Assessment of Water Quality and Watershed Planning for the Leona River - *Bacterial Source Tracking*

TSSWCB Project 11-50

Prepared by:

Emily C. Martin
Terry J. Gentry

Funding provided by
Texas State Soil and Water Conservation Board

TEXAS A&M AgriLIFE

Teaching • Research • Extension • Service

TABLE OF CONTENTS

Background 1
Technical Approach 2
Sample Collection and Processing 4
Bacterial Source Tracking 4
  Library-Independent BST 4
  Library-Dependent BST 4
Results 7
  Samples Processed for BST 7
  Library-Independent BST Results 8
  Library-Dependent BST Results 10
Summary and Discussion 11
References 13

LIST OF FIGURES

Figure 1. Monitoring sites on the Leona River and tributaries 4

Figure 2. *Bacteroidales* PCR marker occurrence in stream samples (n=42) 9

Figure 3. *Bacteroidales* PCR marker occurrence in fish hatchery discharge samples (n=28) 10

Figure 4. Identification of *E. coli* isolates (n=77) from the Leona River watershed using a 3-way split for source classification 11

Figure 5. Identification of *E. coli* isolates (n=77) from the Leona River watershed using a 7-way split for source classification 12

LIST OF TABLES

Table 1. Water samples processed for BST analysis 3

Table 2. Texas *E. coli* BST Library (ver. 6-13, cross-library validation) composition and rates of correct classification (RCCs) by Jackknife analysis of ERIC-RP composite data sets using an 80% similarity cutoff and three and seven-way splits 7

LIST OF ACRONYMS AND ABBREVIATIONS

ARCC – Average Rate of Correct Classification

AU – Assessment Unit

BST – Bacterial Source Tracking

DNA – Deoxyribonucleic Acid

E. coli – *Escherichia coli*

ERIC - Enterobacterial Repetitive Intergenic Consensus Sequence

mTEC - Modified Membrane Thermotolerant *E. coli* Medium

MUG - Methylumbelliferyl-b-D-glucuronide

NA – Nutrient Agar

PCR – Polymerase Chain Reaction

RP - RiboPrinting

SAML – Soil and Aquatic Microbiology Laboratory

SOP – Standard Operating Procedure

TIAER – Texas Institute for Applied Environmental Research, Stephenville, TX

TSSWCB – Texas State Soil and Water Conservation Board

WWTF – Wastewater Treatment Facility

USEPA – United States Environmental Protection Agency

UTSPH-EP – University of Texas School of Public Health, El Paso

Background

The Leona River (Segment 2109) is a tributary of the Frio River within the Nueces River Basin. The river flows about 90 miles from US 83 in Uvalde County, through Zavala County, then to its confluence with the Frio River in Frio County. The watershed is approximately 429,000 acres. Cities within the watershed include Uvalde in Uvalde County and Batesville in Zavala County, both of which have wastewater discharge permits to the river. The Leona River watershed is rural and land use is predominantly agriculture, including cropland and pastureland. According to the USDA NASS 2007 Census of Agriculture, approximately 2.4 million acres of land in Frio, Uvalde, and Zavala counties are farmland. Leading animal operations that exist in all three counties are beef cattle and sheep. Winter wheat production, oats, sorghum, and cotton are among the leading crops harvested in all three counties. Large amounts of land are also used to grow forages such as hay, grass silage, and greenchop in Uvalde and Frio Counties, and Frio County had more than 11,600 acres in peanut production in 2007. While mainly rural, the cities of Uvalde and Batesville are located within the watershed. Uvalde has an estimated population of 16,000, while about 1,300 people reside in Batesville.

The Leona River was first listed as having a bacteria impairment for contact recreation in the *2006 Texas Water Quality Inventory and 303(d) List*. It was listed as having a concern for bacteria in prior reports. The *2010 and 2012 Texas Integrated Reports* include a bacteria impairment for all three assessment units (AU) within the Leona River.

To assess and identify different sources contributing to bacterial loadings in these waterbodies, Texas A&M AgriLife Research – Department of Soil and Crop Sciences – Soil and Aquatic Microbiology Laboratory (SAML) conducted bacterial source tracking (BST). BST is based on the premise that specific microorganisms are selected for in the gut communities of various warm blooded animals due to differences in their physiology and intestinal environment. This specificity can then be exploited through phenotypic and genetic assays to trace fecal contamination back to its source. SAML performed library-independent BST utilizing the *Bacteroidales* polymerase chain reaction (PCR) genetic test for human, ruminant, horse, and hog markers. The *Bacteroidales* PCR method is a culture-independent molecular method, which targets genetic markers of *Bacteroidales* and *Prevotella* spp. fecal bacteria that are specific to humans, ruminants (including cattle and deer), hogs, and horses (Bernhard and Field 2000a; Bernhard and Field 2000b). Results are typically expressed as presence/absence (incidence) of the host-specific genetic markers; therefore, this method is not quantitative.

