class; FN, false negative, isolate not identified to its true source class; TN, test negative, isolate correctly identified as not from the particular source; FP, false positive, isolate incorrectly identified as from a given source. Isolates that were unidentified after cross-validation analyses (<70% minimum similarity) were not used in these calculations.*

7.5 DISCRIMINANT ANALYSIS OF KB-ARA DATA

Data were first analyzed following the designated groups on the data sheets received from EP AREC. Isolates were labeled with the following classifications: avian, wildlife, pet, avian wild, livestock, and human. These were modified following discussion with the other project participants to form a seven-way classification (as presented throughout this report) of domestic sewage ("human"), wildlife avian, wildlife non-avian, pets, cattle, other livestock avian, and other livestock non-avian. The RCCs for the KB-ARA data analyzed with discriminant analysis and the seven-way split of source classifications are included in Table 7-1.

The KB-ARA discriminant analysis RCCs for the two-way split of source classes into domestic sewage and animal were 60 percent and 82 percent, respectively (data not shown). Discriminant analysis of four-way classification – domestic sewage (sewage and pet isolates), livestock (cattle, non-avian livestock), avian (avian wildlife and avian livestock) and wildlife (non-avian only) was also performed. These categories were more easily distinguished than the seven-way split classes and provided categories suitable for management purposes. RCCs ranged from 45 percent to 53 percent for each of these source classes.

The database was examined for appropriateness for antibiotic resistance analysis. Several concerns were identified. First, there was a lack of information regarding many of the sample collections and antibiotic usage. For example, some source samples were collected from state fair livestock animals (35 of 883 four-method composite library isolates), and may have been exposed/treated with antibiotics that would not represent the exposure of the majority of livestock animals impacting the watershed. Second, there were many animal sources that were unlikely to make a significant contribution to fecal contamination of the water bodies, either due to their physical size, habits, or population numbers. Ideally, to include those animal sources, the number of isolates from those sources in the library would have had to have been increased for discriminant analysis. Unfortunately, removing those sources from would have resulted in an unacceptably small database. Compromises were made to optimize the library for discriminant analysis of the KB-ARA within the constraints of the project. A number of isolates were re-categorized, primarily in the avian wildlife group, based on subjective judgment, field information provided, and the desired discriminate analysis classifications. In general, these issues are not problematic for the ERIC-PCR, RiboPrinting and PFGE analyses.

SECTION 8
OBSERVED *E. COLI* LEVELS IN LAKE WACO, BELTON LAKE, AND THEIR MAJOR TRIBUTARIES

Overall, the levels of *E. coli* observed in the samples collected from Lake Waco were far below the contact recreation geometric mean water quality criterion of 126 *E. coli* per 100 ml (Table 8-1). The geometric mean concentrations ranged from 5 cfu/100 ml at the deep water site near the dam to 11 cfu/100 ml near the inlet of the Middle and South Bosque Rivers (Figure 8-1). No samples from Lake Waco exceeded the single sample criterion of 394 cfu/100 ml for *E. coli*. The highest observed concentrations were in the Middle and South Bosque River arm.

In the North Bosque River, *E. coli* concentrations were more dynamic, as expected, since the volume of water for dilution of loads is less than that in the lake. *E. coli* were absent from more than half the samples at the site just upstream of Lake Waco, but one sample collected under runoff influences had a concentration of 8,500 cfu/100 ml. The geometric mean *E. coli* concentration of the samples at this site was 6 cfu/100 ml, and 14 percent of samples exceeded the single sample criterion, which indicates support for contact recreation use. Further upstream on the North Bosque River at Clifton (Station 11956), *E. coli* levels were similar, with a geometric mean concentration of 9 cfu/100 ml and 12 percent of samples exceeding the single sample criterion.

In Belton Lake, *E. coli* levels were very low throughout the lake. *E. coli* were absent from 131 (52%) of the 250 samples collected. Geometric mean *E. coli* concentrations ranged from 1 cfu/100 ml near the dam to 4 cfu/100 ml in the Leon River arm which are far below the contact recreation geometric mean water quality criterion of 126 *E. coli* per 100 ml (Figure 8-2). No samples exceeded the single sample criterion.

The highest *E. coli* levels of any site sampled in this study were observed in the Leon River upstream of Belton Lake (Station 11925). Concentrations ranged from 0 to more than 20,000 cfu/100 ml. The geometric mean *E. coli* concentration was 41 cfu/100 ml, well below the geometric mean criterion of 126, which indicates support of contact recreation use, but the single sample criterion was exceeded in 14 (28%) of the samples, which indicates nonsupport of contact recreation use. The levels were particularly high after runoff events, but on one date (April 20, 2004) *E. coli* levels exceeded the single sample criterion in dry weather, which may indicate the influence of contamination from a point source. A table summarizing the sampling results for each station is provided in Appendix A.

Table 8-1 Observed *E. coli* Levels from Water Quality Samples Collected

Station ID	Site	Sample count	Samples with *E. coli*	Concentration Range (#100 ml)	Geometric Mean Concentration	Samples >WQ Criterion (%)
11942	Lake Waco near Dam	100	74 (74%)	<1 – 78	5	0
11945	Lake Waco N. Bosque Arm	50	43 (86%)	<1 – 170	6	0
11948	Lake Waco Mid/ South Bosque Arm above SH 6	50	43 (86%)	<1 – 327	8	0
TBD3	Lake Waco Mid/ South Bosque Arm near inlet	50	45 (90%)	<1 – 360	11	0
11953	N. Bosque River near Valley Mills	50	22 (44%)	<1 – 8,500	6	7 (14%)
11956	N. Bosque River NE of Clifton	50	35 (70%)	<1 - 860	9	6 (12%)
11921	Belton Lake near dam	100	33 (33%)	<1 – 7	1	0
11922	Belton Lake Cowhouse Creek Arm	50	21 (42%)	<1 – 67	2	0
11923	Belton Lake Leon River Arm	50	32 (64%)	<1 – 361	4	0
TBD4	Belton Lake Owl Creek Arm	50	33 (66%)	<1 – 50	3	0
11925	Leon River at FM 1829 SE of North Fort Hood	50	39 (78%)	<1 - >20,000	41	14 (28%)

Figure 8-1 Geometric Mean of *E. coli* Concentrations - Lake Waco Sampling Sites, September 2003 through June 2004

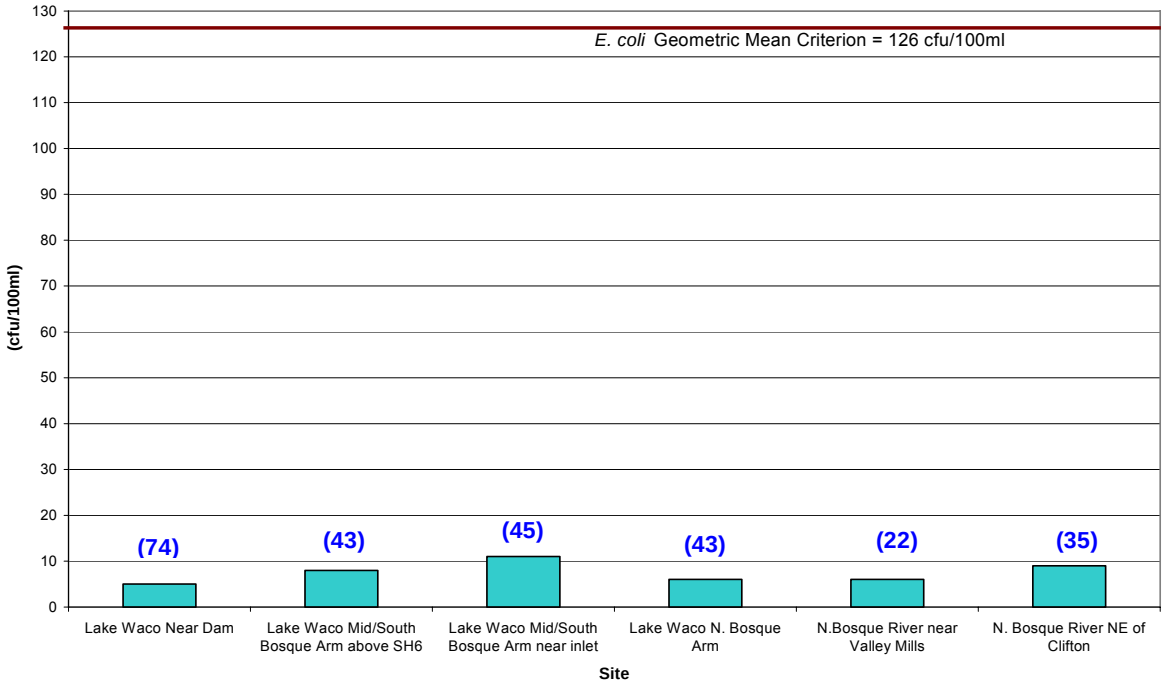
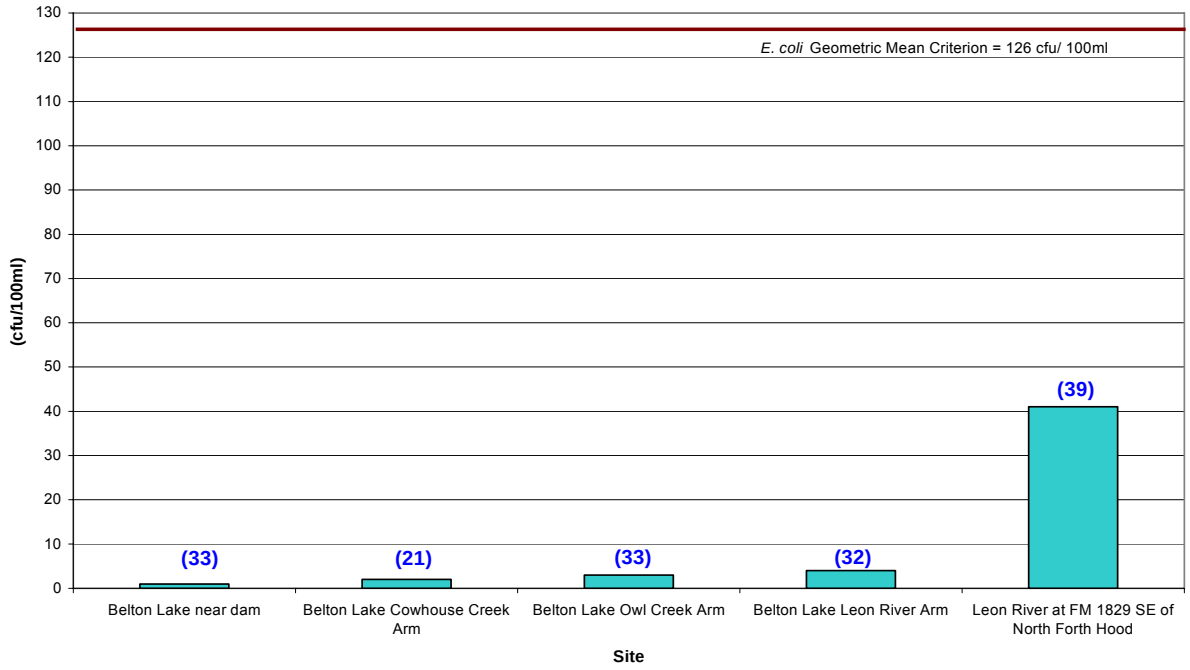


Figure 8-2 Geometric Mean of *E. coli* Concentrations - Belton Lake Sampling Sites, September 2003 through June 2004



THIS PAGE INTENTIONALLY LEFT BLANK

SECTION 9
BST RESULTS

9.1 SOURCE IDENTIFICATION OF WATER ISOLATES

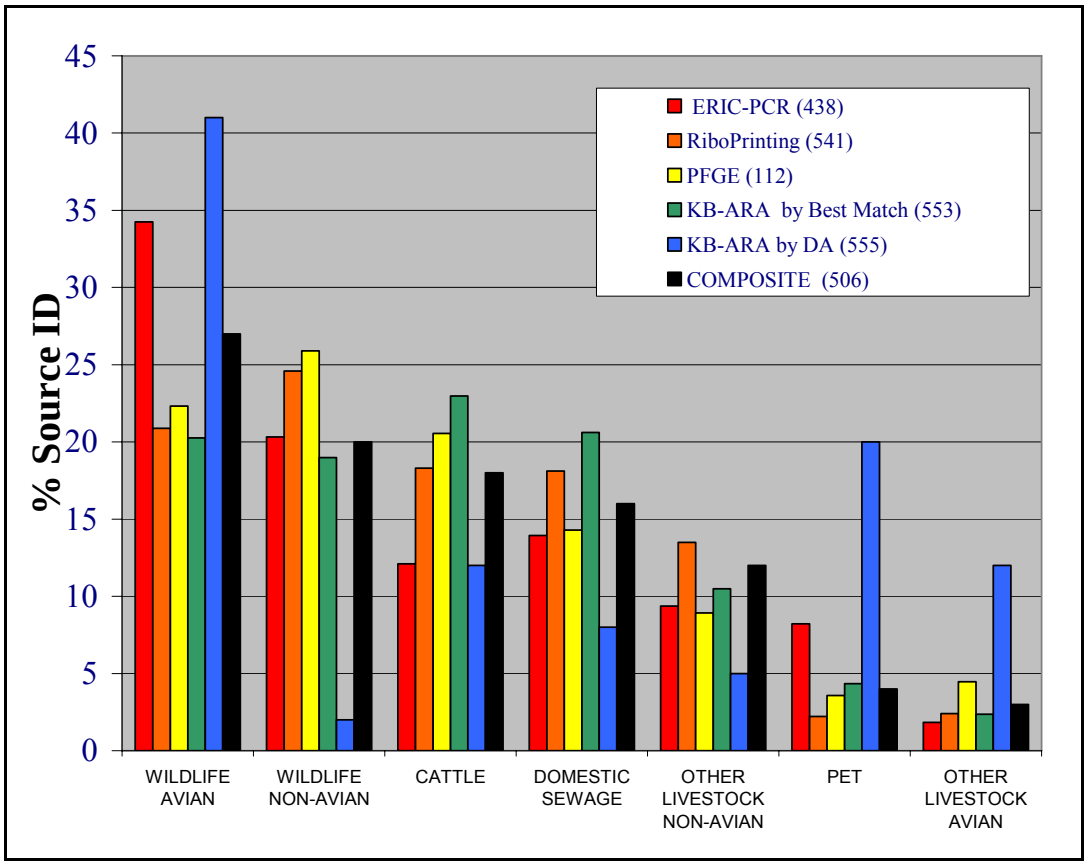
To assess the ability of each BST method and the four-method composite to identify unknown water isolates, a comparison was made between the percent identification to source class for water isolates from both Lakes Waco and Belton Lake watersheds (Figure 9-1). ERIC-PCR, RiboPrinting, PFGE, and KB-ARA (by best match) all identified wildlife, cattle and domestic sewage as the leading sources of contamination. As expected, the four-method composite identified the same major sources of contamination, since it reflects an average of the similarity matrices of the individual methods. The KB-ARA data analyzed using discriminant analysis results also agreed that wildlife is the leading source of contamination. However, identification of other sources of contamination based on discriminant analysis of the KB-ARA data differed from the three molecular methods, as well as the best match analysis of the KB-ARA data. Since best match analysis of KB-ARA data had higher or similar RCCs than discriminant analysis of the data for the potential major sources of pollution (*i.e.,* wildlife, cattle, and domestic sewage), best match analyses of these data were deemed acceptable. KB-ARA data were included in the four-method composite data set and were analyzed using best match.

It should be noted that PFGE patterns could not be generated for approximately 10 percent of the known library isolates and water isolates, limiting the number of library isolates to 883, and the number of water isolates for possible identification to 555. The water isolates that did not generate PFGE fingerprints were analyzed with a three-method composite data analysis, and had a similar distribution of source class identifications of water isolates as the four-method composite.

The four-method composite performed better than any single method for the QC study, had some of the highest RCCs in the seven-way split of source classes, and appeared representative of the identification of water isolates as compared to the individual methods. Given these findings, the identification of water isolates was based on the four-method composite data set.

Percent source identifications in Figure 9-1 are based on the number of isolates identified to each source class out of the number of total identification attempts (shown in parentheses for each method). Isolates that were unidentified because they did not match a library isolate at the minimum similarity are not shown.

Figure 9-1 Identification of Water Isolates from the Lakes Waco and Belton Watersheds using the Individual Methods and Four-Method Composite (555 water isolates vs. 883 known source isolates)



9.2 INTERPRETATION OF COMPOSITE BST RESULTS AND DISCUSSION

To facilitate the interpretation of results, it is helpful to have some understanding of the uncertainties involved in the source estimates. As the number of isolates identified increases, so does the confidence in the estimate of the source contribution. At individual sampling sites, the sources of only a few dozen *E. coli* isolates were identified, so there is a low confidence in the precision of source estimates for individual sites. This is especially true at some sites where *E. coli* were seldom observed. Table 9-1 illustrates the 95 percent confidence intervals for several different contributing source strengths and number of isolates typed. Ninety-five percent confidence implies that 95 times out of 100, the true result is expected to fall within the range indicated. Note that the 95 percent confidence interval estimates are very broad for source identifications based on 100 or fewer *E. coli*. Thus, results for individual monitoring locations must be viewed as rough approximations only. All figures in this section display the number of isolates on which the source contribution estimates are based, in order to facilitate estimation of the uncertainties involved.

Table 9-1 Example 95% Confidence Intervals for Various Contributing Source Magnitudes and Number of *E. coli* Isolates

Number of Isolates Identified	Source Contribution	Source Contribution 95% Confidence Interval
25	3%	0-10%
25	10%	0-22%
25	30%	12-48%
50	3%	0-8%
50	10%	2-18%
50	30%	17-43%
100	3%	0-6%
100	10%	4-16%
100	30%	21-39%
200	3%	1-5%
200	10%	6-14%
200	30%	24-36%
400	3%	1-5%
400	10%	7-13%
400	30%	26-34%

Summaries of sources identified based on the four-method composite at each individual sampling site and for each watershed are provided in Table 9-2. There was a wide variety of contributing sources at each site, and no single source category dominated. Considering all sites and dates combined (Figure 9-2), wild birds were identified as the contributors of 25 percent of the *E. coli* to water, followed by non-avian wildlife at 18 percent of the *E. coli*. Thus, wildlife contributed 43 percent of the *E. coli* from all sites combined. Livestock contributed 31 percent of the *E. coli*, with cattle contributing 17 percent, other livestock - avian at 3 percent, and other non-avian livestock at 11 percent. Sewage was identified as the source of an estimated 14 percent of the *E. coli*. Pets (dogs and cats) were estimated to contribute approximately 3 percent of the *E. coli*. The source of 9 percent of the *E. coli* could not be identified with acceptable confidence.

In a comparison of the source contributions identified in the single sampling event for which a strong runoff influence was questionable (November) versus those from all other sampling events, the source contributions were similar (Figure 9-3). *E. coli* from cattle were less abundant and pets more abundant, in a relative sense, in the runoff samples, but the differences were not statistically significant. It should be noted that identifying sources specifically under runoff conditions was not a goal of this project. Instead, the goal was to sample on a regular basis to capture various climatic conditions at their natural frequency in order to best represent the typical sources of *E. coli*.

The sources of 74 *E. coli* isolates did not generate a PFGE fingerprint and were identified with a three-method composite analysis are provided in Table 9-3. The sources identified are similar to those of the four-method composite analysis provided in Table 9-2.

Table 9-2 Sources of *E. coli* Isolates Based on Four-Method Composite Identification

Site	Station ID	Sewage/ Human	Livestock – cattle	Other Livestock - avian	Other Livestock - nonavian	Pets	Wildlife-avian	Wildlife-nonavian	Unidentified	Total *E. coli* Isolates
Lake Waco near dam	11942	27%	15%	3%	9%	4%	18%	18%	6%	67
Lake Waco N. Bosque River Arm	11945	14%	11%	5%	14%	0%	23%	14%	19%	57
Lake Waco Middle/South Bosque River Arm above SH 6	11948	16%	18%	4%	2%	0%	16%	18%	27%	45
Lake Waco Middle/South Bosque Arm near inlet of Middle/South Bosque River	TBD3	22%	24%	2%	9%	6%	22%	11%	4%	54
North Bosque River at SH 56 near Valley Mills	11953	8%	13%	2%	8%	8%	38%	21%	4%	53
North Bosque River at FM 219 northeast of Clifton	11956	13%	17%	3%	14%	3%	24%	18%	10%	72
Belton Lake near dam	11921	7%	7%	11%	4%	7%	26%	26%	11%	27
Belton Lake Cowhouse Creek arm	11922	17%	21%	0%	8%	8%	13%	17%	17%	24
Belton Lake Leon River arm near headwater	11923	14%	25%	2%	7%	0%	32%	20%	0%	44
Belton Lake Owl Creek arm	TBD4	17%	24%	0%	17%	0%	17%	26%	0%	42
Leon River at FM 1829 southeast of North Fort Hood	11925	4%	11%	0%	20%	4%	36%	19%	6%	70
Belton Lake watershed total		11%	17%	2%	13%	3%	28%	21%	5%	207
Lake Waco watershed total		17%	16%	3%	10%	3%	23%	17%	11%	348
Grand Total		14%	17%	3%	11%	3%	25%	18%	9%	555

Note that the precision of source identification estimates increases with the increasing number of E. coli isolates, and those for individual sites have broad 95 percent confidence intervals.

