Texas A&M AgriLife Research | Extension

Texas A&M University

Texas Water Resources Institute TR-456
May 2014

Attoyac Bayou Bacterial Source Tracking Report

Emily C. Martin and Terry J. Gentry, Texas A&M AgriLife Research, Department of Soil and Crop Sciences

Texas Water Resources Institute
make every drop count

Attoyac Bayou Bacterial Source Tracking Report

Funded by:
Texas State Soil and Water Conservation Board: Project 09-10

Investigating Agencies:
Texas A&M AgriLife Research, Department of Soil and Crop Sciences
Texas A&M AgriLife Research, Texas Water Resources Institute

Prepared by:
Emily C. Martin and Terry J. Gentry, Texas A&M AgriLife Research, Department of Soil and Crop Sciences

Texas Water Resources Institute Technical Report TR-456
May 9, 2014

Funding support for this project was provided in part through a Clean Water Act §319(h) Nonpoint Source Grant from the Texas State Soil and Water Conservation Board and the U.S. Environmental Protection Agency

TABLE OF CONTENTS

List Of Figures

List Of Tables

List of Acronyms

ARCC	Average Rate of Correct Classification
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
RCC	Rates of Correct Classification
RP	RiboPrinting
SAML	Soil and Aquatic Microbiology Laboratory
SFASU	Stephen F. Austin State University
SOP	Standard Operating Procedure
TSSWCB	Texas State Soil and Water Conservation Board
WWTF	Wastewater Treatment Facility
UTSPH EP	University of Texas School of Public Health, El Paso

Background

The Attoyac Bayou, a sub-watershed within the Upper Neches River Watershed, extends approximately 82 miles through Rusk, Nacogdoches, San Augustine and Shelby counties before emptying into Sam Rayburn Reservoir. With several rural communities in the area, the majority of the land in the watershed is used for cattle and poultry operations, forestry or recreational and wildlife uses. The bayou is listed as an impaired water body on the Texas Integrated Report for Clean Water Act Sections 305(b) and 303(d) due to high levels of *E. coli*. Three monitoring stations managed by the Angelina & Neches River Authority, U.S. Geological Survey, and Texas Commission on Environmental Quality have provided water quality data on the bayou for a number of years. Beginning in 2000, data collected for *E. coli* have consistently shown elevated *E. coli* levels that exceed the applicable Texas Water Quality Standards. Through the Development of a Watershed Protection Plan for Attoyac Bayou project, additional water quality and stream flow data was collected to better understand *E. coli* loadings to the water body.

To assess and identify different sources contributing to bacterial loadings in these water bodies, 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 (ERIC-RP) patterns. Therefore, the source of an *E. coli* isolate can be determined by comparing its barcode pattern (ERIC-RP) to those in the Texas *E. coli* BST Library (containing ERIC-RP patterns for *E. coli* collected from over 1,000 various animal and human sources from throughout Texas). Isolates can be classified as originating from a 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 Stephen F. Austin State University (SFASU) beginning in August 2010 thru June 2012. A total of 10 sampling locations were allocated across the watershed (Figure 1). Monthly monitoring sites included the 10 stream samples as well as 4 wastewater treatment facilities (WWTF). Stream samples included Attoyac Bayou at FM 138, SH 21, SH 7, US 59 and US 84 respectively, Big Iron Ore Creek at FM 354, Naconiche Creek at FM 95, Terrapin Creek at FM 95, Waffelow Creek at FM 95, and West Creek at FM 2913. WWTF included the Cities of Center and Garrison as well as Martinsville and Chireno ISDs. Two sites were also identified for storm sampling events, Attoyac Bayou at SH 21 and Big Iron Ore Creek at FM 354.

Sample Collection and Processing

SFASU collected and processed the water samples for downstream BST analysis within 8 hours of sample collection using UTSPH EP standard operating procedures (SOPs). For *E. coli* isolations, water samples were processed using USEPA Method 1603 and modified membrane thermotolerant *E. coli* (mTEC) medium (USEPA 2006). 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 SFASU staff and shipped overnight to SAML for processing. *E. coli* were isolated from the fecal samples and processed and archived using US EPA 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 densiometric curve-based Pearson-product similarity coefficients. Isolates deemed source-specific through self-validation (described below) were added to the Texas *E. coli* BST Library.