In addition, SAML conducted limited library-dependent BST and analyzed *E. coli* isolates utilizing the enterobacterial repetitive intergenic consensus sequence PCR (ERIC-PCR) and RiboPrinting (RP) combination method (ERIC-RP). ERIC-PCR and RP are genetic fingerprinting methods used in previous BST studies as well as many microbial ecology and epidemiological studies (Jones et al. 2009). They generate DNA banding patterns or fingerprints which look similar to barcode patterns. Different strains of *E. coli* bacteria have differences in their DNA sequences and produce different barcode-like patterns. Therefore, the source of an *E. coli* isolate can be

determined by comparing its barcode-like pattern to those in the Texas *E. coli* BST Library (containing ERIC-RP patterns for *E. coli* collected directly from over 1,000 various animal and human sources from throughout Texas). Isolates collected through this method can be classified as originating from domestic animals (including livestock), domestic sewage, or wildlife (3-way split) or further classified as originating from cattle, avian livestock, other non-avian livestock, avian wildlife, non-avian wildlife, domestic sewage, or pet sources (7-way split).

Technical Approach

Water samples were collected by TIAER beginning in July 2011 thru May 2013 (Table 1). A total of 23 sampling locations were located within the three assessment units (Figure 1). Due to a lack of rainfall, 11 of the original sampling locations were not flowing or did not have substantial pools when sites were monitored during the study (see McFarland et al., 2013). The majority of water samples collected was wastewater treatment plant and fish hatchery discharge at the sampling location just downstream of both of those outfalls in AU 03. Sampling locations in the BST analysis from AU 01 included main stem sites on the Leona River at the confluence of the Frio River (Site ID 21044) and at FM 1581 (Site ID 12985); AU 02 included main stem sites on the Leona River at Loma Vista Road (Site ID 12986), at US 57 (Site ID 12987), at FM 1866 (Site ID 21064), and at CR 1005B near the confluence of Camp Lake Slough (Site ID 21066), as well as tributary sites at Live Oak Creek at US 57 (Site ID 21062) and Galina Slough on CR 117 (Site ID 21063); AU 03 included a main stem site on the Leon River at Hoags Dam below the confluence of Cooks Slough (Site ID 12989), a tributary site at Cooks Slough at FM 117 (Site ID 12956), as well as two discharges from the wastewater treatment plant in Uvalde (Site ID WWTF1 and WWTF2) and two discharge sites from the Uvalde National Fish Hatchery (Site ID FH001 and FH002). Of note, FH002 was located at a road crossing about a quarter mile downstream of FH001. The area between FH001 and FH002 crossed a small pasture, which at times was noted to contain cattle.

Table 1. Water samples processed for BST analysis

Parameter (# sites)	2011	2012	2013	Total Collected
	Sept- Dec	Jan - Dec	Jan - May	
Bacteroidales				
Stream (10)	**16**	**18**	**8**	**42**
WWTFs (2)	**8**	**20**	**8**	**36**
Fish Hatchery (2)	**2**	**19**	**8**	**29**
Bacteroidales Total	**26**	**57**	**24**	**107**
***E. coli* (ERIC-RP)**				
Stream (10)	**23**	**24**	**6**	**53**
Fish Hatchery (2)	**1**	**15**	**8**	**24**
E. coli Total	**24**	**39**	**14**	**77**