Table 9-3 Sources of 74 *E. coli* Isolates Based on a Three-Method Composite Identification (No PFGE Data Available)*

Site	Sewage/ Human	Livestock – cattle	Other Livestock - avian	Other Livestock - nonavian	Pets	Wildlife-avian	Wildlife-nonavian	Unidentified	Total *E. coli* Isolates
Belton Lake watershed total	5%	14%	10%	10%	0%	33%	29%	0%	21
Lake Waco watershed total	9%	4%	4%	11%	4%	36%	32%	0%	53

**Results based on composite data set using ERIC-PCR, RiboPrinting, and KB-ARA*

Figure 9-2 Source Contribution Estimates using Four-Method Composite Data for all Sites and Dates Combined (555 *E. coli* Isolates)

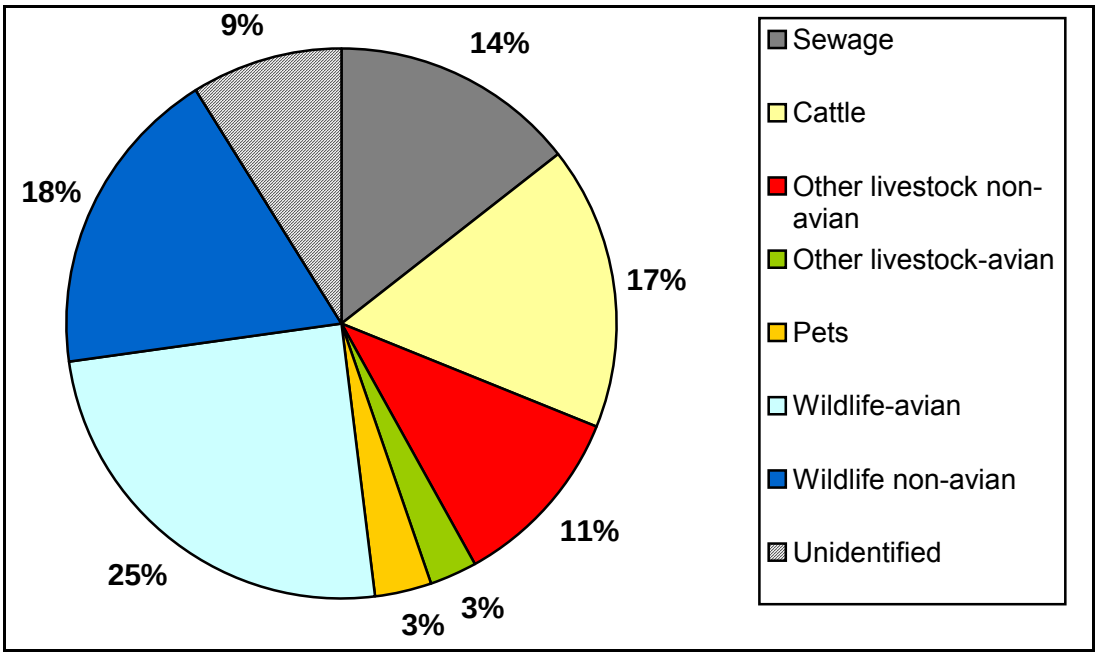
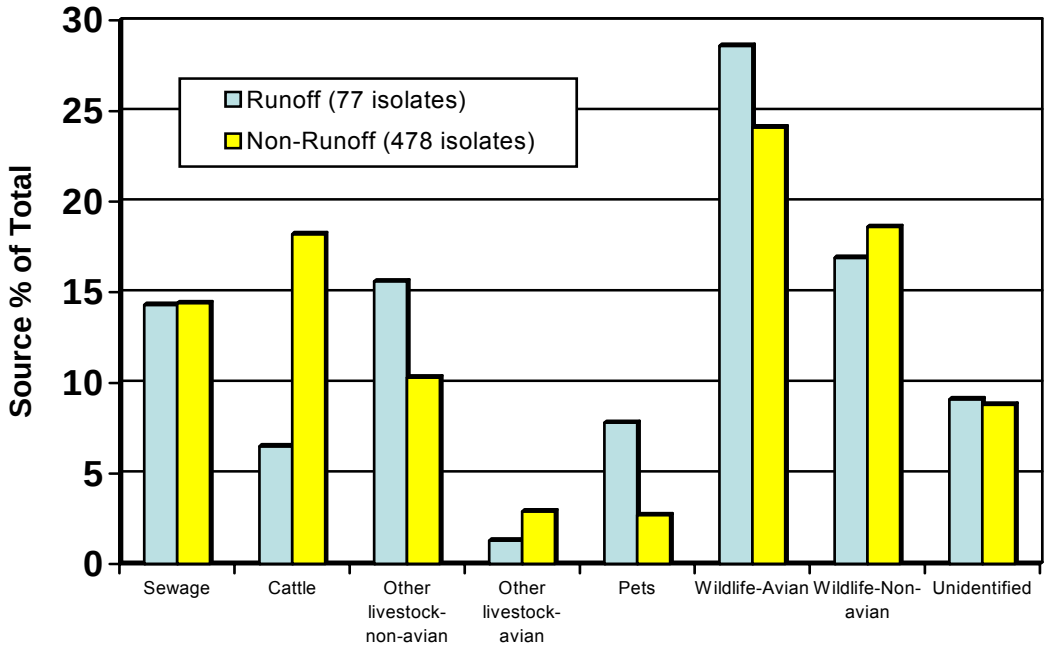


Figure 9-3 Comparison of Source Contribution Estimates for Runoff Samples (November) from all Sites vs. all Other Dates Combined Using Four-Method Composite Data



9.2.1 Aggregate Results Comparison between Lake Waco and Belton Lake

The sources of *E. coli* in Lake Waco and Belton Lake were similar (Figures 9-4 and 9-5). Figure 9-6 displays the same data in a different format by providing a side-by-side comparison of source contribution by category. Sewage contributed a larger percentage of *E. coli* in the Lake Waco watershed, and wildlife represented a larger contribution in the Belton Lake watershed, but the differences are not statistically significant at the 95 percent confidence level.

Figures 9-7 and 9-8 show the identified source contributions by month for Lake Waco and Belton Lake, respectively. The source contributions did not vary significantly with time in a systematic and statistically significant way.

Figure 9-4 Source Contribution Estimates using Four-Method Composite Data for all Lake Waco and North Bosque River Sites and all Dates Combined (348 *E. coli* Isolates)

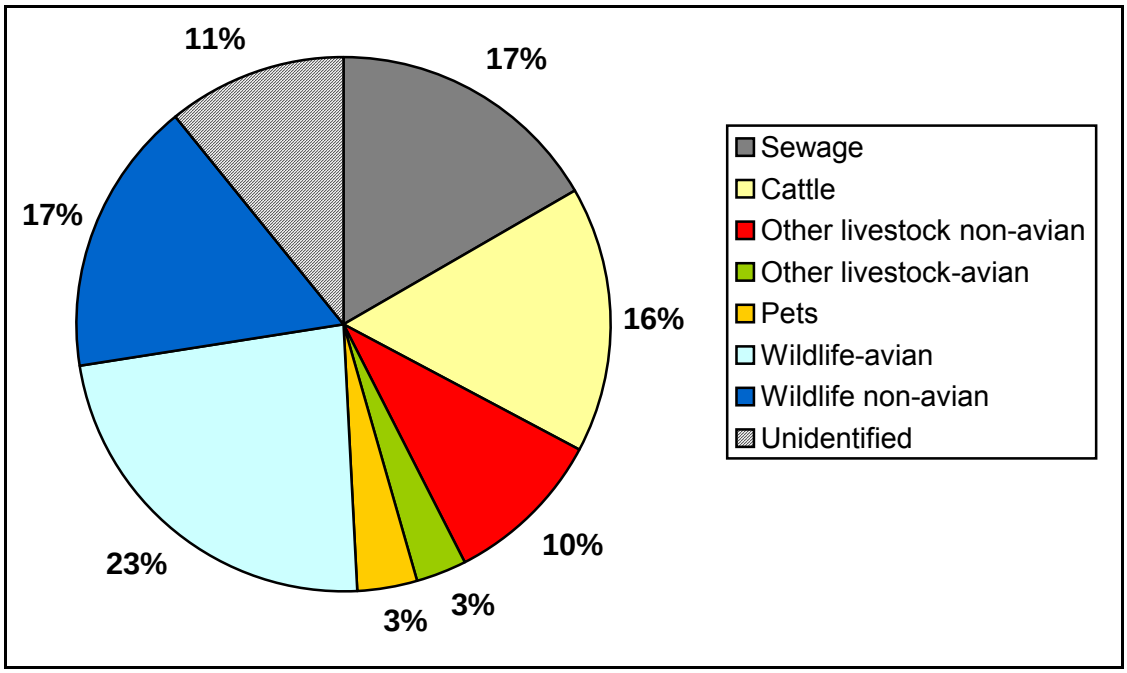


Figure 9-5 Source Contribution Estimates using Four-Method Composite Data for all Belton Lake and Leon River Sites and all Dates Combined (207 *E. coli* Isolates)

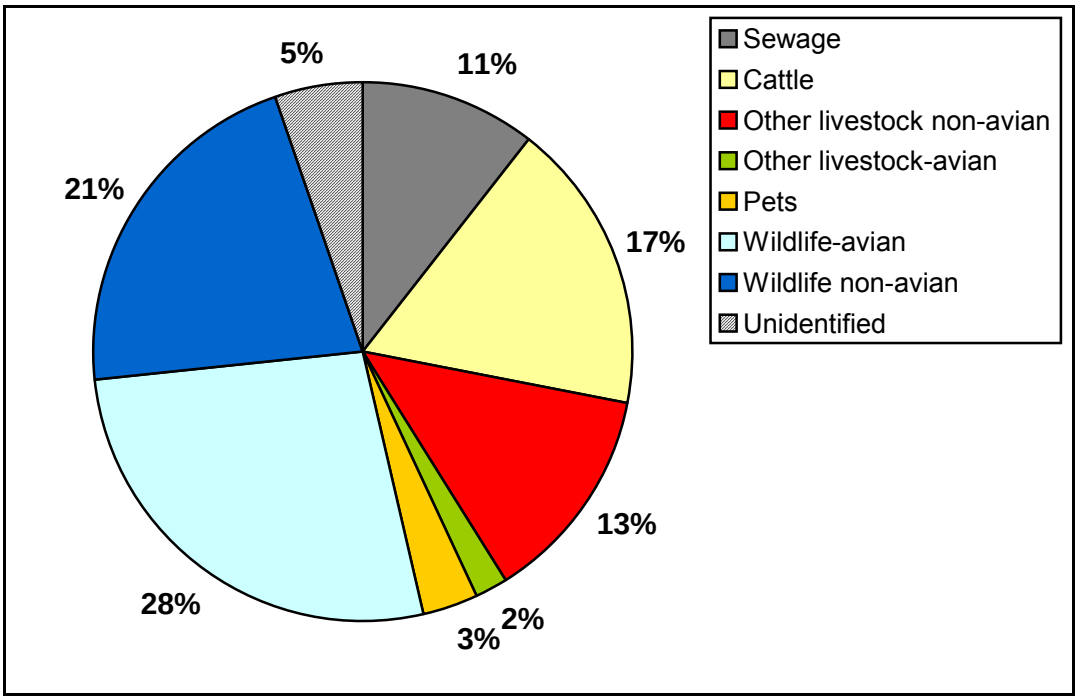


Figure 9-6 Comparison of Source Contribution Estimates for Belton Lake vs. Lake Waco using Four-Method Composite Data

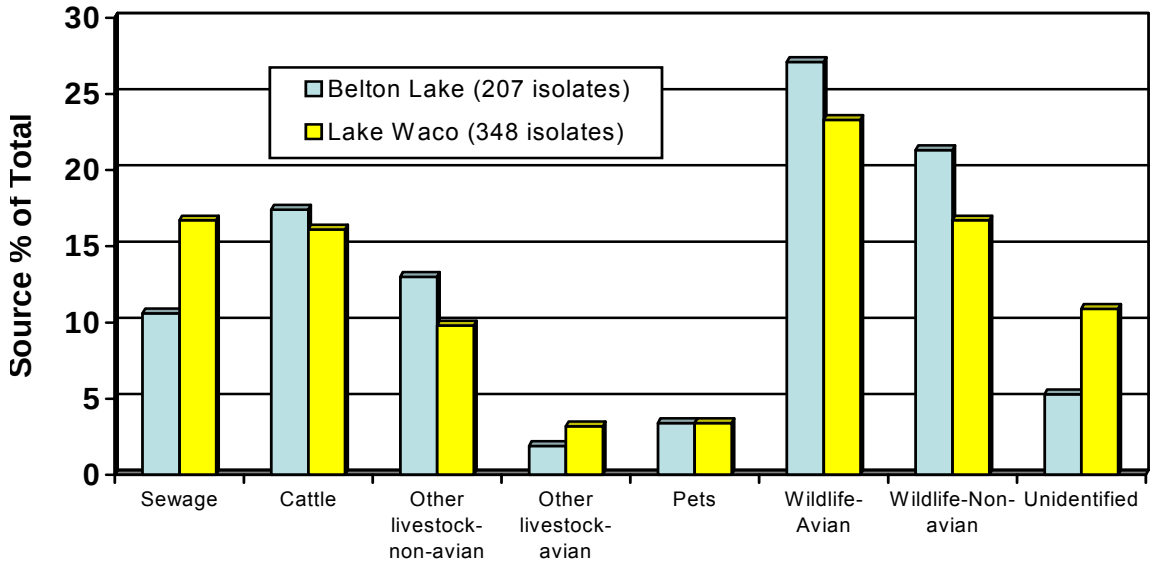


Figure 9-7 Comparison of Source Contribution Estimates by Month using Four-Method Composite Data for all Lake Waco / North Bosque River Sites Combined (348 *E. coli* Isolates)

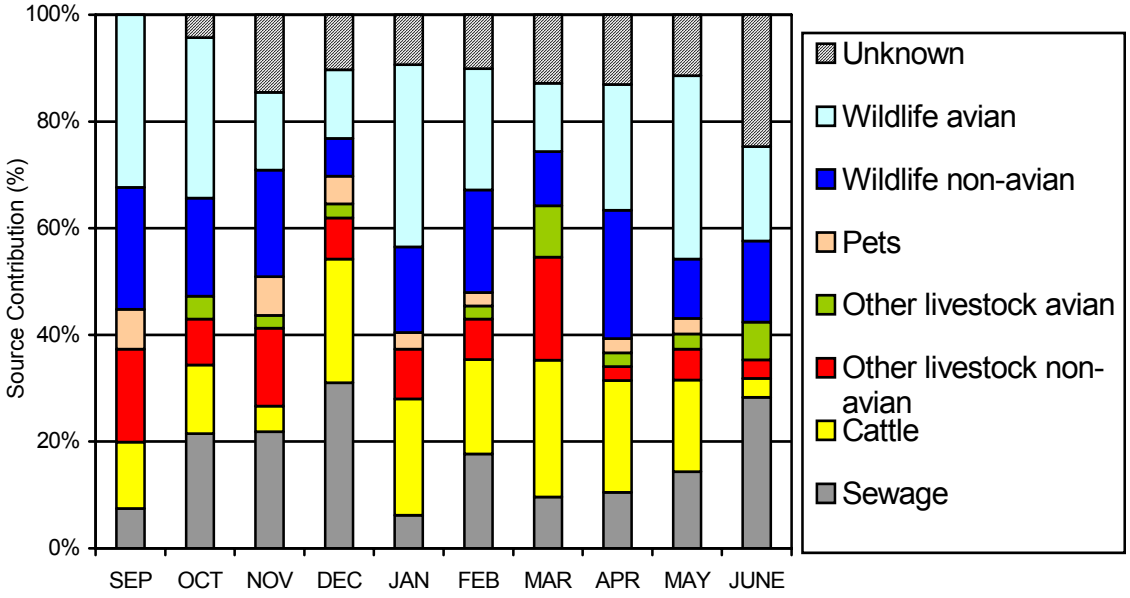
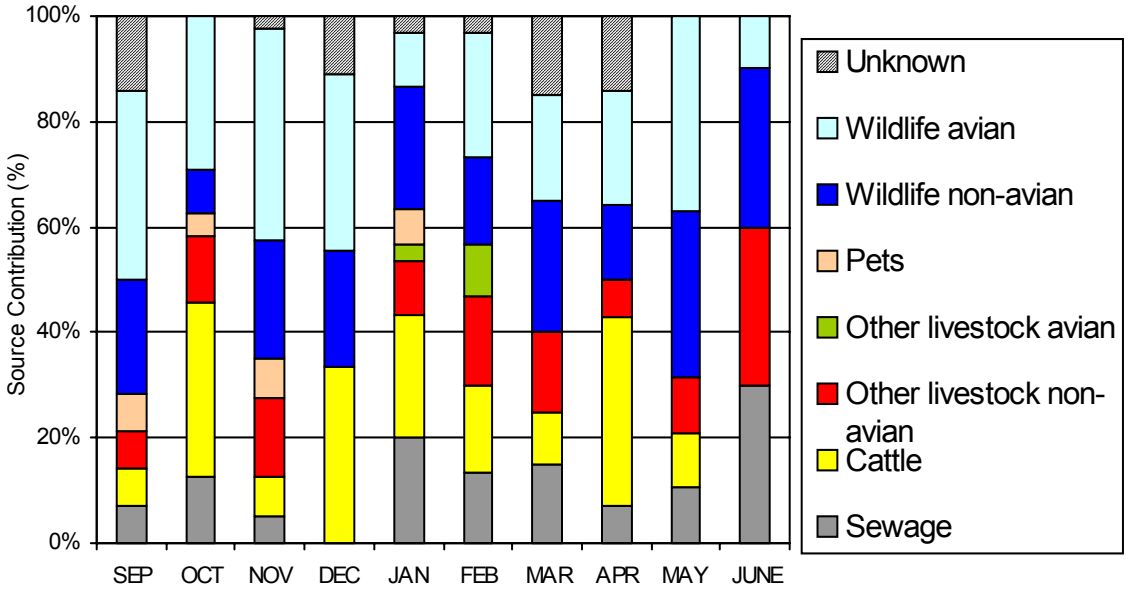


Figure 9-8 Comparison of Source Contribution Estimates by Month using Four-Method Composite Data for all Belton Lake / Leon River Sites Combined (207 *E. coli* Isolates)



9.2.2 BST Classification of Sources for Individual Monitoring Stations

Pie charts representing the source contributions for the 11 water quality monitoring stations associated with Belton Lake and Lake Waco are presented below in Figures 9-9 through 9-19. Although source contribution estimates for individual sites are too imprecise for statistically significant comparisons, the highest sewage contributions were seen in Lake Waco near the dam, a site near heavily populated areas of the City of Waco. The highest contributions from cattle were observed in the Leon River arm of Belton Lake and the Middle/South Bosque River arm of Lake Waco. The largest contributions from other livestock - avian were observed in Belton Lake near the dam, and the highest contributions from other non-avian livestock were observed in the Leon River near north Fort Hood. The highest avian wildlife contributions were observed in the north Bosque River near Valley Mills, perhaps due to a large swallow population observed to be nesting under the bridge. Non-avian wildlife contributions were most substantial in the Owl Creek arm of Belton Lake and near the Belton Lake dam. The highest pet contributions were observed in the North Bosque River near Valley Mills and in the Cowhouse Creek arm of Belton Lake. Wildlife (non-avian wildlife and avian wildlife combined) contributions represented the largest source percentage at all 11 stations and ranged from 30 to 59 percent. Only four of the 11 stations had more than 10 percent of the isolates that were not identifiable with Lake Waco Middle/South Bosque River arm at SH 6 having the highest percentage at 26.