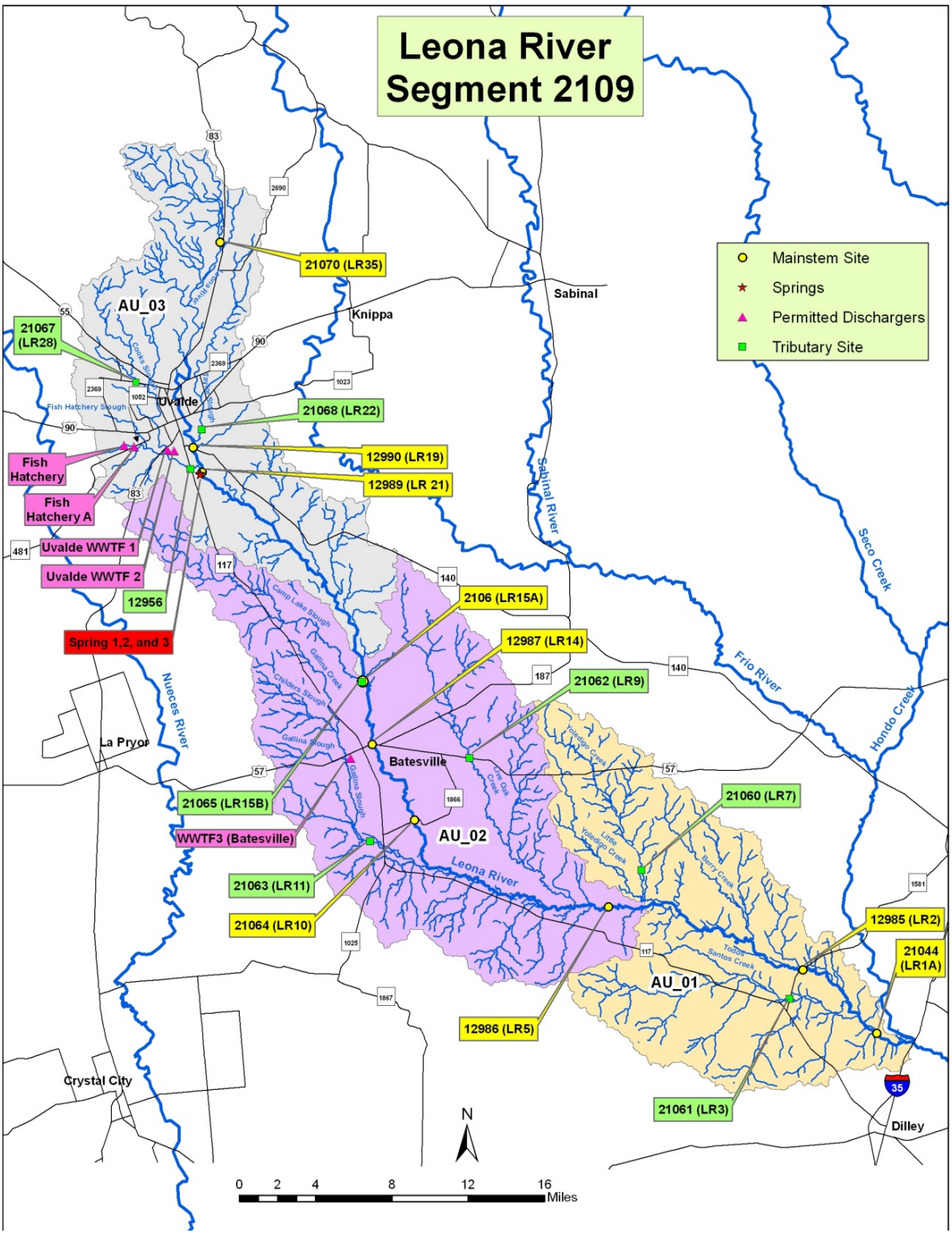

Figure 1. Monitoring sites on the Leona River and tributaries.

Sample Collection and Processing

TIAER collected and processed the water samples (Table 1) for BST analysis within 8 hours of sample collection using standard operating procedures (SOPs) as developed by the University of Texas School of Public Health, El Paso (UTSPH-EP). Project SOPs for BST are provided in the Quality Assurance Project Plan (TIAER, 2013). For *E. coli* isolations, water samples were processed using USEPA Method 1603 and modified membrane thermotolerant *E. coli* (mTEC) medium (USEPA, 2005). Within 48 hours of processing, mTEC plates were shipped overnight to SAML for isolation. *E. coli* colonies were then picked from the modified mTEC medium and streaked onto nutrient agar with MUG (NA-MUG) in order to confirm culture purity. Cultures of selected isolates were archived at -80^{o}C for subsequent BST analyses. For *Bacteroidales* PCR, water samples were filtered, by TIAER, in order to recover bacterial biomass which was then archived at -80^{o}C and shipped to SAML for analysis.

Known-source fecal samples were also collected by TIAER and Nueces River Authority staff and shipped overnight to SAML for processing. *E. coli* were isolated from these fecal samples and processed and archived using USEPA Method 1603 and UTSPH-EP SOPs, as described above for the water samples. In general, one isolate was fingerprinted per fecal sample using ERIC-RP and compared using densitometric curve-based Pearson-product similarity coefficients. Isolates deemed source-specific through self-validation (described below) were added to the Texas *E. coli* BST Library.

Bacterial Source Tracking

Library-Independent BST

Bacteroidales PCR was conducted using UTSPH-EP SOPs. Microbial DNA was extracted from the archived filters and purified. An aliquot of the DNA was then analyzed by PCR for markers specific to humans, ruminants (including cattle, deer, and sheep), hogs (including feral hogs), and horses, in addition to a general marker which detects the *Bacteroidales* order as a whole and is not source specific (Bernhard and Field 2000a; Bernhard and Field 2000b; Dick et al. 2005). For this study, qualitative presence/absence of the host-specific genetic markers was determined; this effectively means that there either was or was not bacteria of a specific type present in the water sample.