Figure 9-9 Source Contribution Estimates using Four-Method Composite Data for Belton Lake near Dam (11921) - all Dates Combined (27 *E. coli* Isolates)

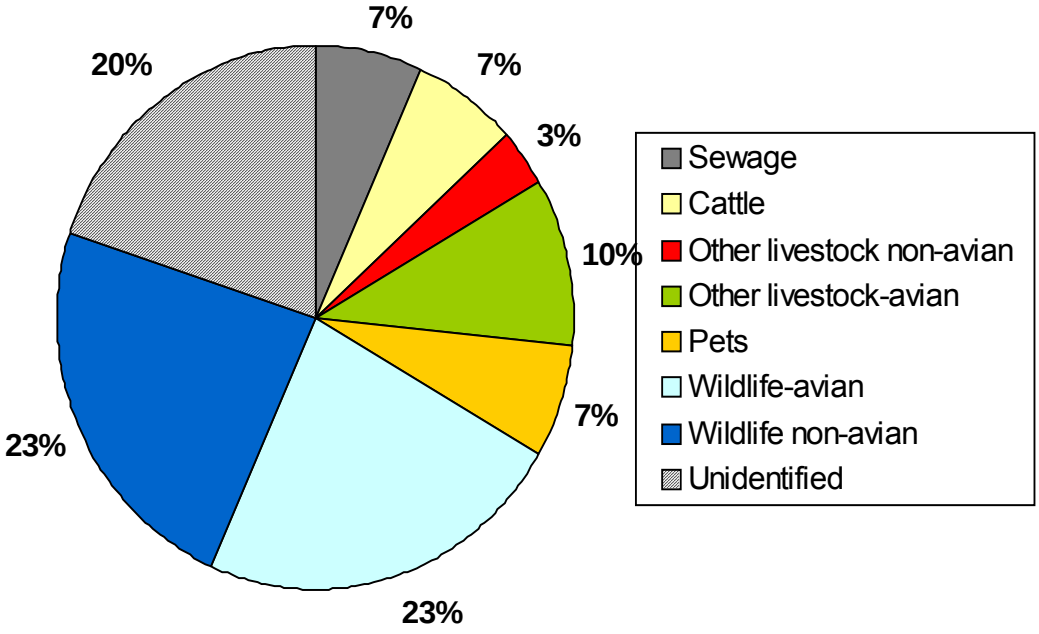


Figure 9-10 Source Contribution Estimates using Four-Method Composite Data for Belton Lake Cowhouse Creek Arm (11922) - all Dates Combined (24 *E. coli* Isolates)

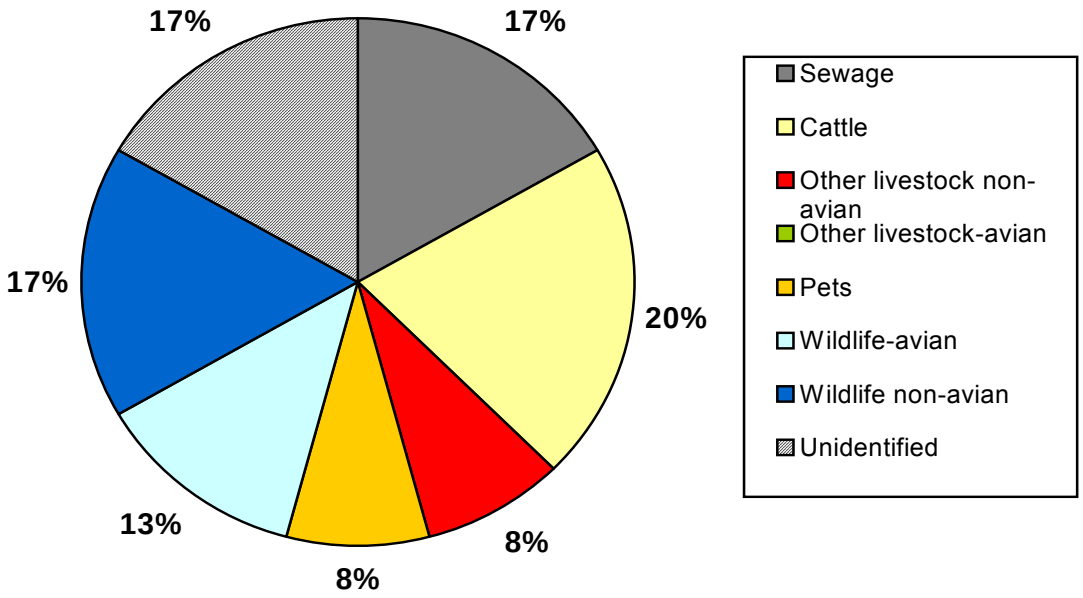


Figure 9-11 Source Contribution Estimates using Four-Method Composite Data for Belton Lake Owl Creek Arm (TBD4) - all Dates Combined (42 *E. coli* Isolates)

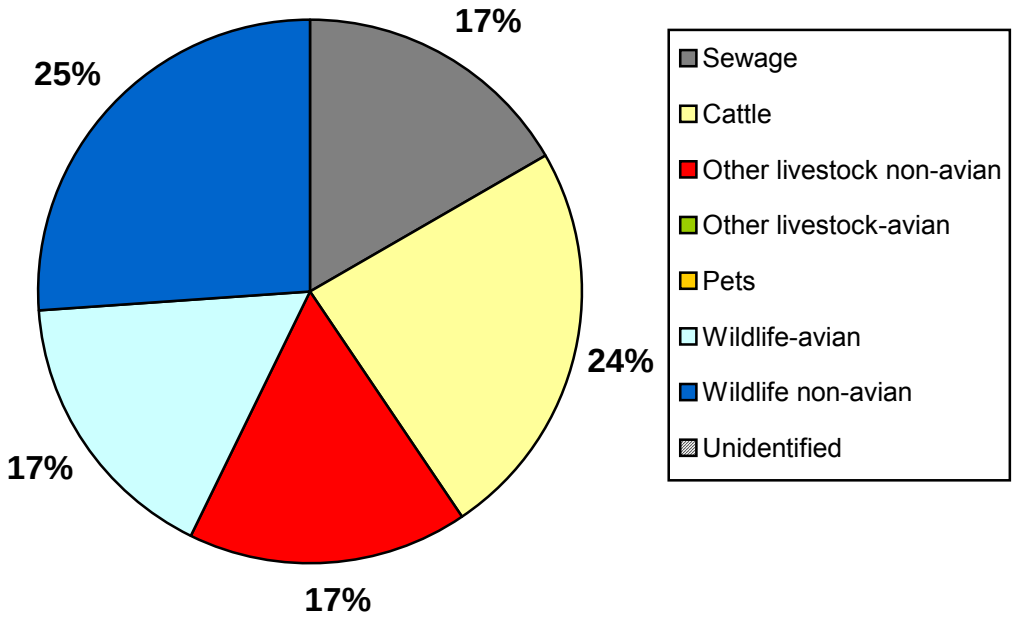


Figure 9-12 Source Contribution Estimates using Four-Method Composite Data for Belton Lake Leon River Arm near Headwater (11923) - all Dates Combined (44 *E. coli* Isolates)

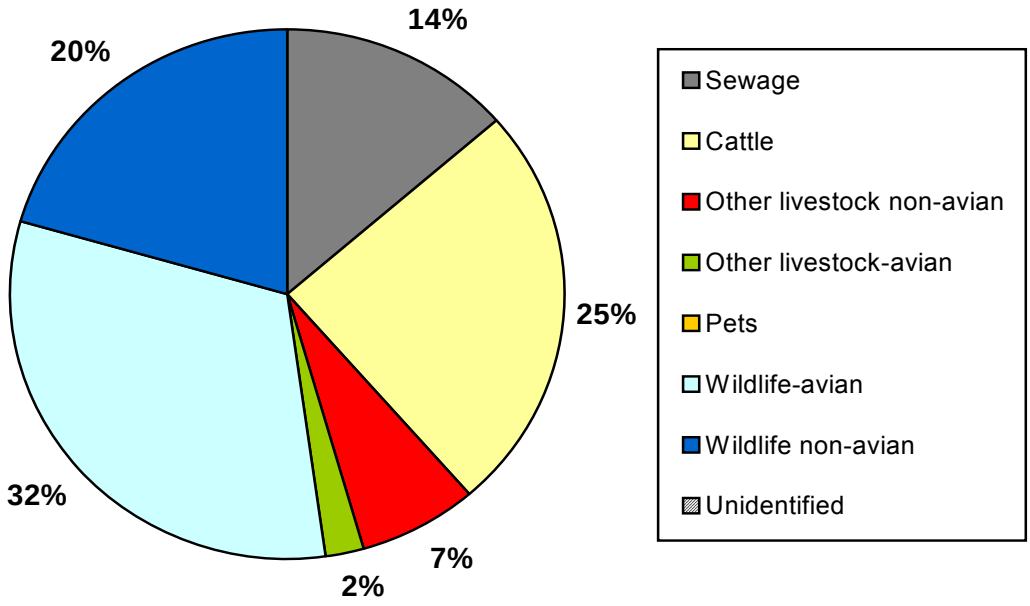


Figure 9-13 Source Contribution Estimates using Four-Method Composite Data for Leon River at FM 1829 southeast of North Fort Hood (11925) - all Dates Combined (70 *E. coli* Isolates)

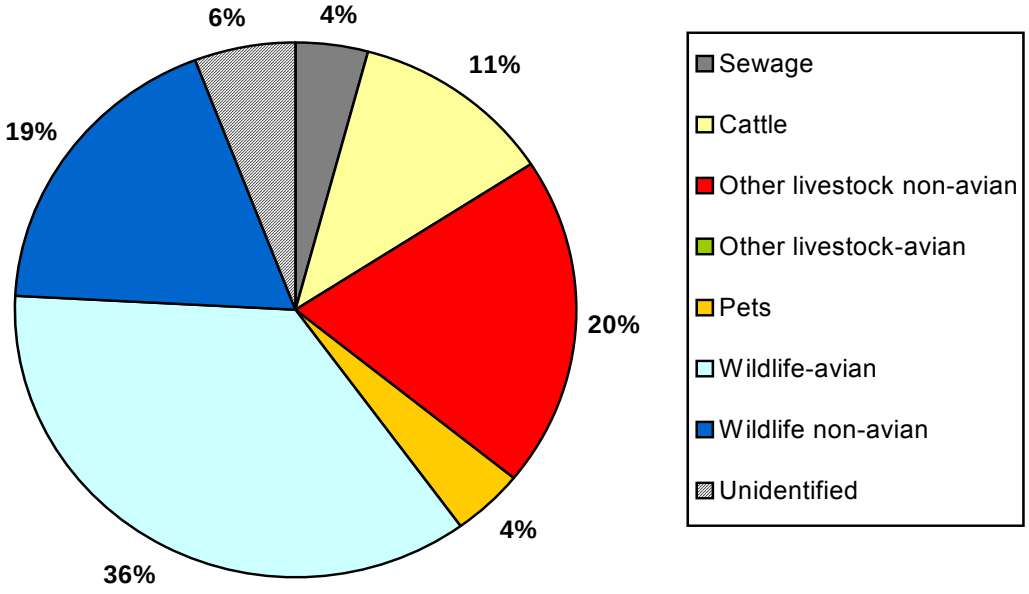


Figure 9-14 Source Contribution Estimates using Four-Method Composite Data for Lake Waco near Dam (11942) - all Dates Combined (67 *E. coli* Isolates)

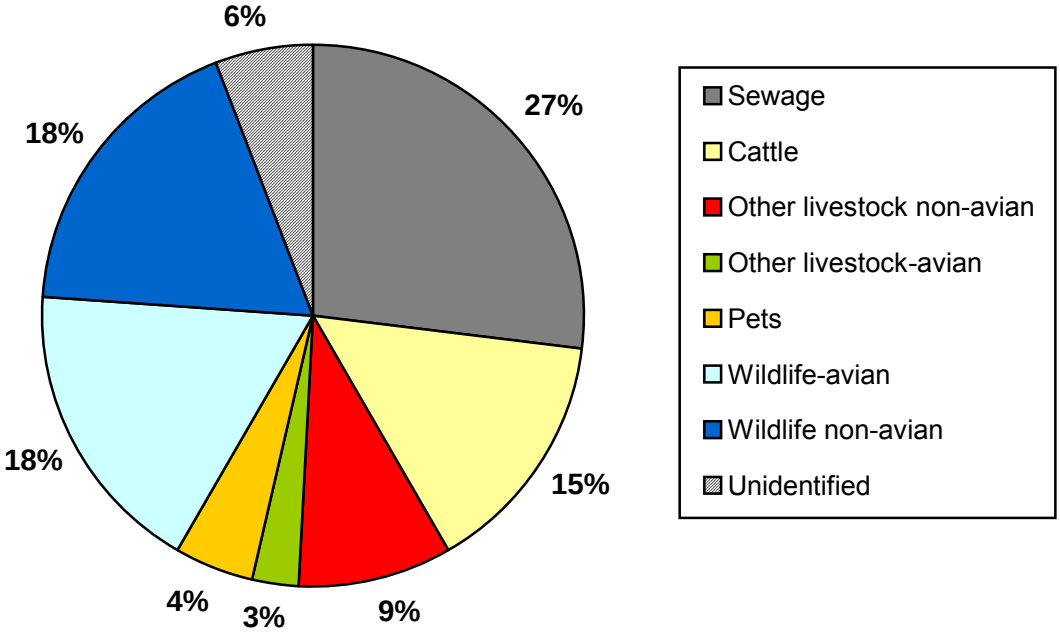


Figure 9-15 Source Contribution Estimates using Four-Method Composite Data for Lake Waco Middle/South Bosque River Arm above SH 6 (11948) - all Dates Combined (45 *E. coli* Isolates)

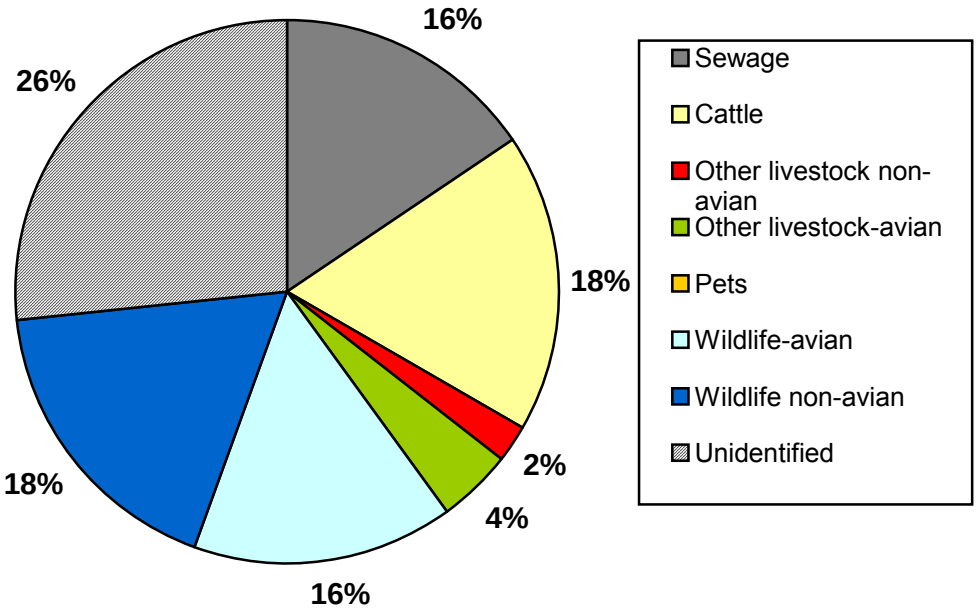


Figure 9-16 Source Contribution Estimates using Four-Method Composite Data for Lake Waco Middle/South Bosque River Arm near Inlet of Middle/South Bosque Rivers (TBD3) - all Dates Combined (54 *E. coli* Isolates)

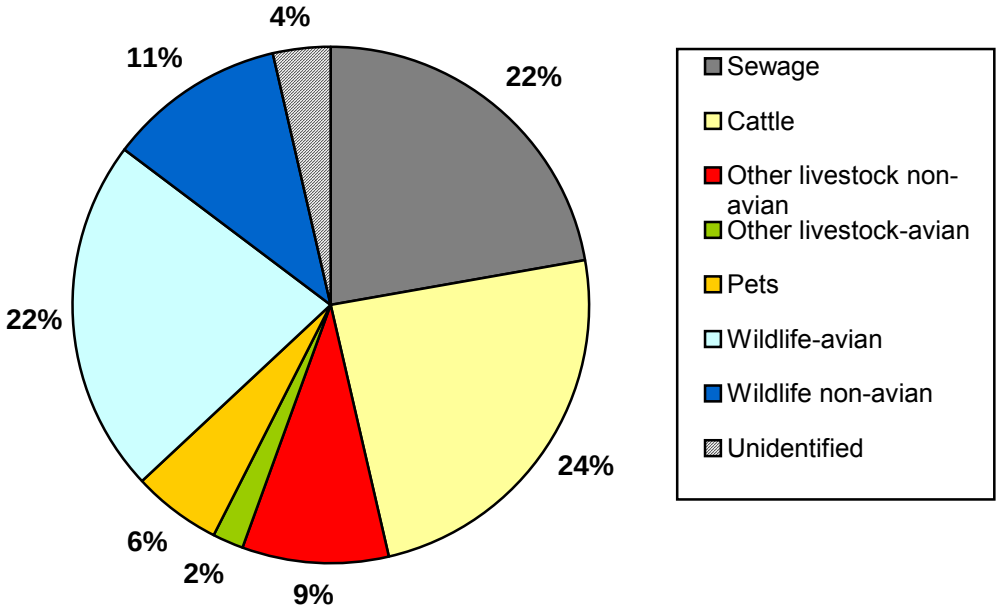


Figure 9-17 Source Contribution Estimates using Four-Method Composite Data for Lake Waco North Bosque River Arm (11945) - all Dates Combined (57 *E. coli* Isolates)

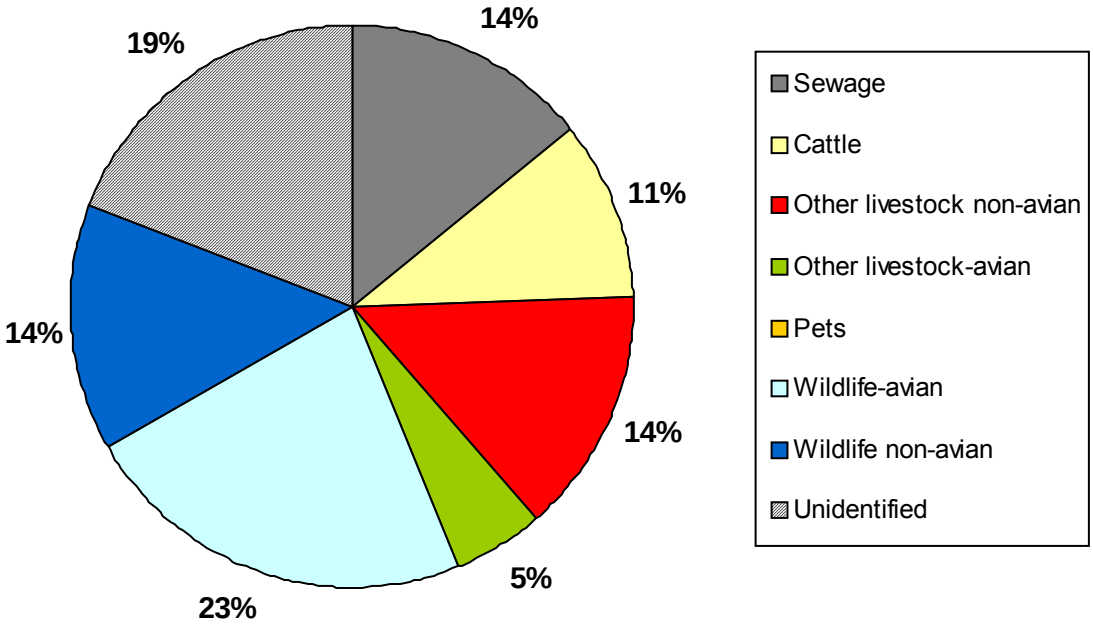


Figure 9-18 Source Contribution Estimates using Four-Method Composite Data for North Bosque River at SH 56 near Valley Mills (11953) - all Dates Combined (53 *E. coli* Isolates)

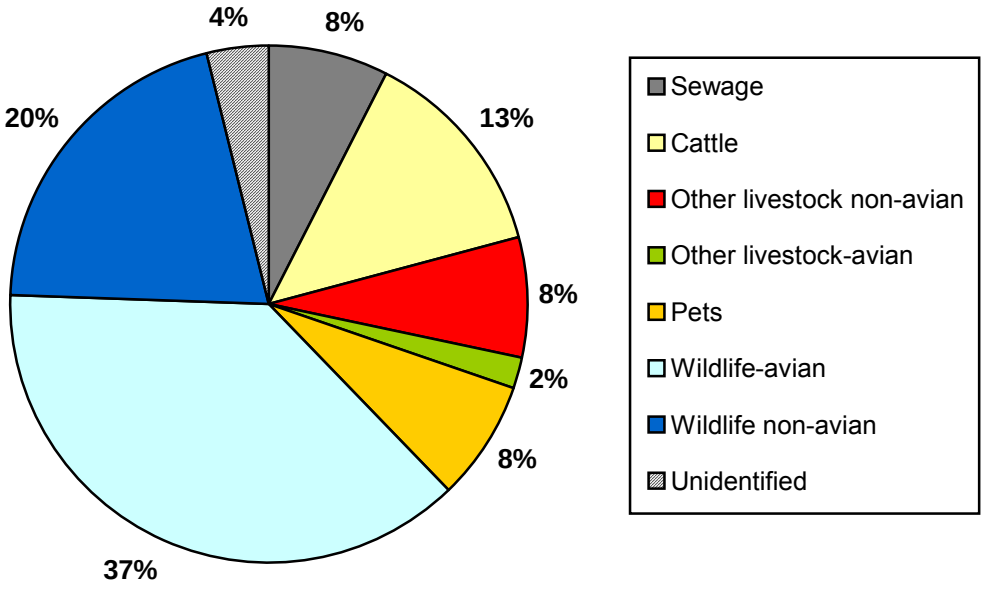
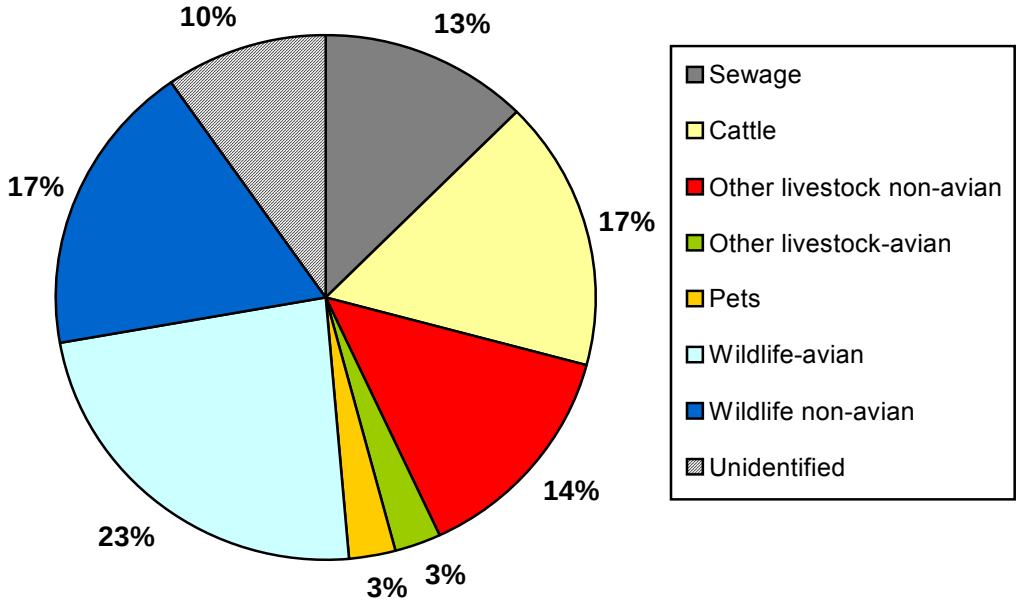


Figure 9-19 Source Contribution Estimates using Four-Method Composite Data for North Bosque River at FM 219 northeast of Clifton (11956) - all Dates Combined (72 *E. coli* Isolates)



THIS PAGE INTENTIONALLY LEFT BLANK

SECTION 10
STAKEHOLDER INVOLVEMENT

10.1 WORKSHOPS

Coordination among the project team members supporting this project was a high priority throughout the study period. The project team also recognized that communication with and feedback from stakeholders in each watershed would strengthen the project design and outcomes.

This project included two 3-hour stakeholder meetings or workshops, one at the beginning of the project in July 2003, and the second in October 2005. Both workshops were held at the TFB headquarters in Waco, Texas, and both were well attended with more than 40 people at each meeting. The lists of stakeholders invited to each workshop are provided in Appendix K. The objectives of the first stakeholder workshop were to:

- Introduce the project team and summarize the project work plan;
- Define what the project was and was not intended to do;
- Provide an introduction on how BST methods provide data necessary to identify specific sources of bacteria; and
- Obtain local knowledge about geographic origin of *E. coli* sources.

During the first workshop in July 2003, Parsons received a number of beneficial suggestions on potential bacteria sources in both watersheds which were taken into account when conducting the sanitary survey and developing the sampling design. The objectives of the second workshop were to:

- Summarize the Lake Waco/Belton Lake BST sampling design;
- Describe laboratory approaches used to analyze and classify *E. coli* bacteria; and
- Provide a summary of the data results that categorize the contributing sources of *E. coli* to Lake Waco and Belton Lake.

10.2 SUMMARY OF QUESTIONS AND RESPONSES FROM WORKSHOP 2

There were a number of questions posed by the stakeholder group at the second workshop in October 2005. Several of these questions are paraphrased below and a brief response follows each question.

1. Are the water quality sampling results transferable to TCEQ for inclusion in the TCEQ surface water quality monitoring database?

   Yes. The data results can be transferred to TCEQ since the data were collected using a TCEQ-approved QAPP. The data will be submitted for inclusion in the TCEQ water quality database.

2. How many of the 10 scheduled sampling events were runoff events?

   Three of the 10 events (November, January, and March) included water samples collected after rainfall or under increased flow conditions.

3. Was the project structured to identify instream bacteria from aquatic species?

   No. The nature of this project was to focus on identifying bacteria sources from warm-blooded species since they have the potential to cause human health concerns. Aquatic species in Texas streams are cold-blood and it is therefore assumed that bacteria from these species would not present public health concerns.

4. With the data collected can estimates be given for the percent contribution of different sources for subwatersheds of Lake Waco and Belton Lake?

   The data collected cannot provide sufficient resolution to quantify bacteria sources for different subwatersheds upstream of each lake. The sampling design was not set up to do this and the cost of providing that level of resolution was beyond the resources of this project.

5. Can bacteria from dairy cattle and beef cattle be differentiated using the data from this study?

   No. The E. coli fingerprints from beef and dairy cattle gut do not appear to be distinct from each other. However, it is possible to obtain more site-specific data to assist BST results in differentiating sources at the subwatershed level.

6. Were bacteria from feral hogs and domestic swine differentiated using the data from this study?

   There was little cross-identification between wildlife non-avian (the source category for feral hog and javelina) and other livestock non-avian (the source category for swine), suggesting that these sources can be differentiated. There were a total of 32 isolates analyzed from these sources with all four BST methods. An evaluation of the data for these isolates suggests that E. coli from feral hogs and javelina can be distinguished from E. coli from swine; however, the total number of isolates is relatively small, and it is unknown if this will continue to be the case as the library is expanded.

7. Why were there especially high levels of unidentified isolates in the one sampling station (Lake Waco Middle/South Bosque arm above SH 6)?

   The percentages can appear high but they may only represent a few (in this case 12) isolates. In other words, the higher the number of isolates used in an analysis, the lower the percentage of unidentified isolates. An allowance will be made for unidentified isolates in the analysis instead of trying to force a match to a fingerprint. It is possible that the source isolate was not available in the library, or the animal was represented, but this was a unique isolate not found in the library.

8. You used a composite of the four BST methods; has this been done before?

   To date, no study of this type has been published in a peer-reviewed journal, although other research groups are now looking into this type of approach. Comparing the four BST methods allows all the data gathered to be used and appears to balance out the extremes in method resolution.

9. Will the report discuss how the methods compare and what is best to use when you cannot afford (time, money, or expertise) to do all four methods as a composite?

   Yes. Comparison of the different method combinations is still being done, and will be included in the final report. Method comparisons and congruence measurements show that a combination of ERIC-PCR and RiboPrinting may work just as well as the four-method composite.

10. You mentioned there were many swallows under the bridge at one of the sampling sites, and that a majority of those isolates from that water segment were identified as wildlife avian. Do you try to take other water samples up or down stream in the vicinity when you see something like that?

    The general procedure is to try to take the samples 10-20 yards upstream of the bridge to minimize the effects of the bridge on water quality measurements, but 1) going farther upstream or downstream would have required trespassing on private land, and 2) the QAPP mandates where samples will be taken, and deviations from that plan cannot be made.

11. How different are the livestock categories?

    Library isolates were divided into different source categories based on management practices. The livestock was grouped into 1) cattle, 2) other livestock non-avian, and 3) other livestock avian. However, the resolution of the BST techniques may not be sufficient to distinguish the other livestock non-avian and avian categories as evidenced by the levels of cross-identification. Also, the number of other livestock avian library isolates was small and may not have been sufficient for source class evaluation of the library.

12. Will the differences in the number of wildlife and human isolates used in the library impact the identification of the unknowns (the influence of the number of isolates over the overall analysis of the samples)?

    More important than the number of isolates in the library from a given source is the representativeness of the isolates. It is not known how large an optimum library needs to be. The library was designed to be diverse by limiting the clones (identical isolates from the same sample) and choosing unique isolates from known source categories to represent the diversity of E. coli found in the feces of the different hosts. Humans and wildlife, in general, seem well-represented as seen by the library analysis. A larger library has a better chance of identifying unknown water isolates, but because the analysis done in this study allows for "unidentified" isolates (instead of force grouping) that should limit false identifications. Statistical methods to evaluate the library are still being explored, and those results will be included in the final report.

THIS PAGE INTENTIONALLY LEFT BLANK

SECTION 11
RECOMMENDATIONS FOR FUTURE BST STUDIES

11.1 BST PROJECT DESIGN

Well defined DQOs are an important first step with any water quality assessment project including BST projects. This section identifies key issues and recommendations that should be acknowledged when considering undertaking a BST project. However some of the key issues which affect DQOs and the sampling design of BST projects that must be given special attention are:

- budget,
- discriminatory capability desired of BST methods,
- size and diversity of known source library, and
- long-term maintenance of library isolates.

Library size and the representativeness of *E coli* strains in a known source library are two major considerations that need to be carefully assessed before embarking on any BST study (USEPA 2005). While there are an array of factors that influence the appropriate size of a known source library, (size of watershed, number and diversity of sources, migration of animal species, etc.) Texas' recent completion of BST projects in different parts of the state has provided a valuable start to a useful library.

Regarding the costs of BST projects, there are too many variables with each project to provide typical ranges one could use in setting a budget for the use of BST. Library dependent methods such as RiboPrinting and ERIC-PCR do require significant funding but the discriminatory capability of these methods can now provide data of greater value to water resource managers. It should be recognized that the investment in infrastructure by the state and USEPA needed to make this project possible has value to future BST projects in the state. EP AREC and TAMU now have the equipment necessary to perform additional studies and the *E. coli* culture collection obtained through this project is being maintained by EP AREC. The availability of this library will make it possible to reduce the number of source samples that need to be collected in future studies. Further, the use of recently developed library-independent methods as part of the BST "toolbox" may also reduce the sampling requirements. Together, these may allow future BST studies to be conducted at significantly less expense. Parsons, Texas Agricultural Experiment Station and TAMU are committed to assisting the state with future BST projects as funding and resources are made available.

The level of discriminatory capability desired will also drive DQOs. BST projects are typically designed to differentiate:

5) human vs. all other sources,
6) species specific results (human vs. cows vs. horses vs. deer, *etc*.),
7) group comparisons (human vs. livestock vs. wildlife), and

8) specific individual hosts (cows from a certain farm vs. other farms vs. other livestock on farms vs. human etc. (USEPA 2005).

Setting up well defined DQOs guided by one of the four options listed above is a critical step in any project design and has a direct bearing on project costs. An assessment method employed by this study that provided significant utility was the sanitary survey. The sanitary survey provided valuable watershed-specific data that improved the sampling design. All BST projects should adequately invest in conducting a thorough sanitary survey. The sanitary survey can also be useful in determining which level of discriminatory capability is most appropriate to use in a project.

Finally, when considering the use of library dependent methods, special attention should be given to the long-term maintenance of the library isolates for future use and application. Building a regionally representative and diverse library of known source isolates for the state of Texas should be given consideration. Special attention should also be given to ensuring that future BST projects in Texas adhere to a QAPP that will allow the samples (isolates) collected to be used in conjunction with the known source library developed for this project.

11.2 CORRESPONDENCE BETWEEN METHODS

Using the 883 known source isolate library, the congruence between individual BST techniques, all two and three-method combinations, and the four-method composite data set were determined using BioNumerics (Figure 11-1). The dendrogram in Figure 11-1 reveals the relationship between the different techniques and composite data sets, and the percent similarity of each method or combination to the four-method composite is included. As expected, the methods do not agree 100 percent with each other due to the differences in resolution between the methods and the fact that they each measure different attributes of *E. coli*.

It may not be feasible in terms of both cost and time considerations to perform all four BST methods used here for future BST studies. The congruence measurement revealed that the two individual techniques most similar to the four-method composite data set were ERIC-PCR and RiboPrinting, with 77.9 percent and 60.8 percent similarity, respectively. This was also not surprising, since PFGE and KB-ARA are at the extremes of the spectrum for their ability to resolve differences between bacterial isolates. The ERIC-PCR RiboPrinting composite data set was found to be the closest two-method combination (90.7 percent similar) to the four-method composite data set. These results suggest that a combination of just ERIC-PCR and RiboPrinting may be suitable for future library-based BST studies. In addition, additional analysis of the QC study data demonstrated that a composite ERIC-RP data set correctly identified 100% of the replicate QC cultures, and had 90% accuracy for the *E. coli* strain and source class identification of the isolates.

It should also be noted that since the beginning of this study, there have been significant developments in library-independent source tracking methods. While some of these developments appear promising, it is still too early to tell if they will be capable of replacing library-dependent BST methods. One particular weakness of these recently

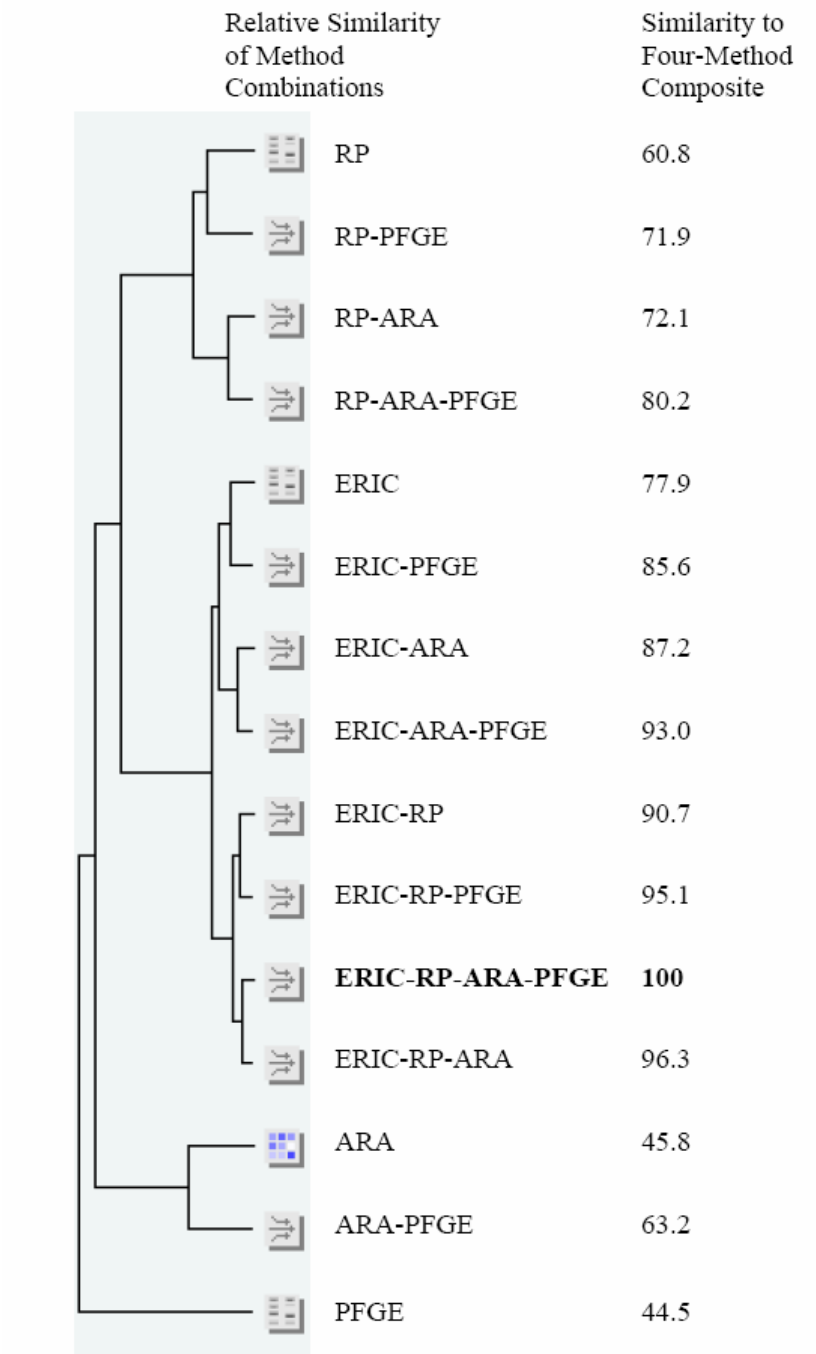
developed approaches is the difficulty of interpreting results in relation to regulatory water quality standards and microbial risk, since they often target microorganisms that are not regulated and that have not been included in prior microbial risk assessments. Nonetheless, for future studies in Texas, it is recommended that these methods be given consideration, not as stand-alone methods, but rather as part of a "toolbox" approach to BST, along with the library-dependent methods.

The use of four different BST methods in this study provided valuable insights to BST methods. This is believed to be one of the first studies to directly compare BST methods by using the same collection of *E. coli* isolates. The methods used covered the spectrum in cost, ease of use, and discriminatory ability. As hypothesized, data from the combined methods (composite data set) were more useful than any individual method. Further, congruence measurements suggest that an ERIC-PCR and RiboPrinting composite data may be as useful as the four methods combined. In future studies, ERIC-PCR and RiboPrinting appear to be the methods of choice since they provide the highest rate of correct classification, and the other methods used to further characterize specific groups of isolates as needed. This would provide for cost, labor, and time savings while not compromising integrity of the BST results.

Key Recommendations

- Special attention should also be given to ensuring that future BST projects in Texas adhere to a TCEQ- or TSSWCB-approved QAPP.
- BST projects should adequately invest in conducting a thorough sanitary survey.
- When considering the use of library dependent methods, special attention should be given to the long-term maintenance of the library isolates for future use and application.
- The combination of ERIC-PCR and RiboPrinting appear to be the most suitable as well as accurate methods for future library-based BST studies.

Figure 11-1 Congruence of Individual BST Methods and Composite Data Sets*



* Labels in Figure 11-1 are as follows: ERIC= ERIC-PCR; RP=RiboPrinting; ARA= KB-ARA; PFGE=PFGE, ERIC-RP-ARA-PFGE = 4-method composite data set; other composites are hyphenated.

SECTION 12
CONCLUSIONS FROM BST RESULTS

The TFB recognized that alternative water quality assessment methods would be required to achieve the objectives of this project. In most Texas watersheds, existing data are insufficient to provide water resource managers with detailed information about which sources of bacteria to target to achieve beneficial load reductions. Reaching beyond the standard suite of techniques used to assess bacteria, the TFB, Parsons, and TAMU worked together to promote and utilize BST methods. As a result, this study provides a wealth of information to state and local agencies in Texas from which future BST studies can be designed. The data now available from a BST study such as this validate the wide array of bacteria sources contributing loading to surface water. While the costs and time requirements associated with the different BST analytical methods used vary, the quality-assured data derived from this study provide water resource managers with the confidence needed to gain support from stakeholders to address specific sources of bacteria. It should be recognized that data from BST studies such as this have limitations to the spatial and temporal questions that can be answered. This study has accomplished the project objectives and adhered to the data quality objectives set forth in the QAPP. The key conclusions derived from this BST project are:

- BST results indicate that wildlife (avian and non-avian) is the major contributor to fecal pollution for both Lake Waco and Belton Lake. Wildlife was identified as the source of 40 percent and 49 percent of the *E. coli* isolated from Lake Waco and Belton Lake water samples, respectively.
- It had been previously speculated that livestock, in particular cattle, and other agricultural activities were responsible for the majority of fecal pollution in these watersheds. Contributions to fecal pollution by cattle as determined by BST are similar for both Lake Waco and Belton Lake watersheds, with 16 percent and 17 percent of the *E. coli* isolates from water identified as cattle source, respectively. Looking at the contributions from cattle at each of the 11 monitoring stations at both lakes, the percentage is less than or equal to 25 percent. Thus, these results demonstrate that cattle are not the major contributing source of bacteria to Lake Waco and Belton Lake. If the cattle, other livestock - avian, and non-avian livestock source classes are combined into a “livestock” class, livestock appear to be the second leading source of fecal pollution, with 29 percent and 32 percent of the water isolates from Lake Waco and Belton Lake, respectively, identified to this source.
- Unexpectedly, domestic sewage is identified as the third leading contributor to fecal pollution in the watersheds, with 17 percent of the Lake Waco and 11 percent of the Belton Lake *E. coli* isolates identified to this source. Of particular concern is the finding that the sampling site with the highest occurrence (27%) of *E. coli* isolates identified as sewage is at Lake Waco near the dam (Station 11942), which is also near the drinking water treatment plant intake. While microorganisms capable of causing disease in humans can be found in animal feces, domestic sewage typically contains much

higher levels of these organisms. Therefore, from a human health perspective, this site may be at greatest risk.