Library-Dependent BST

Both ERIC-PCR and RP were performed as previously described (Casarez et al., 2007). *E. coli* isolates were first DNA fingerprinted using ERIC-PCR (Versalovic et al., 1994). Following ERIC-PCR analysis, *E. coli* isolates were Riboprinted using the automated DuPont Qualicon RiboPrinter® system and the restriction enzyme *Hind*III. Analysis of composite ERIC-RP DNA fingerprints was performed using Applied Maths BioNumerics software (Casarez et al., 2007).

Collecting known source fecal samples were a significant portion of the BST efforts in order to add Leona watershed specific isolates into the Texas *E. coli* BST library. Of the 260 total known source fecal samples collected and processed from the watershed, *E. coli* were successfully isolated from 201 individual samples. All 201 of these isolates (one isolate per known source sample) were screened using ERIC-RP and included in the local watershed library. Jackknife analysis of the ERIC-RP was used to identify isolates that correctly classified using a 7-way split of source classes (i.e., human, pets, cattle, other non-avian livestock, avian livestock, avian wildlife, and non-avian wildlife). Isolates with unique fingerprints (left unidentified using an 80% similarity cutoff) were also included to create the local self-validated library. In total, 94 isolates were self-validated in the local library.

The 94 local, self-validated source isolates from the watershed were then added to the current library of Texas *E. coli* BST self-validated source isolates from twelve previous watershed projects across Texas. The Texas *E.coli* BST library represents thousands of archived and screened known source samples. A series of Jackknife analyses were run on the combined libraries, removing all isolates that cross-identified between human, domestic animals, and wildlife. After each removal, the Jackknife was run again with the goal of 100% average rate of correct classification (ARCC) using a 3-way split of source classes. The updated library began with the 1,807 isolates of the combined self-validated local watershed libraries. After four iterations of cross-watershed validation, the resulting Texas *E. coli* BST Library (ver. 6-13) contained 1,524 isolates from 1,358 samples, resulting in a 100% ARCC with a 3-way split of source classes and a 92% ARCC using the 7-way split of source classes (Table 2). A total of 20% of the isolates were left unidentified as their best match in the library was less than the 80% similarity cutoff, but they were still included in the library in order to reflect the diversity of patterns potentially seen in unknown water samples (Table 2). After cross-watershed validation, 77 of the local library isolates from the Leona known source samples (82% or 77 of the original 94 local library isolates) were included in the Texas *E. coli* BST Library (ver. 6-13). These 77 isolates were comprised of individual fecal samples from cattle (15); avian livestock, including yard chickens (4); non-avian livestock including sheep, goats, and horses (12); pets (3); non-avian wildlife including feral hogs, coyote, and deer (36); and avian-wildlife including small birds (7).

This version of the statewide library was used to identify the source classes for water isolates in the watershed. If a water isolate was not at least 80% similar to a library isolate, it was considered to be unidentified. Although fingerprint profiles were considered a match to a single entry, identification was to the host source class and not to the individual animal represented by the best match. Water isolates were identified in a 3-way split as domestic animals (including livestock and pets), domestic sewage, and wildlife (3-way split) and as a more detailed, 7-way split as cattle, avian livestock, non-avian livestock, avian and non-avian wildlife, domestic sewage and pet sources.

Table 2. Texas *E. coli* BST Library (ver. 6-13, cross-library validation) composition and rates of correct classification (RCCs) by Jackknife analysis of ERIC-RP composite data sets using an 80% similarity cutoff and three and seven-way splits.

Source Class	Number of Isolates	Number of Samples	Library Composition and Expected Random Rate of Correct Classification*	Calculated Rate of Correct Classification (RCC)	RCC to Random Ratio***	Unidentified (unique patterns)
HUMAN	**364**	**315**	**24%**	**100**	**4.2**	**22**
DOMESTIC ANIMALS	**531**	**474**	**35%**	**100**	**2.9**	**19**
Pets	86	76	6%	83	13.8	40
Cattle	237	207	16%	93	5.8	11
Avian Livestock	96	83	6%	89	14.8	25
Other Non-Avian Livestock	112	108	7%	90	12.9	14
WILDLIFE	**629**	**569**	**41%**	**100**	**2.4**	**19**
Avian Wildlife	239	221	16%	85	5.3	21
Non-Avian Wildlife	390	348	26%	92	3.5	17
Overall	**1524**	**1358**		**ARCC** ** **= 100% (3-way) 92% (7-way)**		**20% (3-way)**

*RARCC, expected random average rate of correct classification

**ARCC = average rate of correct classification: the proportion of all identification attempts which were correctly identified to source class for the entire library, which is similar to the mean of the RCCs for all source classes when the number of isolates in each source class is similar

***An RCC to Random Ratio greater than 1.0 indicates that the rate of correct classification is better than random. For example, the rate of correct classification for human is 4.0-fold greater than random chance.