- Levels of *E. coli* were low in both Lake Waco and Belton Lake, and *E. coli* were absent from approximately one third of the samples collected.
- Additional data collection would be necessary to present more definitive conclusions of how stormwater runoff might influence percent contributions from source categories.
- The BST results provide valuable information that will assist water resource managers in targeting future management strategies to address bacteria contributions from specific source categories in each watershed.

SECTION 13
REFERENCES

AMVA 2002. *U.S. Pet Ownership and Demographics Sourcebook (2002 Edition)*. Schaumberg, IL.

Johnson, L.K., Brown, M.B., Carruthers, E.A., Ferguson, J.A., Dombek, P.E., Sadowsky, M.J. 2004. *Sample size, library composition, and genotypic diversity among natural populations of Escherichia coli from different animals influence accuracy of determining sources of fecal pollution*. Appl. Environ. Microbiol. 70(8): 4478-85.

Mott and Lehman 2001. DNA Fingerprinting to Identify Sources of Bacteria in Coastal Waters of Texas, Final Report - Phase II. Coastal Coordination Council, National Oceanic and Atmospheric Administration Award No. NA870Z0251.

Reed, Stowe and Yanke 2001. *Study to Determine the Magnitude of, and Reasons for, Chronically Malfunctioning On-Site Sewage Facility Systems in Texas*. Austin, TX. September 2001.

Stoeckel, D.M., Mathes, M.V., Hyer, K.E., Hagedorn, C., Kator, H., Lukasik, J., O'Brien, T.L., Fenger, T.W., Samadpour, M., Strickler, K.M., Wiggins, B. A.; 2004. *Comparison of seven protocols to identify fecal contamination sources using Escherichia coli*. Environ. Sci. Technol. 38(22): 6109-6117.

TCEQ 2002. An Implementation Plan for Soluble Reactive Phosphorus in the North Bosque River Watershed. Texas Commission on Environmental Quality. December 2002. Austin, Texas.

TCEQ 2004. Guidance for Assessing Texas Surface and Finished Drinking Water Quality Data, 2004. Texas Commission on Environmental Quality. August 2003. Austin, Texas.

TCEQ 2005a. Draft 2004 Texas Water Quality Inventory and 303(d) List. http://www.tceq.state.tx.us/compliance/monitoring/water/quality/data/wqm/305_303.html#2004 May 13, 2005.

TCEQ 2005b. Texas Water Utility Database. http://www3.tnrcc.state.tx.us/iwud/ .

TCEQ 2005c. Water Quality in the North and Upper North Bosque Rivers: March 2005: An Updated Status of Activities to Address Elevated Nutrient Concentrations. Texas Commission on Environmental Quality. March 1, 2005. http://www.tceq.state.tx.us/.

TPWD 2005. 2005-2006 Texas Waterfowl Digest. PWD BK K0700-1014. Austin, Texas.

USACE 2005. U.S. Army Corps of Engineers. Fort Worth District, Waco Lake. http://www.swf-wc.usace.army.mil/waco/

U.S. Census Bureau. 2003. 2000 Census of Population and Housing, Population and Housing Unit Counts PHC-3-45, Texas. Washington, DC.

USDA 2002. 2002 Census of Agriculture, National Agricultural Statistics Service, United States Department of Agriculture. http://www.nass.usda.gov/census/census02/.

USEPA 2002. Method 1603: *Escherichia coli* (*E. coli*) in Water by Membrane Filtration Using Modified Membrane-Thermotolerant *Escherichia coli* Agar (Modified mTEC). Washington, DC, Office of Research and Development, Government Printing Office.

USEPA 2005. Microbial Source Tracking Guide Document. Cincinnati, OH, Office of Research and Development, EPA/600-R-05-064. June 2005.

USGS (U.S. Geological Survey). 1999. National Land Cover Dataset. Reston, VA. http://landcover.usgs.gov/natllandcover.asp.

APPENDIX A
AMBIENT WATER QUALITY DATA

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix A
Ambient Water Quality Data

Station ID	Station Description	Date	Number of Samples	Minimum EC Conc*	Maximum EC Conc*	Geometric Mean EC Conc.*	Number of Samples with EC present	Number of Samples with EC Conc. > 394*	Total Suspended Solids (mg/L)
11921	Lake Belton near dam	9/15/2003	10	<1	2	0.66	3	0	10
11922	Lake Belton Cowhouse Creek Arm	9/15/2003	5	<1	<1	<1	0	0	7
11923	Lake Belton Leon River Arm near headwater	9/15/2003	5	<1	<1	<1	0	0	34
11925	Leon River at FM 1829	9/16/2003	5	<1	12	1.00	2	0	42
TBD4	Lake Belton Owl Creek Arm	9/15/2003	5	<1	1	0.66	2	0	15
11942	Lake Waco near dam	9/16/2003	10	<1	3	0.64	2	0	23
11945	Lake Waco North Bosque arm	9/16/2003	5	<1	3	1.25	4	0	24
11948	Lake Waco Middle/South Bosque arm above SH6	9/16/2003	5	<1	<1	<1	0	0	9
11953	North Bosque River at SH 56 near Valley Mills	9/16/2003	5	<1	70	4.42	3	0	28
11956	North Bosque River at FM 219 northeast of Clifton	9/16/2003	5	<1	30	6.83	4	0	
TBD3	Lake Waco Middle/South Bosque arm near inlet of Middle/South Bosque	9/16/2003	5	<1	6	1.25	4	0	9
11921	Lake Belton near dam	10/20/2003	10	<1	<1	<1	0	0	4
11922	Lake Belton Cowhouse Creek Arm	10/20/2003	5	<1	1	0.57	1	0	5
11923	Lake Belton Leon River Arm near headwater	10/20/2003	5	12	29	22.88	5	0	42
11925	Leon River at FM 1829	10/21/2003	5	1	82	21.04	5	0	22
TBD4	Lake Belton Owl Creek Arm	10/20/2003	5	<1	20	8.07	4	0	29
11942	Lake Waco near dam	10/21/2003	10	27	53	34.84	10	0	31
11945	Lake Waco North Bosque arm	10/21/2003	5	12	21	15.75	5	0	22
11948	Lake Waco Middle/South Bosque arm above SH6	10/21/2003	5	42	58	48.33	5	0	27
11953	North Bosque River at SH 56 near Valley Mills	10/21/2003	5	<1	2	0.60	1	0	5
11956	North Bosque River at FM 219 northeast of Clifton	10/21/2003	5	<1	15	0.91	1	0	3
TBD3	Lake Waco Middle/South Bosque arm near inlet of Middle/South Bosque	10/21/2003	5	45	64	52.35	5	0	25
11921	Lake Belton near dam	11/19/2003	10	1	7	2.51	10	0	2
11922	Lake Belton Cowhouse Creek Arm	11/19/2003	5	2	12	4.36	5	0	5
11923	Lake Belton Leon River Arm near headwater	11/19/2003	5	2	8	4.58	5	0	19
11925	Leon River at FM 1829	11/18/2003	5	710	1,060	852.06	5	5	77
TBD4	Lake Belton Owl Creek Arm	11/19/2003	5	2	12	6.45	5	0	20
11942	Lake Waco near dam	11/19/2003	10	12	25	19.67	10	0	18
11945	Lake Waco North Bosque arm	11/19/2003	5	15	24	18.91	5	0	8
11948	Lake Waco Middle/South Bosque arm above SH6	11/19/2003	5	18	31	21.37	5	0	18
11953	North Bosque River at SH 56 near Valley Mills	11/18/2003	5	<1	550	85.75	4	2	12
11956	North Bosque River at FM 219 northeast of Clifton	11/18/2003	5	320	860	609.43	5	4	12
TBD3	Lake Waco Middle/South Bosque arm near inlet of Middle/South Bosque	11/19/2003	5	9	18	13.07	5	0	13
11921	Lake Belton near dam	12/17/2003	10	<1	<1	<1	0	0	2
11922	Lake Belton Cowhouse Creek Arm	12/17/2003	5	<1	1	0.57	1	0	6
11923	Lake Belton Leon River Arm near headwater	12/17/2003	5	<1	2	0.76	2	0	27
11925	Leon River at FM 1829	12/16/2003	5	20	48	30.51	5	0	4
TBD4	Lake Belton Owl Creek Arm	12/17/2003	5	<1	<1	<1	0	0	8
11942	Lake Waco near dam	12/17/2003	10	18	31	24.93	10	0	6
11945	Lake Waco North Bosque arm	12/17/2003	5	2	11	4.99	5	0	8
11948	Lake Waco Middle/South Bosque arm above SH6	12/17/2003	5	10	19	14.54	5	0	10
11953	North Bosque River at SH 56 near Valley Mills	12/16/2003	5	<1	13	0.89	1	0	2
11956	North Bosque River at FM 219 northeast of Clifton	12/16/2003	5	1	13	3.00	5	0	5
TBD3	Lake Waco Middle/South Bosque arm near inlet of Middle/South Bosque	12/17/2003	5	25	32	27.29	5	0	7
11921	Lake Belton near dam	1/20/2004	10	1	5	2.97	10	0	5
11922	Lake Belton Cowhouse Creek Arm	1/20/2004	5	12	20	15.54	5	0	6
11923	Lake Belton Leon River Arm near headwater	1/20/2004	5	282	361	329.10	5	0	22
11925	Leon River at FM 1829	1/20/2004	5	126	680	317.55	5	1	35

Appendix A
Ambient Water Quality Data

Station ID	Station Description	Date	Number of Samples	Minimum EC Conc*	Maximum EC Conc*	Geometric Mean EC Conc.*	Number of Samples with EC present	Number of Samples with EC Conc. > 394*	Total Suspended Solids (mg/L)
TBD4	Lake Belton Owl Creek Arm	1/20/2004	5	31	50	39.19	5	0	10
11942	Lake Waco near dam	1/21/2004	10	<1	78	14.76	8	0	7
11945	Lake Waco North Bosque arm	1/21/2004	5	<1	170	29.13	4	0	17
11948	Lake Waco Middle/South Bosque arm above SH6	1/21/2004	5	1	327	87.15	5	0	17
11953	North Bosque River at SH 56 near Valley Mills	1/20/2004	5	<1	420	5.32	2	1	6
11956	North Bosque River at FM 219 northeast of Clifton	1/20/2004	5	<1	640	25.18	4	1	5
TBD3	Lake Waco Middle/South Bosque arm near inlet of Middle/South Bosque	1/21/2004	5	278	360	313.41	5	0	21
11921	Lake Belton near dam	2/17/2004	10	<1	2	1.23	8	0	3
11922	Lake Belton Cowhouse Creek Arm	2/17/2004	5	<1	2	0.76	2	0	6
11923	Lake Belton Leon River Arm near headwater	2/17/2004	5	14	25	17.40	5	0	18
11925	Leon River at FM 1829	2/17/2004	5	<1	12	1.00	2	0	10
TBD4	Lake Belton Owl Creek Arm	2/17/2004	5	5	14	7.87	5	0	21
11942	Lake Waco near dam	2/16/2004	10	<1	17	3.03	7	0	5
11945	Lake Waco North Bosque arm	2/16/2004	5	<1	4	2.05	4	0	2
11948	Lake Waco Middle/South Bosque arm above SH6	2/16/2004	5	<1	17	7.07	4	0	5
11953	North Bosque River at SH 56 near Valley Mills	2/17/2004	5	<1	6	0.77	1	0	2
11956	North Bosque River at FM 219 northeast of Clifton	2/17/2004	5	<1	5	1.66	4	0	3
TBD3	Lake Waco Middle/South Bosque arm near inlet of Middle/South Bosque	2/16/2004	5	<1	22	1.07	1	0	10
11921	Lake Belton near dam	3/15/2004	10	<1	<1	<1	0	0	9
11922	Lake Belton Cowhouse Creek Arm	3/15/2004	5	55	67	61.85	5	0	13
11923	Lake Belton Leon River Arm near headwater	3/15/2004	5	<1	18	2.05	3	0	19
11925	Leon River at FM 1829	3/16/2004	5	10,600	20,000	17,615	5	5	494
TBD4	Lake Belton Owl Creek Arm	3/15/2004	5	<1	6	1.35	3	0	19
11942	Lake Waco near dam	3/16/2004	10	3	20	9.51	10	0	7
11945	Lake Waco North Bosque arm	3/16/2004	5	14	30	18.03	5	0	12
11948	Lake Waco Middle/South Bosque arm above SH6	3/16/2004	5	7	24	15.54	5	0	14
11953	North Bosque River at SH 56 near Valley Mills	3/16/2004	5	<1	8,500	756.81	4	4	29
11956	North Bosque River at FM 219 northeast of Clifton	3/16/2004	5	225	800	294.37	5	1	1
TBD3	Lake Waco Middle/South Bosque arm near inlet of Middle/South Bosque	3/16/2004	5	55	67	62.42	5	0	15
11921	Lake Belton near dam	4/19/2004	10	<1	1	0.57	2	0	1
11922	Lake Belton Cowhouse Creek Arm	4/19/2004	5	<1	1	0.57	1	0	18
11923	Lake Belton Leon River Arm near headwater	4/19/2004	5	<1	4	1.78	4	0	10
11925	Leon River at FM 1829	4/20/2004	5	331	470	407.86	5	3	25
TBD4	Lake Belton Owl Creek Arm	4/19/2004	5	1	4	1.64	5	0	18
11942	Lake Waco near dam	4/20/2004	10	<1	7	2.78	9	0	8
11945	Lake Waco North Bosque arm	4/20/2004	5	<1	3	1.08	3	0	7
11948	Lake Waco Middle/South Bosque arm above SH6	4/20/2004	5	1	4	1.32	5	0	8
11953	North Bosque River at SH 56 near Valley Mills	4/20/2004	5	66	169	115.60	5	0	11
11956	North Bosque River at FM 219 northeast of Clifton	4/20/2004	5	21	38	25.77	5	0	2
TBD3	Lake Waco Middle/South Bosque arm near inlet of Middle/South Bosque	4/20/2004	5	1	4	2.00	5	0	12
11921	Lake Belton near dam	5/17/2004	10	<1	<1	<1	0	0	5
11922	Lake Belton Cowhouse Creek Arm	5/17/2004	5	<1	1	0.57	1	0	17
11923	Lake Belton Leon River Arm near headwater	5/17/2004	5	<1	3	1.08	3	0	9
11925	Leon River at FM 1829	5/18/2004	5	<1	46	3.20	3	0	25
TBD4	Lake Belton Owl Creek Arm	5/17/2004	5	<1	6	1.30	2	0	19
11942	Lake Waco near dam	5/18/2004	10	<1	2	0.81	4	0	7
11945	Lake Waco North Bosque arm	5/18/2004	5	<1	6	2.83	4	0	11
11948	Lake Waco Middle/South Bosque arm above SH6	5/18/2004	5	3	11	6.62	5	0	11

Appendix A
Ambient Water Quality Data

Station ID	Station Description	Date	Number of Samples	Minimum EC Conc*	Maximum EC Conc*	Geometric Mean EC Conc.*	Number of Samples with EC present	Number of Samples with EC Conc. > 394*	Total Suspended Solids (mg/L)
11953	North Bosque River at SH 56 near Valley Mills	5/18/2004	5	<1	80	1.38	1	0	8
11956	North Bosque River at FM 219 northeast of Clifton	5/18/2004	5	<1	26	1.10	1	0	3
TBD3	Lake Waco Middle/South Bosque arm near inlet of Middle/South Bosque	5/18/2004	5	8	18	11.89	5	0	9
11921	Lake Belton near dam	6/21/2004	10	<1	<1	<1	0	0	<1
11922	Lake Belton Cowhouse Creek Arm	6/21/2004	5	<1	<1	<1	0	0	<1
11923	Lake Belton Leon River Arm near headwater	6/21/2004	5	<1	<1	<1	0	0	5
11925	Leon River at FM 1829	6/22/2004	5	<1	93	3.46	2	0	97
TBD4	Lake Belton Owl Creek Arm	6/21/2004	5	<1	2	0.76	2	0	19
11942	Lake Waco near dam	6/22/2004	10	<1	2	0.71	4	0	16
11945	Lake Waco North Bosque arm	6/22/2004	5	<1	13	6.24	4	0	26
11948	Lake Waco Middle/South Bosque arm above SH6	6/22/2004	5	<1	5	1.20	4	0	8
11953	North Bosque River at SH 56 near Valley Mills	6/22/2004	5	<1	<1	0.45	0	0	10
11956	North Bosque River at FM 219 northeast of Clifton	6/22/2004	5	<1	7	1.00	1	0	11
TBD3	Lake Waco Middle/South Bosque arm near inlet of Middle/South Bosque	6/22/2004	5	1	3	2.00	5	0	20

* *E. coli* (EC) Concentrations are reported in colony forming units in each 100 milliliters of water

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B
LABORATORY PROTOCOL FOR ISOLATION AND CONFIRMATION OF *ESCHERICHIA COLI* FROM WATER SAMPLES

THIS PAGE INTENTIONALLY LEFT BLANK

1. Follow the EPA Modified mTEC procedure described in USEPA Method 1603 (modified mTEC agar plate (USEPA Method 1603, http://www.epa.gov/nerlcwww/1603sp02.pdf) for isolation of *E. coli* colonies.

2. After the Modified mTEC 44.5±0.2°C incubation, the plates should be immediately stored at 4°C until shipment to prevent growth of non-*E. coli* coliforms on the plates.

3. Plates with red or magenta colored colonies should be parafilmed or taped closed, placed in plastic bags and then secured with tape to prevent the plates from being disturbed during shipment.

4. Ship plates in insulated coolers with ice packs sufficient to keep the plates between 1–4°C and ship by next day courier to:

   Dr. George D. Di Giovanni

   Texas A&M Agricultural Research and Extension Center

   1380 A&M Circle

   El Paso, TX 79927

   915-859-9111

5. Presumptive *E. coli* from the Modified mTEC plates will be isolated and confirmed as described in the protocol for fecal specimens.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C
SUMMARY OF KNOWN FECAL SAMPLES COLLECTED FROM LAKE WACO AND BELTON LAKE WATERSHEDS

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix C
Summary of Known Fecal Samples Collected

Lake Waco & Belton Lake			Sampling	Target No.	Actual Collected	Actual Collected	Actual Collected	Actual Collected	Actual Collected	Actual Collected	Actual Collected	Actual Collected	Actual Collected	Actual Collected		Sum of Collected	Negative Samples (did not produce at El Paso)	Potential Collection Areas	
Grouping	Animal		Lead Agency	Fecal Samples	Sep-03	Oct-03	Nov-03	Dec-03	Jan-04	Feb-04	Mar-04	Apr-04	May-04	Oct-04		Fecal Samples			
Human																			
	human - raw sewage/influent WW		Parsons	140				40	56	36			60	39		231	55	headv	human - raw sewage/influent WW
	human - septage w/o chemicals		Parsons	60						32				27		59		unloa	human - septage w/o chemicals
	human - treated WW effluent		Parsons	40				4								4	4	outfal	human - treated WW effluent
Pets																			
	cat		Parsons	25			2						19			21		cat	
	dog		Parsons	40				1		1	3		23	2		30	1	dog	
	other		Parsons	20														other	
		ferrets, gerbil, guinea pig	Parsons										3			3			ferrets, gerbil, guinea pig
		birds	Parsons													0			birds
		rabbit	Parsons			1							1			2			rabbit
Livestock																			
	cattle, dairy		TX Fm Bureau	60		1		14			40		15			70	7	cattle, dairy	
	cattle, beef		TX Fm Bureau	90		3		45			26	2	27			103	13	cattle, beef	
	chicken		TX Fm Bureau	15		2	3	4			4	1	2			16	8	chicken	
	turkey		TX Fm Bureau	5		3										3		turkey	
	horse		TX Fm Bureau	25		1		3			2	1	21			28	3	horse	
	goat, meat		TX Fm Bureau	15		3					7		23			33	2	goat, meat	
	goat, dairy		TX Fm Bureau	10							4					4	1	goat, dairy	
	sheep		TX Fm Bureau	25		2		1			27					30	3	sheep	
	llama		TX Fm Bureau			1										1		llama	
	pony, donkey		TX Fm Bureau			1							1			2		pony, donkey	
	guinea		TX Fm Bureau			1		3								4	3	guinea	
	pig/hog (domestic)		TX Fm Bureau	20		3					9		5			17		pig/hog (domestic)	
Wildlife (mammals/reptiles)																			
	racoon		Parsons	35			6	2		1	1	2	8	3		23	2	racoon	
	deer		Par+TxFB	40			17	5		1	3	6				32	5	deer	
	hog (feral)		Par+TxFB	20							1	2	1			4	2	hog (feral)	
	mouse		Par+TTU	15			1	2		1		3				7	7	mouse	
	rat		Par+TTU	15								1				1	1	rat	
	rabbit		Par+TTU	15		1	1		1	1	1	1	3			9	4	rabbit	
	opossum		Par+TTU	8		2	2	1	2		2	1		1		11	2	opossum	
	squirrel		Par+TTU	8		1	2	3			4	1	1			12	5	squirrel	
	armadillo		Par+TTU	8			1	1	1		1		3			7	1	armadillo	
	coyote		Anyone	5		1	1			1	11	6				20	1	coyote	
	fox		Par+TTU	5							1	1		1		3		fox	
	beaver		Par+TTU	2								2				2	1	beaver	
	nutria		Par+TTU	2												0		nutria	
	skunk		Par+TTU	5	1			1		1			1			4	1	skunk	
	bobcat		Anyone	2							1	1				2		bobcat	
	javalina		Anyone	1												0		javalina	
	other (exotics, cat (feral), etc.)		Anyone	20								1				1	1	other (exotics, cat (feral), etc.)	
Wildlife (bird/waterfowl)																			
	duck		Parsons	25		1	10					2	6	17		36	5	duck	
	swallow		Parsons	20			5					12				17	6	swallow	
	pigeon		Parsons	20		1	8							7		16	1	pigeon	
	heron		Parsons	15			5						1			6	2	heron	
	grackle		Parsons	15		1		1				1		17		20	1	grackle	
	turkey		Parsons	10			9				2	4				15	3	turkey	
	egret		Pars+Rehab	10									3			3		egret	
	martin		Parsons	10												0		martin	
	sparrow		Parsons	10								1				1	1	sparrow	
	dove		Parsons	10			5				1		1			7	2	dove	
	goose		Parsons	10		1	10						7			18		goose	
	other (birds)		Parsons	18														other (birds)	

Appendix C
Summary of Known Fecal Samples Collected

Lake Waco & Belton Lake			Sampling	Target No.	Actual Collected	Actual Collected	Actual Collected	Actual Collected	Actual Collected	Actual Collected	Actual Collected	Actual Collected	Actual Collected	Actual Collected		Sum of Collected	Negative Samples (did not produce at El Paso)	Potential Collection Areas	
Grouping	Animal		Lead Agency	Fecal Samples	Sep-03	Oct-03	Nov-03	Dec-03	Jan-04	Feb-04	Mar-04	Apr-04	May-04	Oct-04		Fecal Samples			
		wren	Parsons										1			1			wren
		warbler	Parsons										1			1			warbler
		cardinal	Parsons										3			3			cardinal
		crow	Parsons										1			1			crow
		blackbird	Parsons				6						1			7	4		blackbird
		meadowlark	Parsons				3	1								4	2		meadowlark
		cowbird	Parsons													0			cowbird
		mockingbird	Parsons			1										1			mockingbird
		flycatcher	Parsons													0			flycatcher
		snipe	Parsons				1									1	1		snipe
		vulture	Pars+Rehab					1		4	1					6	1		vulture
		killdeer	Parsons				2						1			3	1		killdeer
		hawk	Pars+Rehab													0			hawk
		owl	Pars+Rehab					1								1	1		owl
		starling	Parsons											15		15			starling
	other (waterfowl,shorebirds)		Parsons	18									1			1		other (waterfowl,shorebirds)	
		cormoran	Parsons										1			1			cormoran
		swan	Parsons													0			swan
		pelican	Parsons				9									9	2		pelican
		seagull	Parsons				1									1			seagull
		plover	Parsons													0			plover
		crane	Pars+Rehab													0			crane
Zoo/Pettir	various		Parsons	13												0		zoo, carnival, petting zoo	
	SUM			1000	1	32	110	134	60	79	152	52	245	129		994	166		
															ck	994			

All samples collected from this species were negative

No samples were collected from this species

Sum of Collected

Negative Samples

wildlife only:

APPENDIX D
LABORATORY PROTOCOL FOR ISOLATION AND CONFIRMATION OF *ESCHERICHIA COLI* FROM FECAL SPECIMENS

THIS PAGE INTENTIONALLY LEFT BLANK

Note: All collection and handling of fecal specimens should be performed using protective gear (i.e. latex or nitrile gloves). Specimens should be handled aseptically to ensure sample quality and minimize exposure of personnel to pathogens.

1. Fecal specimens should be refrigerated as soon as possible after collection and shipped in insulated coolers with ice packs sufficient to keep the specimens between 1–4°C. Ship by next day courier to:

   Dr. George D. Di Giovanni

   Texas A&M Agricultural Research and Extension Center

   1380 A&M Circle

   El Paso, TX 79927

   915-859-9111

Note: All handling of fecal specimens and cultures will be performed using a Class 2 biological safety cabinet to minimize the exposure of laboratory personnel to pathogens.

2. Using a bacteriological loop, streak a loopful of fecal material onto a labeled modified mTEC agar plate (USEPA Method 1603, http://www.epa.gov/nerlcwww/1603sp02.pdf) for isolation of *E. coli* colonies.

3. Invert the plate and incubate at 35±0.5°C for 2 h.

4. After the 2-h incubation at 35±0.5°C, incubate the plate inverted at 44.5±0.2°C for 20–24 h.

5. Examine the plate for presumptive *E. coli* colonies, which will appear red or magenta colored.

6. Select up to three presumptive *E. coli* colonies and streak each colony for purity onto a labeled nutrient agar with MUG (NA-MUG) plate.

7. Invert and incubate plates at 35–37°C for 20–24 h.

8. Examine the cultures using a long-wave handheld UV lamp. If there is a mixture of fluorescent and non-fluorescent colonies, select a well isolated fluorescent colony and streak again onto NA-MUG for purity.

At the discretion of the laboratory, additional biochemical tests such as urease, indole and citrate tests may be performed.

APPENDIX E
ARCHIVAL OF *ESCHERICHIA COLI* ISOLATES

THIS PAGE INTENTIONALLY LEFT BLANK

Note: All handling of cultures will be performed using a Class 2 biological safety cabinet to minimize the exposure of laboratory personnel to pathogens.

1. Select a well isolated colony of purified *E. coli*.

2. Using a bacteriological loop, transfer the colony to a labeled sterile cryovial containing 1 mL of tryptone soy broth (TSB) with 20% reagent grade glycerol. Verify that the cells have been resuspended.

3. Firmly cap the cryovial and plunge into liquid nitrogen until frozen.

4. Immediately transfer to a cryostorage box and place in -70 to -80°C freezer. Cultures may be stored for several years under these conditions.

5. To recover cultures from frozen storage, remove the cultures from the freezer and place the cryovials in a freezer block.

    a. Using a bacteriological loop, scrape the topmost portion of the culture and transfer to growth medium, being careful not to contaminate the top or inside of the vial.

    b. Reclose the cryovial before the contents thaw and return to the freezer.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX F
LABORATORY PROTOCOL FOR ERIC-PCR FINGERPRINTING OF *ESCHERICHIA COLI*

THIS PAGE INTENTIONALLY LEFT BLANK

Laboratory Protocol for Enterobacterial Repetitive Intergenic Consensus Polymerase Chain Reaction (ERIC-PCR) Fingerprinting of *Escherichia coli*

1. Select isolated colonies from overnight cultures of *E. coli* isolates on BHI plates.

2. Transfer colonies using 1 µL loop to sterile microfuge tubes containing 100 µL of sterile molecular grade water, vortex briefly to suspend cells.

3. Prepare sufficient PCR Master Mix for samples, including one blank per 10 samples to account for volume loss due to repeat pipetting. Prepare Master Mix for each sample as follows. One full PCR batch on the MJ Research Cycler 48 well-plate will have 46 samples, *E. coli* QC101, and a no template control (NTC).

ERIC-PCR Master Mix – 24 samples + 2 blanks

(prepare X 2 for full 48-well plate)

Table F-1 ERIC-PCR Master Mix

MASTER MIX	Amt (uL)	Final Calc	Final Units
dH2O	**31.5**		
10X PCR buffer I w Mg	**5**	1	X
20 mM dNTP	**0.5**	200	uM each
ERIC Primer Mix	**5**	600	nM each
BSA (30 mg/ml)	**2.5**	1.5	ug/uL
AmpliTaqGold (Units)	**0.5**	2.5	Units/rxn

4. Dispense 45 µl of Master Mix for each sample into the appropriate well of PCR plate.

5. Briefly vortex cell suspensions, then add 5 µl of each cell suspension to the appropriate PCR well.

6. Carefully seal plate using an adhesive PCR cover.

7. Load the plate into the thermal cycler and run under the "ERIC-PCR" program with the following cycling conditions:

    a. Initial denaturation at 95°C for 10 min
    b. 35 Cycles:

i. Denaturation at 94°C for 30 sec
ii. Annealing at 52°C for 1 min
iii. Extension at 72°C for 5 min

c. Final Extension at 72°C for 10 min

8. Store completed reactions at -20°C until analyzed by gel electrophoresis.

9. Prepare a 250 mL, 2% agarose gel using a 500 mL bottle. Add 250 mL of 1 X TBE buffer and 5.0 g agarose. Microwave until agarose is fully dissolved, tighten cap and let cool 1-2 minutes, then pour agarose into casting tray with 30-tooth, 1 mm thick comb.

10. Allow gel to solidify for approximately 30 minutes on the bench, then without removing comb place in Ziploc bag and solidify overnight in the refrigerator. The next day carefully remove comb, transfer to gel tank in cold room (4°C) containing pre-cooled 1X TBE buffer. Replace TBE in gel tank after it has been used twice.

11. The following items will be needed for electrophoresis:

100 bp ladder (0.33 µg/10 µL) (1500 µL final, enough for 150 lanes)

200 µL Roche DNA Marker XIV (Cat. #1721933) 0.25 µg/µL 100 bp ladder (add reagents below to a full tube of marker)

300 µL 6X ERIC-PCR loading buffer (see recipe below)

150 µL 10X PCR buffer

850 µL molecular grade water

Store in cold room

6X ERIC-PCR Loading Buffer

25 mg bromphenol blue (0.25%)

1.5 g ficoll 400 (15%)

Add molecular grade water to 10 mL, divide into 1 mL aliquots and freeze, the aliquot currently being used can be stored in the cold room

<u>ERIC-PCR Blank</u>

100 µL 10X PCR buffer

200 µL 6X ERIC-PCR loading buffer

900 µL molecular grade water

Store in cold room

<u>Ethidium Bromide Stain (0.5 µg/mL)</u>

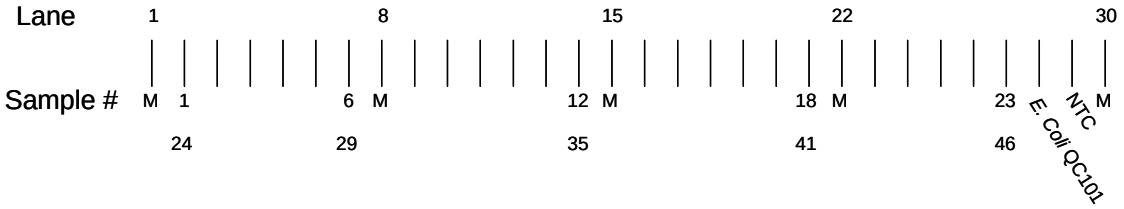
1250 mL 1X TBE

62.5 µL ethidium bromide (Sigma, 10 mg/mL)

Store covered at room temp, can use up to 5 times by adding 10 µL ethidium bromide each additional use

12. Mix 10 µL of 6X ERIC-PCR Loading Buffer to each PCR well and mix with pipette tip.

13. Load the gel in the cold room as follows (max. of 23 samples + QC101 + NTC per gel):
    a. Load 10 µl of 100 bp ladder (0.33 µg) into the first lane
    b. Load 10 µl of sample ERIC-PCR reactions into next 6 lanes
    c. Load 10 µl of 100 bp ladder (0.33 µg)
    d. Load 10 µl of sample ERIC-PCR reactions into next 6 lanes
    e. Load 10 µl of 100 bp ladder (0.33 µg)
    f. Load 10 µl of sample ERIC-PCR reactions into next 6 lanes
    g. Load 10 µl of 100 bp ladder (0.33 µg)
    h. Load 10 µl of sample ERIC-PCR reactions into next 5 lanes
    i. Load PCR Batch *E. coli* QC101 and NTC into next 2 lanes
    j. Load 10 µl of 100 bp ladder (0.33 µg)



If running a gel with fewer samples, follow steps above until last sample, followed by *E. coli* QC101, NTC and ladder, then load ERIC-PCR Blank into remaining lanes on gel.

14. Start electrophoresis power supply set at 100 volts, run for 1 hour.

15. Stop power supply, set time to "000", set voltage to 200 and start circulating pump at setting #2, run for 4 hours.

16. After electrophoresis, stain gel in Ethidium Bromide Stain for 20 minutes with agitation (save stain, see Step 13).

17. Destain gel for 10 minutes in 1X TBE buffer. Save destain, can be used 3 times then discard.

18. Follow Gel Logic 200 SOP for image capture. Save digital photograph as a TIFF file (default) and print a hardcopy for notebook.

APPENDIX G
SELECTION OF ISOLATES FOR LIBRARY INCLUSION BASED ON ERIC-PCR FINGERPRINTS

THIS PAGE INTENTIONALLY LEFT BLANK

Revised 12-20-04

Note: The terms "isolates" and BN "entries" are used interchangeably.

Ambient Water Isolates

1. At present, only 1 isolate per sample (in alphabetical order) is being analyzed by ERIC-PCR and RiboPrinting for the TCEQ San Antonio project, while all isolates from each water sample are being analyzed for the TSSWCB Waco project.

Source Sample Isolates

1. At least one isolate from each sample will be selected for RiboPrinting, even if it is identical to a previously selected isolate.
2. If there is only one isolate for a sample, select it for RiboPrinting – complete Step 7, then go to Step 10.
3. Select BioNumerics (BN) entries of all isolates for a single sample, create a temporary Comparison in BN (do not name or save).
4. Maximize the Comparison window and select the ERIC-PCR icon on the bottom menu bar to display the gel strips.
5. Click on the Dendrogram icon on the top menu bar, select Calculate Cluster Analysis, curve-based Pearson similarity coefficient and UPGMA Dendrogram. Sometimes you have to repeat steps 4 and 5 to get the Dendrogram to display.
6. Determine how many different ERIC-PCR fingerprint types there are (80% sim cutoff, i.e. <80% sim is unique, >80% sim is similar to previous entries). *Interpret the results, keeping in mind that isolates with visibly obvious identical banding patterns may cluster differently due to smearing in the lane when using curve based analyses. Consider these isolates the same ERIC-PCR type, and note "smear" after the information added to the Excel database in Step 4 below. Select the entry/isolate that does not have a smear in its lane for RiboPrinting, even if it is not in alphabetical order.* Designate, in order of abundance and alphabetically, the types as 1, 2 and 3. Example possible combinations for 1 to 3 isolates per sample are:

   1

   1, 1

   1, 1, 1

   1, 1, 2

   2, 1, 1

   1, 2

   1, 2, 1

   1, 2, 3

7. Enter the "ERIC-PCR Types" designation for each isolate entry into the Excel BN Import database.
8. Close the temporary comparison window without saving.
9. Select the BN entry of one representative isolate (in alphabetical order) of each ERIC-PCR type for that sample. For example, if Isolate A, B, C were ERIC-PCR Types 1, 2, 1, respectively, BN entries Isolates A and B would be selected.
10. Add entries to the Comparison list that contains only source isolates previously selected for RiboPrinting. The curve-based Pearson cluster analysis is automatically updated.
11. Determine if the added entries are unique or cluster with existing entries (80% sim cutoff, i.e. <80% sim is unique, >80% sim is similar to previous entries). Each isolate that is unique should be selected for RiboPrinting. *Use some judgment here – if all or a majority of the isolates appear to be unique (<80% sim to existing entry), please contact Dr. Di Giovanni.*
12. If all of the isolates from a sample cluster with existing entries, select one isolate (in alphabetical order) from the most abundant ERIC-PCR type for that sample (e.g. Isolate A from example given in Step 6 above). Also, each cluster should be composed of at least two entries. Therefore, if an isolate is >80% similar to a *single existing entry* it should be selected for RiboPrinting, regardless of the selection of other isolates from the same sample. If all isolates are already represented in the BN selected database ("Library" for the watershed), then select the isolate that is least represented in the database.
13. Designate the isolate(s) to be RiboPrinted in the Excel BN Import database, save.
14. Remove the unwanted entries from the Comparison window (*not the database*), the dendrogram is automatically updated. Save the updated comparison.
15. Repeat for isolates from the next sample.

1. When finished with a gel, go and update the Main Excel database with the ERIC-PCR types and Isolates to be RiboPrinted information.

APPENDIX H
LABORATORY PROTOCOL FOR AUTOMATED RIBOTYPING OF *ESCHERICHIA COLI* USING THE DUPONT QUALICON RIBOPRINTER

THIS PAGE INTENTIONALLY LEFT BLANK

Storing and Handling Disposables

Check the lot expiration date on each label for details and rotate the stock to optimize use.

Heating MP Base

After storage and the temperature changes that occur during shipment, the oxygen in the buffer loaded in the MP base may need to be removed before use. This is called degassing and is accomplished by heating the base pack overnight in your incubator.

To degas buffer:

1. Place enough MP base packs for the next day's production in their storage pouches in an incubator set at 37°C.
2. Allow the base pack to degas for 16-24 hours prior to loading in the characterization unit. You may do this while you are incubating samples, since the base packs are sealed in their pouches. This procedure allows you to start a batch immediately at the beginning of the next shift.
3. If you do not use the heated base packs, you can return them to storage and reuse them. These base packs should be heated again before reuse since temperature cycling affects oxygen content in the buffer.

Preparing Lysing Agent

Lysing agent (A and B) is shipped frozen and must be stored at -20°C.

Lysing agent must be thawed before use. This only takes about 5 minutes. If the lysing agent will not be used again for more than 2 hours, the material should be returned to the freezer. Lysing agent can be re-frozen several times with no effect on performance.

Sample Preparation Procedures

1. Incubate and Inspect the Samples

Use BHI (Brain-Heart Infusion) agar plates prepared within the last 30 days. Do not use plates that appear dry or dehydrated. Such plates can cause problems when you attempt to "pick" the colonies for use in the RiboPrinter® system.

1. Using a pure isolated colony as the source, streak BHI agar plates heavily in the upper portion of the plate to create a lawn. Streak the remainder of the plate lightly to create single colonies.
2. Follow standard laboratory techniques. Heat plates for 18-30 hours in a humidified incubator at 37 °C.

2. Transfer Sample Buffer to Intermediate Tubes

1. Locate the 250 mL twist-top bottle of sample buffer supplied in Pack # 1 Install the twist cap.
2. Transfer about 5 mL of buffer to a sterilized disposable 15 mL intermediate working tube.

3. Add sample buffer to microcentrifuge tubes

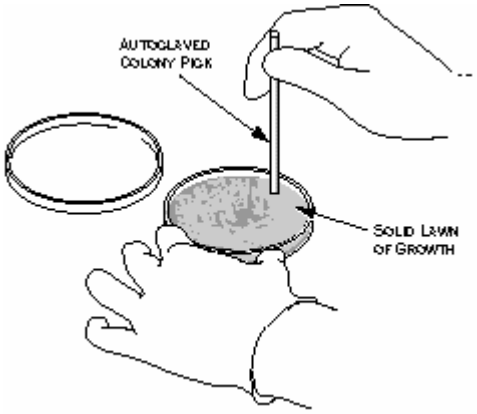
1. Place a sterile 0.65 mL microfuge tube in each of the eight holes in the lower row of the sample preparation rack.
2. For Gram negative samples (including *E. coli*), add 200 µL of sample buffer from the intermediate tube.

For Gram positive samples (*e.g. S. aureus* and *L. innocua* QC strains), add 40 µL of sample buffer.

3. Close the lids on the tubes.

4. Harvest the Samples

1. Using autoclaved colony picks and making certain not to gouge the agar, carefully place the pick into one of the single colonies or the lawn. You need a sample area at least equal to that of the bottom of the colony pick. In most cases you will need to harvest from the lawn area of the plate. If you are working with large colonies, a single colony will be adequate.