Results

Samples Processed for BST

A total of 107 water samples were assayed using *Bacteroidales* PCR and 77 *E. coli* isolates were assayed using ERIC-RP (Table 1). Even though the proposed sampling regime and RUAA were based on three AUs, a large portion of the samples were collected in the northern part of the watershed. In total, 88 of 107 *Bacteroidales* samples (82%) were collected from AU 03, which included the fish hatchery as well as City of Uvalde wastewater treatment plant samples. *Bacteroidales* results are shown as a percentage of positive samples for the watershed for all of the stream samples (n=42), fish hatchery discharge (n=29), and WWTF discharge (n=36).

For the *E. coli* isolates, limited library-dependent BST (ERIC-RP) was utilized for this project and a relatively limited number of isolates were identified; therefore, there is not sufficient data to analyze these results for each AU individually. Instead, the ERIC-RP results were summarized across the entire study area and include both stream isolates as well as fish hatchery discharge isolates. Each water sample processed and having archived *E. coli* had at least one isolate identified. Further, some sampling sites had only one sampling event collected over the course of the project. Three isolates were identified from these single event locations and include sites 12985, 12986, and 21063 from October 2011, as well as site 21062 from May 2012. Additional isolates were collected from the water samples and archived, but were not processed. These isolates could be analyzed in the future should it be decided that more extensive library-dependent BST is required to characterize the sources in the study area. The source identifications of *E. coli* isolates, based upon the Texas *E. coli* BST Library, is presented using a 3-way split and a 7-way split for all samples.

It is valid to compare the *E. coli* and *Bacteroidales* BST results as they are complementary techniques; however, it is important to note that identified pollution source classes are not identical. They are derived utilizing two different methods. For example, one of the *E. coli* source classes is domestic animals, which includes cattle but not deer, while the *Bacteroidales* ruminant marker includes both of these animal sources.

Library-Independent BST Results

Bacteroidales PCR marker occurrence for all stream samples (n=42) is shown in Figure 2. There were 3 samples from AU 01, 16 from AU 02, and 23 from AU 03. The general marker was detected in 96% of samples, the ruminant marker was detected in 19% (n=8) of samples, the hog marker was detected in 2% (n=1) of samples, and the human and horse markers were not detected in any of the 42 total samples.

Bacteroidales PCR marker occurrence for all fish hatchery discharge samples (n=29) is shown in Figure 3. This site is upstream of the WWTF sites north of the City of Uvalde. The general marker was detected in 100% of samples. The only other marker detected was the ruminant marker, which occurred in 10% of samples.

WWTF discharge samples (n=36) were also included in the sampling regime (data not shown). The human marker was detected in 53% of samples (n=19). The ruminant marker was the only other marker detected in WWTF discharge samples and was found in 5 samples or 15% of the total. WWTF samples were the only samples where the human marker was detected across the entire study.

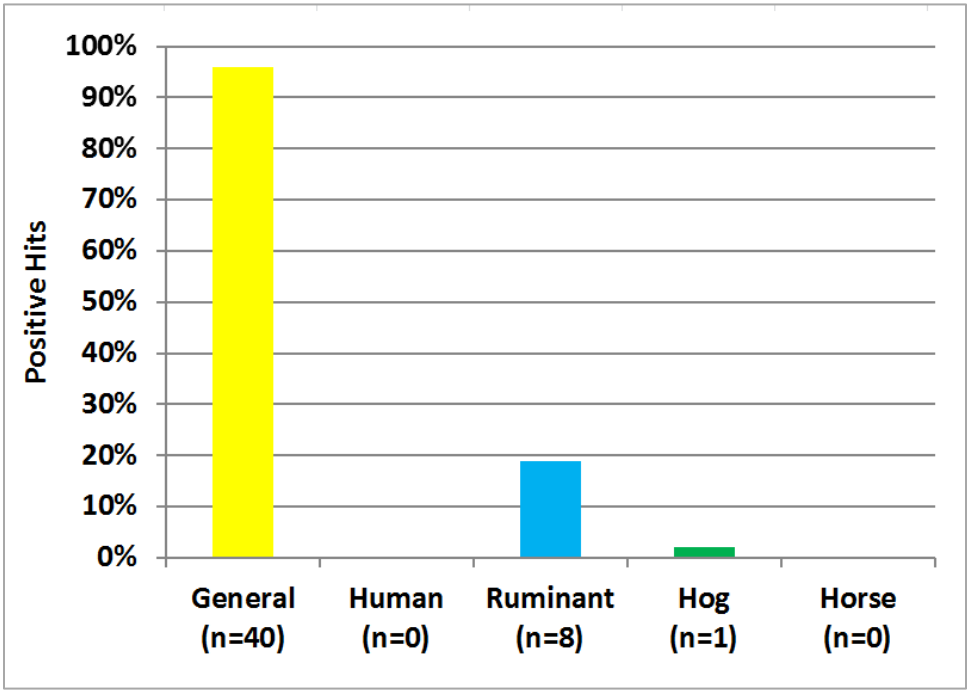

Figure 2. *Bacteroidales* PCR marker occurrence in stream samples (n=42).