2. For Gram negative samples (e.g. *E. coli*), perform 1 pick placed into 200 µL of sample buffer.

CAUTION! Do not try to use the same pick twice on a plate. You need to harvest only enough sample to cover the bottom surface of the pick. Make sure the end of the pick is flat, if not, use a different pick.

CAUTION! Do not overload the harvesting pick. Collect only enough sample to cover the base of the pick. Over sampling will cause inaccurate results. Over sampling is a particular problem with *Staphylococcus*.

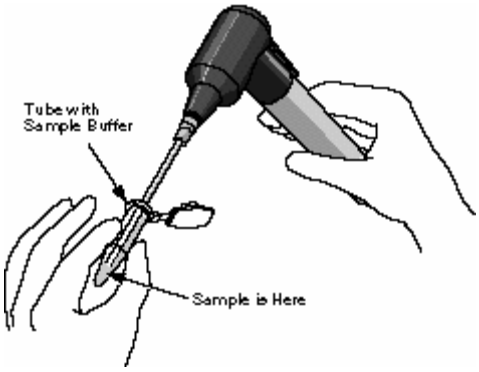
5. Mix the Samples

WARNING! *Perform sample preparation using a Class 2 biological safety cabinet since aerosols may be formed during mixing of the samples.*

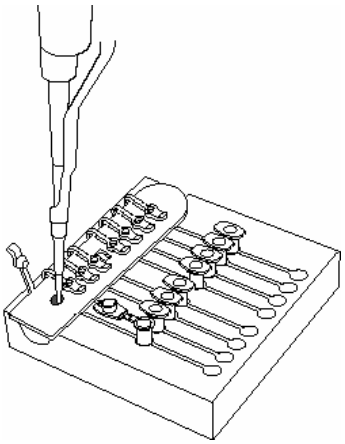
1. Making certain not to touch the sample end of the pick, place the pick into one of the filled sample tubes.
2. While holding the tube with the open end facing away from you, carefully attach the pick to the hand-held mixer. The fit of the pick in the coupling will be loose.

WARNING! Do not turn on the mixer unless the pick is inside the sample tube and below the surface of the liquid. Turning the unit on at other times will cause the sample to aerosolize and may cause contamination.

3. Press the ON lever on the mixer for about 5 seconds.
4. Release the lever and carefully remove the colony pick. The sample liquid should appear turbid.
5. For **Gram positive samples only**, (e.g. *Staphylococcus* and *Listeria*) locate a new colony pick and repeat the steps for harvesting and mixing samples, adding a second sample to the original tube. Discard the used picks in a biowaste bag.
6. Cap the sample tube.
7. Move the tube to the top row of the sample preparation rack. This indicates that the tube is filled.



6. Transfer the Samples to the Sample Carrier



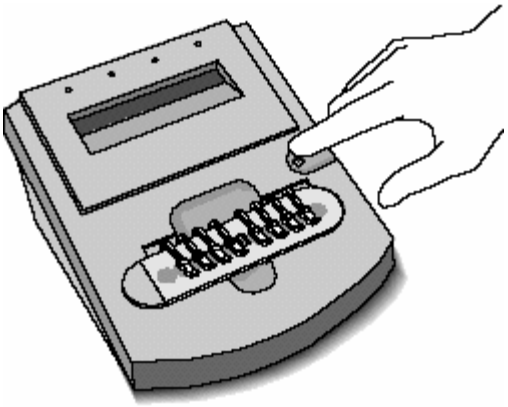
1. Open the lid covering the first well of the sample carrier.
2. Using a 100 µL pipetter, pipette 30 µL of sample from the microcentrifuge tube into the well.
3. Close the lid cover for the well.
4. Repeat for remaining samples using a new pipet tip for each sample.

CAUTION! Transfer the sample carrier to the Heat Treatment Station within 2 hours. If you wait longer than 2 hours, you will have to discard the sample carrier and begin again for this batch.

6. Lightly wipe down the outer surfaces of the sample carrier with a lab wipe wetted with surface disinfectant (10% bleach or 70% alcohol).
7. Write down the name or code you use to identify the sample and the well number in the sample carrier for each sample using a sample log sheet.

7. Place the Sample Carrier in the Heat Treatment Station and Process the Sample Carrier

1. Place the sample carrier into the Heat Treatment Station. The display on the Heat Treatment Station will show **Insert**, if power is available. If the display is blank, make certain that the power cord on the back of the station is properly connected.



After you insert the carrier, the display shows **Press Button.**

2. Press the button on the Heat Treatment Station.

   The display shows **Warm up** and counts down from **10** while the station is warming up. The actual warm up cycle varies with the condition of the room and the heat treatment station. Normal time is about 4 minutes.

   When the station reaches operating temperature, the display changes to **Heat** and counts down from **13.** This represents each minute of heat treatment.

   The indicator message changes to **Cool.** The display counts down from **9**, indicating the minutes remaining in the cooling cycle**.** If necessary, you can remove the carrier as soon as the **Cool** message appears.

3. The heat treatment step is finished when the display shows **READY** and counts down from **90.** The display will flash and an audible beep will sound three times. The alarm will then beep once every 10 minutes until the sample is removed or 90 minutes elapses.

Caution! The heat-treated samples must be used within the 90-minute period at room temperature or they must be discarded. The heat-treated samples may be stored at this point (prior to adding Lysis Agents) for 1 week at 4 °C, or for several months at -70 °C.

8. Add the Lysing Agents

1. Using a 10-µL pipetter and new tips for each addition, add 5 µL of Lysing Agents A and B to each sample. Note: this step may be omitted for *E. coli* if no effect on ribopatterns is demonstrated. Lysing Agents were specifically developed for *Staphylococcus* and Lactic-Acid bacteria samples.

Caution! This step must be performed just prior (within 10 minutes) of loading the samples into the RiboPrinter and starting the run.

Creating and Loading a Batch

There are three options under the Operations menu for creating standard batches;

- *Eco*RI batches (VCA)
- *Pst*I batches (VCB)
- *Pvu*II batches (VCC)

You can also create special batches:

- Restriction Enzyme Flexibility batches
- **Substitute Enzyme batches (including *Hin*d III)**

From the Instrument Control Base Window:

1. Move the pointer to Operations and click with the mouse button. The Operations menu appears.
2. Move the pointer to Create Substitute Enzyme Batch and click with the mouse button.
3. Use the View menu to remove any optional items you do not wish to fill in. The system requires at least Sample Type and RiboGroup Library information for each sample. You cannot remove these options. The **Clear** option de-selects the **Use Default ID Libraries.** You will have to enter a DuPont ID and Custom ID library name for all samples. These become required fields and the system will make you enter data before you can save the information in this window.

CAUTION! If you change the display after you have entered information, you will lose all the information in the window. The window will redraw with a new blank display showing the items you have selected.

4. To enter information about the sample, click on the **View** button with the mouse button, then click on **Sample Items**. Click on the options you want to display.
5. Enter your initials and any comment you want to record about the batch.
6. Select the lot number fields and record for all reagents.

CAUTION! All fields must be completed or the system will not let you start processing the batch.

7. For each well in the sample carrier, choose the type (Sample or Control [QC Number]) from the Sample Type field. The system defaults to Sample.
8. Once you define the Sample Type as Sample, type in the name you actually want to use. This information will appear as Sample Label in the Data Analysis software screens.
9. You can change the RiboGroup library name if needed. Do this by clicking on the button next to the field with the mouse button. A pop up menu appears listing your choices. If you want to add a new library name, move the pointer to the line and click with the mouse button to get a cursor, then type in the new library name. Once you have saved this file, the new name will be added to the pop up list for future use. Do **NOT** change the DuPont ID field. If you select one of the QC strains, the system automatically enters QC in the DuPont ID and RiboGroup Library fields. Do not change these names. If you wish, you may enter a name for the Custom ID library.
10. Repeat for the other seven samples.
11. Click on Save and Submit Batch to Instrument.

Loading Disposables

Follow the screen prompts to load disposables and check the DNA Prep Waste. The icons on the window will flash red to tell you to remove and load an item. The screen prompts you about which Separation and Transfer chamber to use for the membrane and gel cassette. The LDD Pipette will move to physically block you from placing samples in the wrong chamber.

CAUTION! *Do not try to move the pipette manually. You will cause the system to lose the step count. This can result in the loss of batch data. If the pipette is blocking the S/T chamber that you are instructed to use, STOP. Call Customer Support.*

CAUTION! *Do not load disposables until you are prompted by the system. If you try to load them earlier, the alarm will sound as long as the doors are open. If you do load disposables ahead of time, the MP Base will be moved to the wrong position and you will not be able to begin processing the batch. You will not be able to move the MP base manually.*

1. Check the DNA Preparation Waste Container

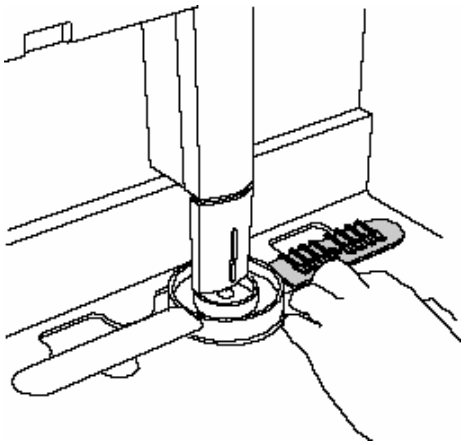
1. The DNA Prep waste container must be visually checked before every batch. If the container looks nearly full (about 1 inch from the top), remove the container, unscrew the cap and empty into the liquid biohazard waste.

WARNING! Do not tip the DNA Preparation waste container when you remove it.

WARNING! Do not unscrew the cap from the DNA Preparation waste container if the fluid level has risen into the cap. First pour the excess waste liquid into the liquid biohazard waste.

WARNING! When replacing container make sure that the cap is properly threaded in place. If the cap is only partially threaded, it can snag the pipette during operation.

2. Load the Sample Carrier



1. Place the sealed carrier into the labeled slot on the far right of the characterization unit.
2. Push the sample carrier down firmly until it snaps into place.

CAUTION! Place the rounded edge of the sample carrier on your right as you view the characterization unit. Position the carrier this way to insure correct identification of the sample wells.

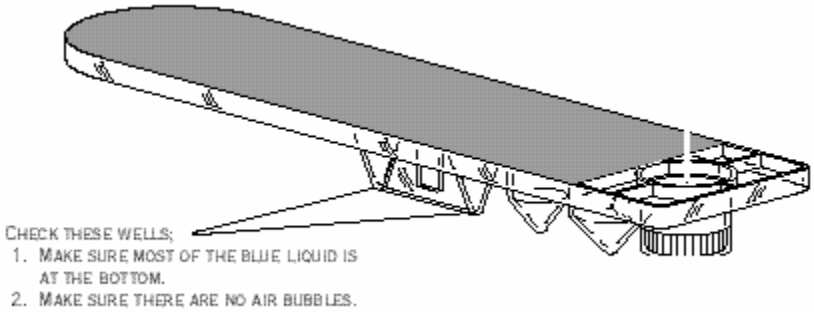
3. Load the DNA Prep Carrier

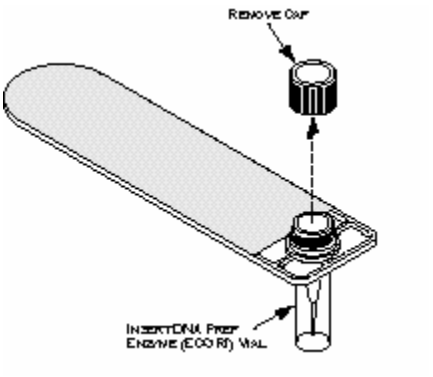
1. Remove the DNA Prep carrier from the refrigerator.
2. Check the wells in the carrier. If most of the liquid appears to be in the bottom of the wells and there are no bubbles, go to step 3. Otherwise **lightly tap the side of the carrier a few times with your finger to release any material that has adhered to the lid.**
3. CAUTION! Do not tap the carrier briskly. This may cause the marker to degrade which can create inaccurate results.
4. Remove a vial of DNA Prep Enzyme (*Hin*d III or *Eco*R I) from the freezer. ***Hin*d III (NEB Cat. #R0104M) is prepared in a Sarstedt 500-µL microfuge tube (Cat. #72730-005) as either a 100 or 50 U/µL working stock as follows.**

   100 U/µL: 50 µL *Hin*d III and 3 µL of NEB Buffer 2

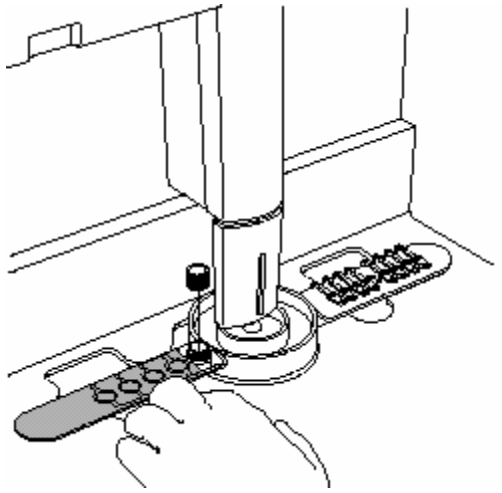
   50 U/µL: 26.5 µL *Hin*d III and 3 µL of NEB Buffer 2

   During addition of the Buffer, mix enzyme and buffer to homogeneity by stirring with the micropipette tip.



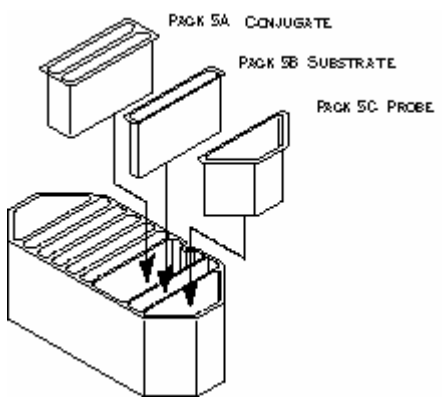


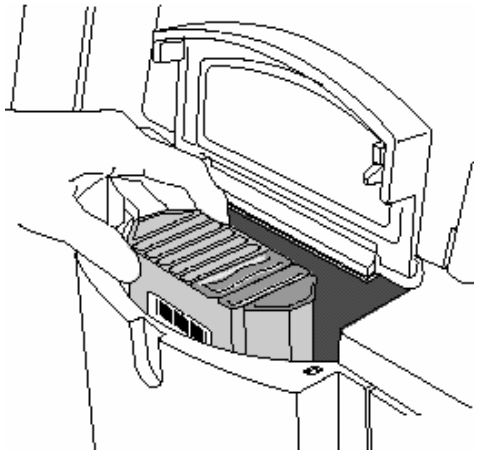
5. Remove the cap from the Enzyme vial.
6. Insert the vial into the carrier.
7. Place the DNA Prep carrier into the slot labeled **Reagent** to the left of the sample carrier slot.
8. Push the DNA Prep carrier down firmly until it snaps into place.



4. Load the MP Base and Carousel

1. Unpack the disposables.
2. Remove the MP base (Pack 5) from the incubator and the Conjugate (Pack 5A), Substrate (Pack 5B), and Probe (Pack 5C) from the refrigerator.
3. Remove each insert from its pouch. Tap the powdered reagent packs gently to bring all powder to the bottom of the packs. Place reagent packs in the MP base and load the base in the carousel.

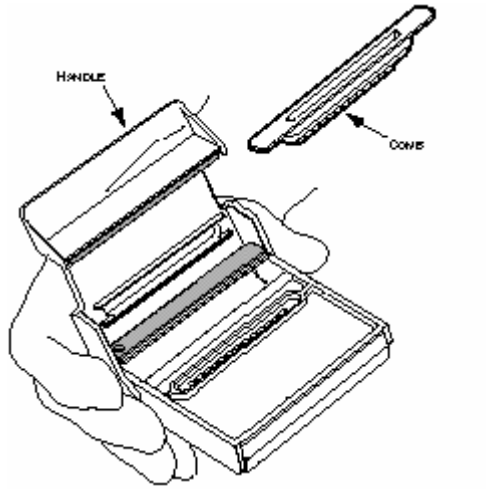


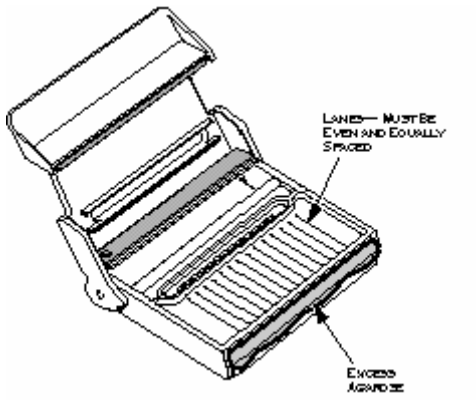


CAUTION! Push each insert firmly into place. If part of the insert extends above the top of the base, it could catch on the bottom of the deck and cause a system error. You could lose one or more batches as a result. Each insert is keyed by shape and cannot be inserted incorrectly.

5. Load the Gel Cassette

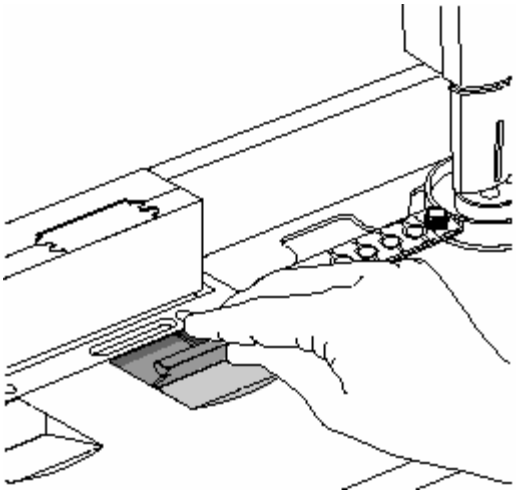
1. Remove the gel cassette from its package.
2. Grasp one end of the rubber comb and gently pull the comb from the cassette.
3. Unfold the handle of the cassette towards you until the handle snaps into place.
4. Check the front edge of the gel cassette and the lanes of the gel.





Warning! *If the cassette shows a build up of excess gel on the front edge, or if you notice any shrinkage of the gel away from the cassette or bubbles, record the lot number and call Customer Support. Use a new cassette for this run.*

5. Insert the gel cassette into the slot labeled **Gel Bay.** The RiboPrinter® system will prevent the insertion of the cassette into the incorrect slot by blocking one slot with the LDD Pipette.

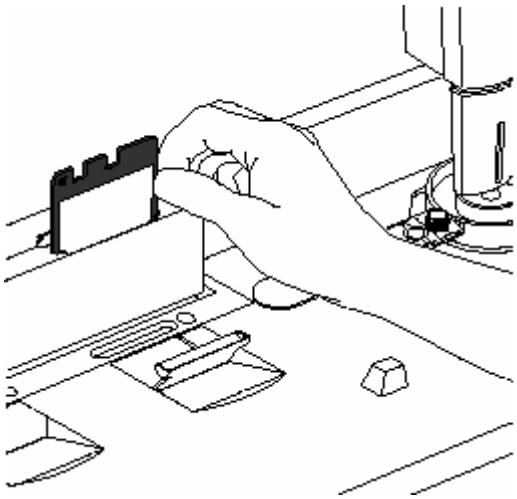


6. Press the cassette forward firmly until it snaps into place.

6. Load the Membrane

1. Grasp the membrane and carefully drop it into the front slot and flip the metal bracket against the back of the membrane.

CAUTION! You can insert the membrane backwards. This will cause an alarm that prevents the sample from being processed until the error is corrected. Always make certain that the two large slots are on top and that the square hole on the side faces your left as you insert the membrane.



7. Close all doors and the instrument will begin sample processing.

8. Load the Next Batch

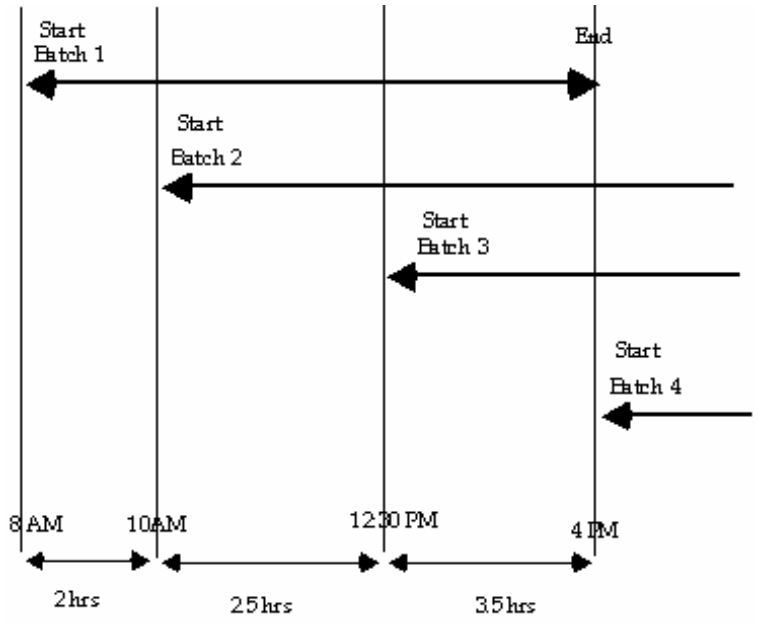
The RiboPrinter® microbial characterization system lets you load up to four VCA batches in an eight hour period. **Other batches** may take longer to process.