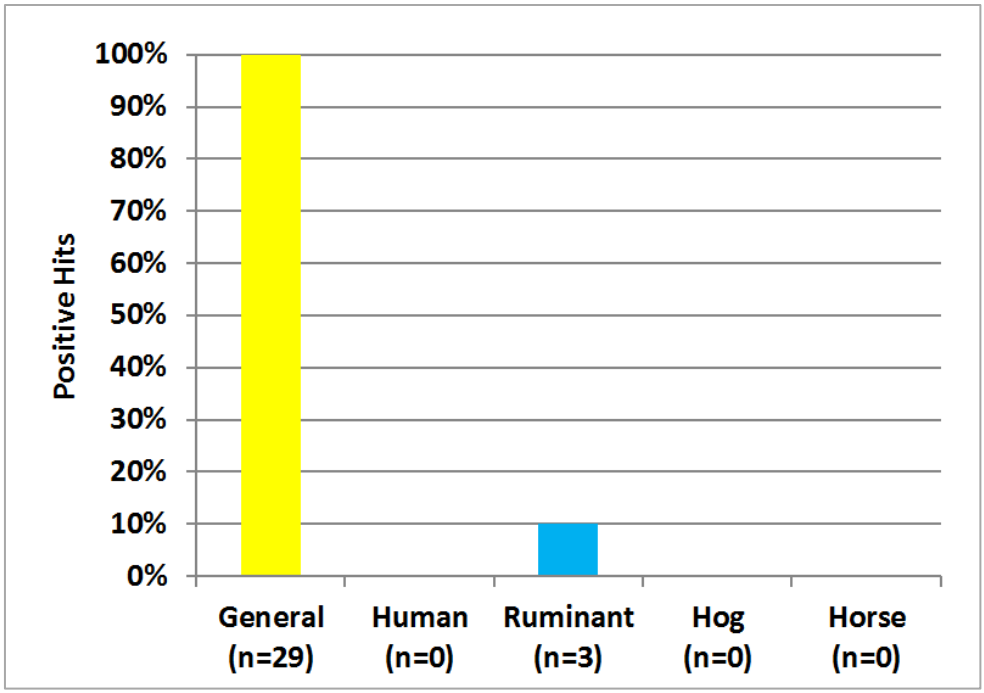

Figure 3. *Bacteroidales* PCR marker occurrence in fish hatchery discharge samples (n=28).

Library-Dependent BST Results

In total, 77 water *E. coli* isolates were classified using the Texas *E. coli* BST library. Using a 3-way split, 55% of the isolates (n=42) classified as originating from wildlife sources, followed by 32% (n=25) from livestock sources, and finally 5% (n=4) from human sources (Figure 4). The originating source could not be found for 8% (n=6) of the isolates.

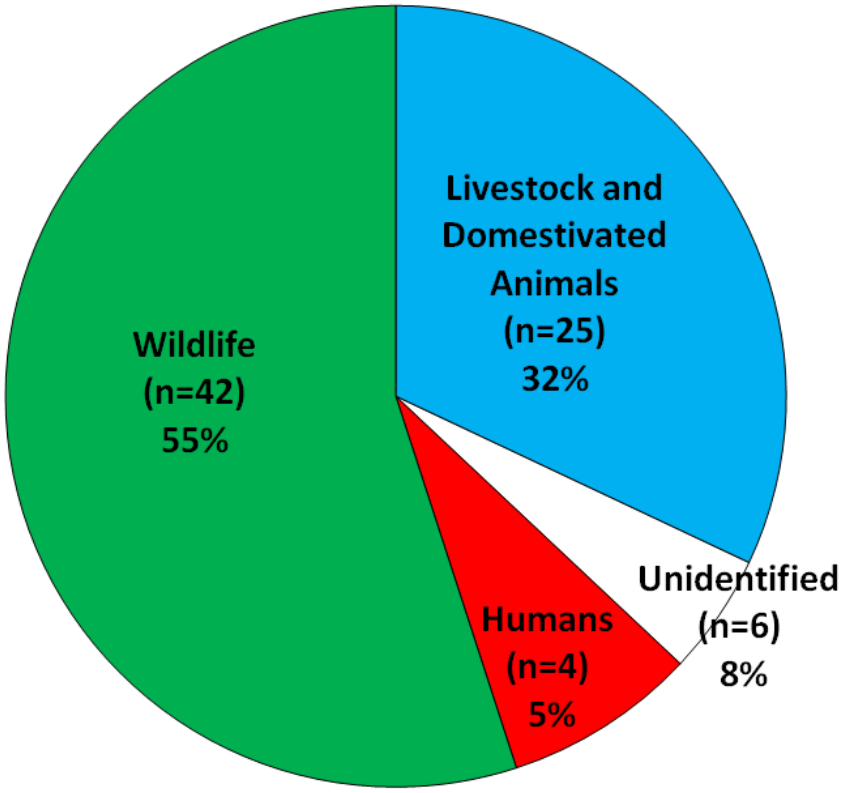

Figure 4. Identification of *E. coli* isolates (n=77) from the Leona River watershed using a 3-way split for source classification.

In the more detailed 7 –way split, the majority of the wildlife isolates were characterized as non-avian wildlife (44%) with 10% from avian wildlife (Figure 5). Avian livestock made up another 14% followed closely by cattle (13%), humans (5%), other non-avian livestock (3%), and pets (3%).

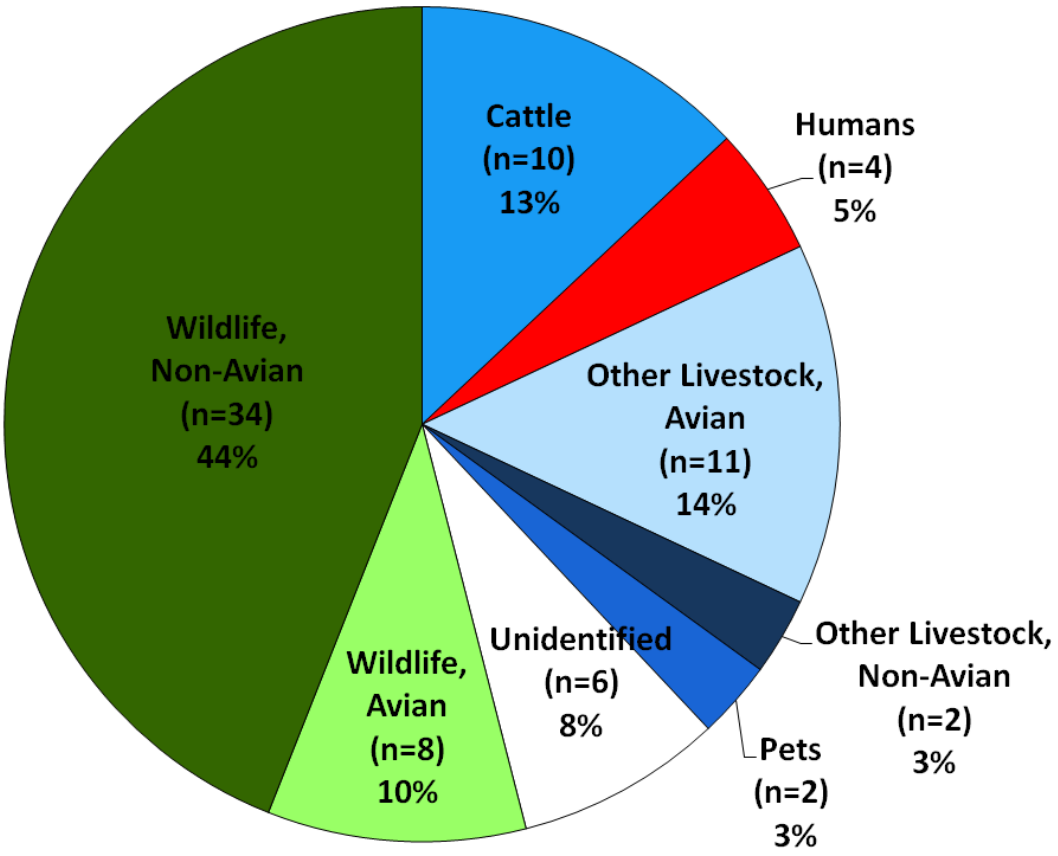

Figure 5. Identification of *E. coli* isolates (n=77) from the Leona River watershed using a 7-way split for source classification.