The chart above shows the approximate loading times for each batch in a work shift using only the VCA protocol.

1. You can now use the **Create Batch** option to set up a new pending batch.
2. When you complete the information window and click on the **Start Normal Batch** option, the window displays a message telling you when you can load the next batch.

Batch Report

After image processing is completed, the system automatically runs a series of analysis functions and generates a Batch Information Report. This task does not require any action on the part of the operator. Reports are automatically saved to the hard disk of the computer and sent to the printer.



THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX I
LABORATORY PROTOCOL FOR MOLECULAR SUBTYPING OF *ESCHERICHIA COLI* BY PULSE FIELD GEL ELECTROPHORESIS (PFGE)

THIS PAGE INTENTIONALLY LEFT BLANK

Day 0 – Bacterial cultures

Streak overnight broth culture of TSB onto TSA for confluent growth using inoculating loop. Incubate cultures at 37º C for 14-18 hrs.

Day 1 – Preparation of agarose plugs (This part will be performed during the lab session of September 27)

1. Prepare 1% SeaKem Gold:1% SDS agarose in TE Buffer (10mM Tris:1mM EDTA, pH-8.0). Melt agarose in microwave and store in 55º C water bath until use.
2. Transfer 5 ml of Cell Suspension Buffer (CSB) (100mM Tris:100 mM EDTA, pH-8.0) to tubes. Use sterile cotton swab that has been moistened with CSB to remove confluent growth from agar plates; suspend cells in CSB by spinning swab gently to disperse cells.
3. Adjust concentration of cell suspension to Optical Density value of 1.35 at 610nm wavelength of light using a spectrophotometer. Concentration of cells can be changed by adding cells or CSB as needed.
4. Transfer 0.4 ml adjusted cell suspension to 2 labeled 1.5-ml microcentrifuge tubes at room temperature.
5. Add 20ul of Proteinase K (20mg/ml stock) to each tube and mix gently with pipet tip. Proteinase K should be maintained in ice and stored at -20º C
6. Add 0.4 ml melted % SeaKem Gold:1% SDS agarose to the 0.4 ml cell suspension; mix by gently pipetting mixture up and down a few times. Maintain temperature of melted agarose by keeping flask in 55º C water bath.
7. Immediately, dispense part of mixture into appropriate wells of plug mold (~0.4 ml per plug). Make two plugs per specimen. Allow plugs to solidify at room temperature for 10 min.
8. Prepare Cell Lysis Buffer (50 mM Tris: 50 mM EDTA, pH 8.0 + 1% Sarcosyl). Calculate the total volume of Cell Lysis/Proteinase K Buffer needed as follows:
    a. 5ml of Cell Lysis Buffer plus 25ul Proteinase K Stock Solution (20mg/ml) per tube.
    b. Measure correct volumes into appropriate size test tube or flask and mix well.
9. Add 5 ml of Proteinase K/Cell Lysis Buffer to a screw-cap tube for cell lysis of plugs.
10. Transfer plugs from plug molds into tubes containing Proteinase K/Cell Lysis Buffer. Place tubes into rack in 54º C shaker water bath for 1.5 hrs.
11. Pre-heat sterile reagent grade water to 50º C so that plugs can be washed two times with 10 ml of water
12. Remove tubes from water bath and pour off lysis buffer into discard container.

13. Add 10 ml of sterile Pre-heated water to each tube and shake tubes in 50º C Water bath for 10 min.
14. Pour off water from the plugs and repeat wash steps once more. Pre-heat TE Buffer (10mM Tris:1mM EDTA, pH-8.0) in 50º C water bath so that the plugs can be washed four times with 10 ml of buffer.
15. Pour off water and add 10 ml pre-heated TE buffer and shake vigorously in 50º C water bath for 10 min.
16. Pour off TE and **repeat wash steps three more times**.
17. Decant last wash step and add 5 ml TE. Store plugs at 4º C until restriction digest.

Day 2 - Restriction digestion of DNA in agarose plugs with XbaI

1. Dilute 10X H Buffer (Roche Molecular Biochemicals) 1:10 with sterile reagent grade water.
2. Add 200 ul diluted H buffer to labeled microcentrifuge tube. Carefully remove plugs from TE with spatula and place in sterile disposable Petri dish.
3. Cut 2.0 mm Slice from plug with scalpel and transfer to tubes with diluted H buffer. Replace remaining plug into TE buffer and store at 4º C.
4. Incubate sample and control plug slices in 37º C water bath for 5 min.
5. After incubation, carefully remove buffer from plug slice using a micropipettor. Dilute 10X H buffer 1:10 and add XbaI restriction enzyme (50U/Sample). Mix thoroughly.
6. Add 200 ul restriction enzyme mixture to each tube. Close tube and mix by tapping gently.
7. Incubate sample and control plug slices in 37º C water bath for 1.5 hrs.
8. Cast agarose gel by preparing 1% SeaKem Gold agarose in 0.5X TBE buffer. Melt agarose completely and place in 50º C water bath until use.
9. Carefully pour agarose into level gel form fitted with comb.
10. Place black gel frame in electrophoresis chamber. Add 2.2 L freshly prepared 0.5X TBE.
11. Turn on cooling unit (14º C) and pump (set at 70 for 1 liter/min flow rate).
12. Remove restricted plug slices from 37º C water bath. Remove enzyme/buffer mixture and add 200 ul 0.5X TBE. Incubate at room temperature for 5 min.
13. Remove comb after gel solidifies for 30 min.
14. Remove restricted plug slices from tubes and load into appropriate wells with spatula. Make sure there are no air bubbles.

15. Cover plugs and fill wells with remaining melted 1% SeaKem Gold agarose. Allow to harden for 5 min. Unscrew end gates from form; remove excess agarose from sides and bottom of stand with a Kimwipe. Carefully place gel inside black gel frame in electrophoresis chamber. Close cover of unit.
16. Run electrophoresis user program 1 in unit.

Day 3 - Restriction digestion

1. When electrophoresis is over, turn off equipment and stain gel with Ethidium Bromide (0.5 ug/ml) made in 400 ml of 0.5X TBE. Stain for 30 min with agitation.
2. Pour off Ethidium Bromide solution and place gel in 400 ml of 0.5X TBE to destain for 15 min. After 15 min. pour off TBE and add fresh 0.5X TBE and continue destaining for 15 min.
3. Place gel on UV light box and photograph using Kodak CDC and software. Save digital photograph and print hardcopy for notebook.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX J
LABORATORY PROTOCOL FOR ANTIBIOTIC RESISTANCE ANALYSIS BY THE KIRBY-BAUER DISK DIFFUSION METHOD

THIS PAGE INTENTIONALLY LEFT BLANK

The Kirby Bauer Disk Diffusion Method is described in the following document from the NCCLS: *Performance Standards for Antimicrobial Disk and Dilution Susceptibility Tests for Bacteria Isolated from Animals; Approved Standard-2nd edition (2002). M31-A2 Vol. 22 No. 6.* A copy of this method is available upon request.

Cryofrozen samples (a bacterial suspension of 0.4 mL of 80% glycerin and 20% water and 0.6 mL of purely isolated *E. coli* in nutrient broth) in 1.0 mL plastic cryovials were delivered overnight to TAMU-CC, from Texas A&M El Paso Agricultural Research and Extension Center, in styrofoam boxes packed with dry ice. TAMU-CC received shipments of samples on 3/26/2004, 7/29/2004, 8/11/2004, 10/5/2004, 1/19/2005, 2/4/2005, 4/15/2005, and 6/27/2005. The samples were immediately inventoried and the date, time, sample identification numbers, temperature, and signature of the receiver were recorded on laboratory sample reception logs. The samples were then stored at -80° C. for later analysis.

Before antibiotic resistance analysis, the frozen isolates were aseptically transferred from the cyrovials onto tryptic soy agar (TSA) plates and incubated overnight (18-24 hr) at 35° C. Each isolate was transferred to tryptic soy broth (TSB) and incubated in a 35° C shaker for 2-6 hours. Antibiotic resistance analysis of the isolates was performed by the standard Kirby Bauer Disk Diffusion method, with a panel of 20 antibiotics (Table J-1), following NCCLS Performance Standards (2000, 2002a, 2002b). Two Mueller Hinton Agar plates were used per isolate, with 10 antibiotics per plate. After incubation the plates were analyzed with an automated plate reader, the BIOMIC® system, which uses color digital image analysis to provide instantaneous reading of inhibition zone diameters and interpretation following NCCLS M100 (2002b). The BIOMIC® system includes software to determine whether each isolate is resistant, intermediate or susceptible (R-I-S) based on published NCCLS guidelines (TableJ-2). The automated image analyzer ensured uniformity for future comparisons with *E. coli* isolates from unknown sources as detailed in the TAMU-CC SOP following NCCLS (2002a), approved in the QAPP for the project (2003). The results were stored electronically in the BIOMIC® system database, and as hard copy in binders, with back up CDs.

American Public Health Association. 1998. Standard methods for the examination of water and wastewater. 20th ed. American Public Health Association, Washington D.C.

NCCLS (2000) Performance Standards for Antimicrobial Disc Susceptibility Tests; Approved Standard-Seventh Edition. NCCLS document M2-A7.

NCCLS (2002a) Performance Standards for Antimicrobial Disk and Dilution Susceptibility Tests for Bacteria Isolated from Animals; Approved Standard-Second Edition. NCCLS document M31-A2.

NCCLS (2002b) Performance Standards for Antimicrobial Susceptibility Testing; Twelfth Informational Supplement. NCCLS document M100-S12

Table J-1 Antibiotics used to develop antibiotic resistance profiles for *E. coli* isolates

Antibiotic	Abbreviation	Concentration
Ampicillin	AMP	10 µg
Augmentin	AMC	30 µg
Cefazolin	CZ	30 µg
Cefotaxime	CTX	30 µg
Ceftazidime	CAZ	30 µg
Ceftriaxone	CRO	30 µg
Chloramphenicol	C	30 µg
Ciprofloxacin	CIP	5 µg
Doxycycline	D	30 mg
Enrofloxacin	ENO	5 µg
Gentamicin	GM	10 µg
Imipenem	IPM	10 µg
Kanamycin	K	30 µg
Nalidixic acid	NA	30 µg
Neomycin	N	30 µg
Spectinomycin	SPT	100 µg
Streptomycin	S	10 µg
Sulfamethoxazole Trimethoprim	SXT	23.75/1.25 µg
Sulfisoxazole	G	0.25 mg
Tetracycline	Te	30 µg

Table J-2 Susceptible (S), Intermediate (I), and Resistant (R) ranges (mm) for *E. coli* using the BIOMIC Microbiology Analyzer System

Antibiotic	S	I	R
AMP	≥ 17	14-16	≤ 13
AMC	≥ 18	14-17	≤ 13
CZ	≥ 18	15-17	≤ 14
CTX	≥ 23	15-22	≤ 14
CAZ	≥ 18	15-17	≤ 14
CRO	≥ 21	14-20	≤ 13
C	≥ 18	13-17	≤ 12
CIP	≥ 21	16-20	≤ 15
D	≥ 16	13-15	≤ 12
ENO	≥ 21	16-20	≤ 15
GM	≥ 15	13-14	≤ 12
IPM	≥ 16	14-15	≤ 13
K	≥ 18	14-17	≤ 13
NA	≥ 19	14-18	≤ 13
N	≥ 17	13-16	≤ 12
SPT	≥ 18	15-17	≤ 14
S	≥ 15	12-14	≤ 11
SXT	≥ 16	11-15	≤ 10
G	≥ 7	NA	≤ 6
TE	≥ 19	15-18	≤ 14

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX K
BST WORKSHOPS STAKEHOLDERS LIST

THIS PAGE INTENTIONALLY LEFT BLANK

BST WORKSHOP ATTENDEE LIST

June 27, 2003

AGENCY	NAME	TITLE
Project Team		
Texas Farm Bureau	Ned Meister	Director, Commodity & Regulatory Activities
Parsons	Mel Vargas	Project Manager
Parsons	Kirk Dean, Ph.D.	Technical Manager
TSSWCB	Kevin Wagner	Project Manager
EP AREC, Texas A&M	George Di Giovanni, Ph.D.	Assoc. Prof, Environmental Microbiology
Texas A&M College Station	Suresh Pillai, Ph.D.	Assoc. Prof., Food and Environmental Microbiology
Texas A&M Corpus Christi	Joanna Mott, Ph.D.	Assoc. Prof, Physical & Life Sciences
City of Waco	Tom Conry	Program Administrator, Lab & Watershed Programs
Brazos River Authority	Kyle Headley	Regional Environmental Planner, Central Basin
State Agencies		
TCEQ	Margaret Hoffman	Executive Director
TCEQ	Mason Miller	TMDL Technical Director
TDH	James Morgan	Regional Director, M.D.
TPWD	Pat Radloff	Water Quality Team Leader
TX Dept Ag.	Bo Spoonts	
TX Dept Ag.	Lisa Eldridge	
TXDOT	Barrie Cogburn	Landscape Architect
Universities		
Texas A&M	Allan Jones	Director, TX Water Resources
TIAER	Larry Hauck	Assoc. Dir. Environ. Sciences
Baylor	Owen Lind	
Federal Agencies		
USEPA Region 6	Shawneille Campbell	Texas TMDL Coordinator
USEPA Region 6	Philip Crocker	
USDA – NRCS	Tim Buscha	
USDA – NRCS	James Abbott	
USGS	Jess Weaver	District Chief
U.S. Congress	Chet Edwards	Congressman
Ft. Hood Environ	Riki Young	Biologist
USACE	Jeff Tripe	Environmental Planner
Regional Entities		
Brazos River Auth	John Baker	Middle Basin Manager
SWC District-Hamilton-Coryell	B.W. Teague	Chairman
SWC District – Bell	B.G. Welch	Chairman
SWC District – McClennan	Max Sturdivant	Chairman
SWC District – Bosque	Phillip Munden	Chairman

AGENCY	NAME	TITLE
SWC District – Erath	Donald Smart	Chairman
SWC District – Comanche/Eastland/Erath	Norman Moore	Chairman
Local Citizens		
Local Citizens	Truman Blum	
Local Citizens	John Merrill	TX & SW Cattle Raisers Assn
Local Citizens	John Hatchell	
Dairy Farmers and Ranchers by Watershed		
Dairy Farmers of America	John Cowan	
Farm Bureau – McLennan Co.	Robert Rush	President
Farm Bureau – Bell County	Robert Fleming	President
Farm Bureau – Bosque County	Alan Day	President
Farm Bureau – Coryell County	Neil Walter	President
Farm Bureau – Hamilton County	Rusty Harris	President
Farm Bureau – Erath County	Jimmy Holeman	President
Cities by Watershed		
Watershed – Leon River – Gatesville	Brandon Emmons	City Manager
Watershed – Leon River – Temple/Belton	Mark Watson	City Manager
Watershed – Bosque – Waco	Ricky Garrett	Water Treatment Superintendent
Watershed – Bosque – Clifton	Jerry Golden	City Administrator
Watershed – Bosque – Stephenville	John Moser	Mayor
Watershed – Bosque – Meridian	Marie Garland	Public Works Director
County Judges		
County Judge – Bosque	Cole Ward	County Judge
County Judge – Erath	Tab Thompson	County Judge
County Judge – Coryell	John Hull	County Judge
County Judge – Comanche	James Arthur	County Judge
County Judge – McLennan	Jim Lewis	County Judge
County Judge – Bell	Jon Burrows	County Judge
County Judge – Hamilton	Fred Cox	County Judge

BST WORKSHOP ATTENDEE LIST

October 27, 2005

AGENCY	NAME	TITLE
Project Team		
Texas Farm Bureau	Ned Meister	Director, Commodity & Regulatory Activities
Parsons	Mel Vargas	TMDL Technical Director
Parsons	Kirk Dean	Principal Scientist
TSSWCB	T.J. Helton	NPS Grant Coordinator
EP AREC, Texas A&M	Elizabeth Casarez, Ph.D.	Postdoctoral Res. Assoc., Environmental Microbiology
Texas A&M College Station	Suresh Pillai, Ph.D.	Professor, Food and Environmental Microbiology
Texas A&M Corpus Christi	Joanna Mott, Ph.D.	Assoc. Prof, Physical & Life Sciences
City of Waco	Tom Conry	Program Administrator, Lab & Watershed Programs
Brazos River Authority	Dr. John Ellis	Regional Environmental Planner, Central Basin
State Agencies		
TCEQ	Mark Vickery	Deputy Executive Director
TCEQ	Jim Davenport	Water Quality Standards
TCEQ	Faith Hambleton	TMDL Program Manager
TCEQ	Robbie Ozment	TCEQ Region 9 Water Quality Specialist
DSHS	James K. Morgan, MD, MPH	Public Health Regional Director
TPWD	Pat Radloff	Water Quality Team Leader
TX Dept Agriculture	Richard Eyster	Hydrogeologist
TXDOT	Barrie Cogburn	Landscape Architect
Universities		
Texas A&M	Allan Jones	Director, TX Water Resources
TIAER	Larry Hauck	Assoc. Dir. Environmental Sciences
Baylor	Renee Massengale	Associate Professor
Baylor	Owen Lind, Ph.D.	Professor of Biology
Federal Agencies		
USEPA Region 6	Shawneille Campbell	TMDL Coordinator
USEPA Region 6	Philip Crocker	Water Quality Assessment Sect. Manager
USDA – NRCS	Claude Ross	State RC&D Coordinator
USGS	Bob Joseph	District Chief
USGS	Ann Ardis	Hydrologist
USGS	Michael Canova	Hydrologist
U.S. Congress	Chet Edwards	Congressman
Texas State Senate	Kip Averitt	Texas State Senator, District 22
Ft. Hood Environ	Roderick Chisholm	Director of Public Works – Ft. Hood
USACE	Jeff Tripe	Environmental Planner
USACE	Becky Griffith	Supervisory Regional Economist

AGENCY	NAME	TITLE
Regional Entities		
Brazos River Auth	John Baker	Middle and Lower Basin Manager
SWC District-Hamilton-Coryell	P M Gerald, Jr.	Chairman
SWC District – McClennan	Max Sturdivant	Chairman
SWC District – Bosque	J Charlie Blue	Chairman
SWC District – Central Texas	Stanley Glaser	Chairman
SWC District – Cross Timbers	Donald Smart	Chairman
SWC District – Upper Leon	Norman Moore	Chairman
Local Citizens		
Local Citizens	Truman Blum	BRA Board Member
Local Citizens	John Merrill	TX & SW Cattle Raisers Assn.
Dairy Farmers and Ranchers by Watershed		
Dairy Farmers of America	John Cowan	Legislative Environmental Manager
Farm Bureau – McLennan Co.	Jimmy Westerfeld	President
Farm Bureau – Bell County	Robert Fleming	President
Farm Bureau – Bosque County	Alan Day	President
Farm Bureau – Coryell County	Neil Walter	President
Farm Bureau – Hamilton County	Rusty Harris	President
Farm Bureau – Erath County	Paul Tyus	President
Farm Bureau – Comanche County	Tommy Elliott	President
Texas Farm Bureau	Darren Turley	Chairman, Dairy Advisory Committee
Cities by Watersheds		
Watershed – Leon River – Gatesville	Brandon Emmons	City Manager
Watershed – Leon River – Temple/Belton	David Blackburn	City Manager
Watershed – Bosque – Waco	Ricky Garrett	Director, Utility Services
Watershed – Bosque – Clifton	Jerry Golden	City Administrator
Watershed – Bosque – Clifton	Jim Burch	Director, Public Works
Watershed – Bosque – Clifton	Leon Smith	Mayor
Watershed – Bosque – Stephenville	Rusty Jergins	Mayor
Watershed – Bosque – Meridian	Marie Garland	Public Works Director
County Judges		
County Judge – Bosque	Cole Ward	Judge
County Judge – Erath	Tab Thompson	Judge

AGENCY	NAME	TITLE
County Judge – Coryell	John Hull	Judge
County Judge – Comanche	James Arthur	Judge
County Judge – McLennan	Jim Lewis	Judge
County Judge – Bell	Jon Burrows	Judge
County Judge – Hamilton	Fred Cox	Judge

THIS PAGE INTENTIONALLY LEFT BLANK