Summary and Discussion

A combination of library-dependent and -independent BST was utilized to characterize sources of fecal contamination in the Leona River watershed. Severe drought in the watershed during the sampling period limited the extent of samples collected across the watershed. The *Bacteroidales* results yielded very low percentages of source specific hits, but the ruminant marker was detected in 19% of the stream samples and 10% of the fish hatchery samples. The human marker was detected in 53% of the WWTF discharge samples, but was not detected anywhere else in the watershed. The hog marker was only detected in 1 of the 107 total samples analyzed, and the horse marker was not detected in any of the tested samples. Low flow rates and the limited number of sampling locations and stream samples analyzed in part helps explain a lack of source specific marker hits (McFarland, personal communication, provisional data). The general marker was detected in a vast majority of samples, thus, indicating that the *Bacteroidales* order as a whole could be detected. But when the samples were screened with the more specific source markers targeting a much smaller portion of the *Bacteroidales* population, a large portion of the samples assayed did not test positive for a

specific source. Given this result, a subset of these samples was further tested to verify that the negative PCR results were not due to PCR-inhibition in the samples. These samples were spiked with known-source fecal DNA and re-tested with *Bacteroidales* PCR. All of the spiked samples were positive for the tested marker, indicating that the negative *Bacteroidales* PCR results for the original water samples were correct (data not shown).

Library-dependent BST characterization of isolates using the Texas *E. coli* BST library indicated that wildlife were a major source contributor of bacterial contamination. A majority of the isolates were classified as originating from wildlife (55%) and more specifically non-avian wildlife (44%) including feral hogs, coyote, deer, opossums, and raccoons. Livestock, including cattle and avian livestock made up an additional 32% of the isolates. The percentage of avian livestock hits were a bit surprising, but a large percentage of the isolates were from AU 03 at locations south of the City of Uvalde and possibly could be attributed to birds and domesticated waterfowl including ducks and geese in the area. The low number of human-classified *E. coli* isolates (5%) from across the watershed was corroborated with a lack of human *Bacteroidales* hits outside of WWTF discharge. However, any human contributions are important as the likelihood of fecal contamination from human sources containing pathogens may be higher as compared to non-human sources.

It should be noted that the *Bacteroidales*-based PCR and *E. coli*-based ERIC-RP differ in their approach and measure two different microbial populations. The results of the two approaches were similar for this watershed, as *Bacteroidales* PCR detected the presence of ruminant markers in 15% of the total samples and the ERIC-RP characterized 13% as originating from cattle. But it is judicious to discuss differences in the two methods. Cattle represent one of the *E. coli* source classes (7-way split), while the *Bacteroidales* ruminant marker does not discriminate between cattle and other ruminants and thus, would not only detect cattle, but also other ruminants including deer. The *Bacteroidales* PCR approach used in this study also only measures the incidence of detection as opposed to quantifying the relative abundance of different sources, as is done using ERIC-RP. In other words, although ruminant (including cattle) fecal contamination existed in 15% of the samples, it is impossible to say, based on the *Bacteroidales* PCR results, whether each of these positive samples had low or high relative amounts of fecal bacteria originating from ruminants. The ERIC-RP results complement the *Bacteroidales* results by indicating that 13% of the total number of *E. coli* characterized was determined to have originated from cattle sources. Even with these methodological differences, both approaches indicated that wildlife and domesticated animals were the primary sources of fecal bacterial contamination found in the samples analyzed for this project.

References

Bernhard, A. E. and K. G. Field (2000a). Identification of nonpoint sources of fecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from fecal anaerobes. Appl Environ Microbiol 66:1587-1594.

Bernhard, A. E. and K. G. Field (2000b). A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Appl Environ Microbiol 66:4571-4574.

Casarez, E. A., S. D. Pillai, J.B. Mott, M. Vargas, K.E. Dean, and G.D. Di Giovanni (2007). Direct comparison of four bacterial source tracking methods and a novel use of composite data sets. J Appl Microbiol 103:350–364.

Dick, L. K., A. E. Bernhard, T.J. Brodeur, J.W. Santo Domingo, J.M. Simpson, S.P. Walters, and K.G. Field (2005). Host distributions of uncultivated fecal *Bacteroidales* bacteria reveal genetic markers for fecal source identification. Appl Environ Microbiol 71:3184-3191.

Jones, C.A., K. Wagner, G. Di Giovanni, L. Hauck, J. Mott, H. Rifai, R. Srinivasan, and G. Ward (2009). Bacteria total maximum daily load task force final report. TR-341. Texas Water Resources Institute, College Station, TX.

McFarland, A., T. Adams, and J. Stroebel (2013). Hydrology and Water Quality of the Leona River. TR1302. Texas Institute for Applied Environmental Research (report in review with the Texas State Soil and Water Conservation Board).

TIAER (2013). Quality Assurance Project Plan for Assessment of Water Quality and Watershed Planning for the Leona River, TSSWCB Project 11-50, Revision No. 2, prepared by the Texas Institute for Applied Environmental Research for the Texas State Soil and Water Conservation Board (approved May 15, 2013).

USEPA (2005). 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.

Versalovic, J., M. Schneider, F.J. De Bruijn, and J.R. Lupski (1994). Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. Meth Mol Cell Biol 5:25-40.