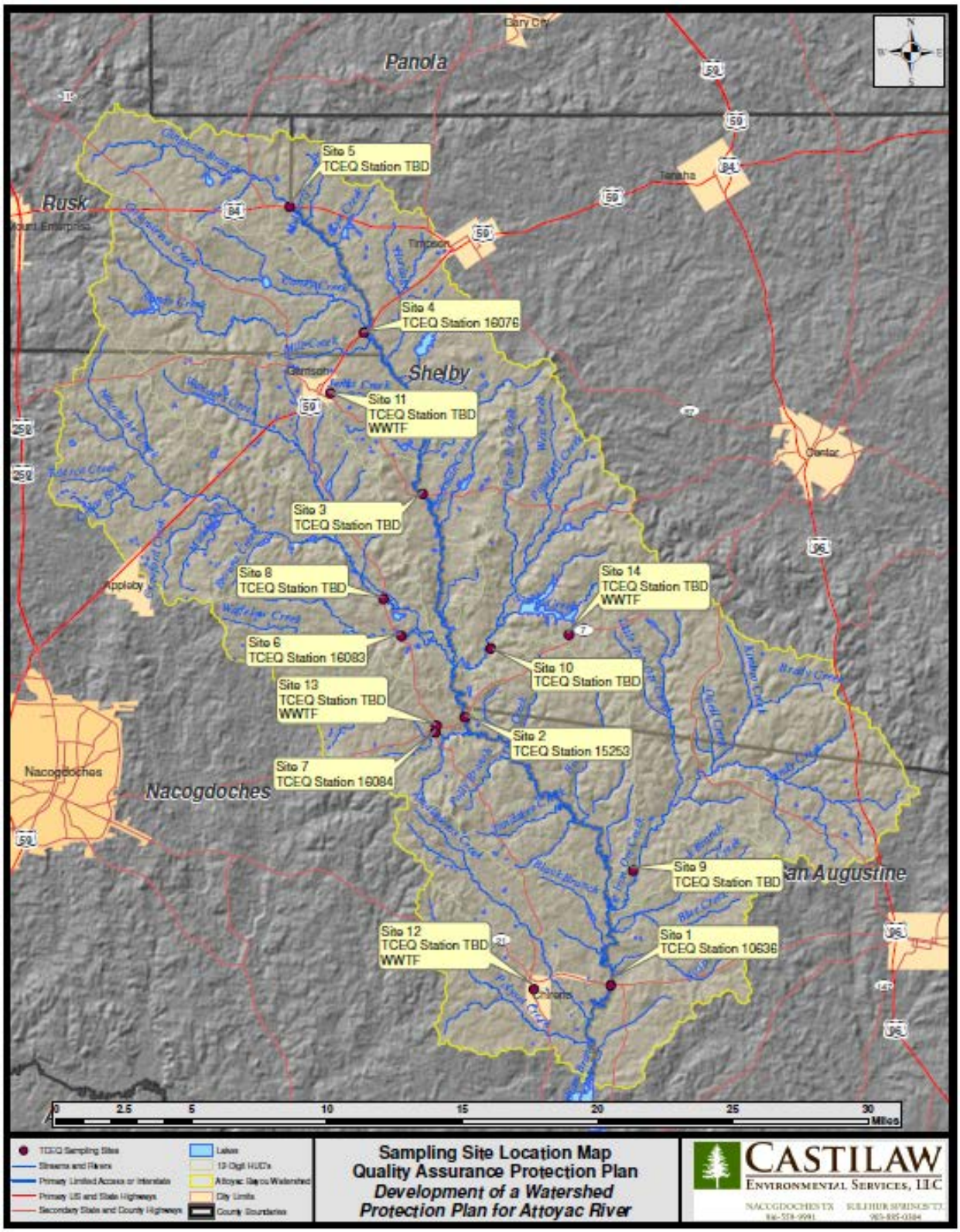

Figure 1. Monitoring sites on the Attoyac Bayou.

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 a specific source (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 by 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 *Hin*dIII. Analysis of composite ERIC-RP DNA fingerprints was performed using Applied Maths BioNumerics software (Casarez et al. 2007).

Known source fecal samples were collected as a portion of the BST efforts to add Attoyac Bayou watershed specific isolates into the Texas *E. coli* BST library. Of the 156 total known source fecal samples collected and processed from the watershed, *E. coli* were successfully isolated from 113 individual samples. All 113 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, 72 isolates were self-validated in the local library.

The 72 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, and 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. After four iterations of cross-watershed validation, the resulting Texas *E. coli* BST Library (ver. 5-13) contained 1454 isolates from 1291 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. A total of 20% of the isolates were identified as singletons (unique fingerprints left unidentified using an 80% similarity cutoff) and were kept in the library in order to reflect the diversity of patterns potentially seen in unknown water samples (See Table 1).

After cross-watershed validation, 59 isolates (82% of the local library samples) were included in the Texas *E. coli* BST Library (ver. 5-13). The 59 isolates were comprised of individual fecal samples from beef cattle (13), raw poultry litter (18), a domestic goose (1), dairy cattle (4), 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 to: domestic animals (including livestock and pets), domestic sewage and wildlife (3-way split) as well as a more detailed, 7-way split, to: cattle, avian livestock, non-avian livestock, avian and non-avian wildlife, domestic sewage, and pet sources.

Table 1. Texas *E. coli* BST Library (ver. 5-13, 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 ***	Left Unidentified (unique patterns)
HUMAN	364	315	25%	100	4.0	22
DOMESTIC ANIMALS	497	442	34%	100	2.9	21
Pets	83	73	6%	88	15.4	41
Cattle	220	192	15%	94	6.2	12
Avian Livestock	93	80	6%	88	13.8	27
Other Non-Avian Livestock	101	97	7%	89	12.8	16
WILDLIFE	593	534	41%	100	2.4	19
Avian Wildlife	232	214	16%	85	5.3	21
Non-Avian Wildlife	361	320	25%	91	3.7	19
Overall	**1454**	**1291**		**ARCC** = 100% 92%**		**20%**

*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/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 267 water samples were assayed using *Bacteroidales* PCR and 104 *E. coli* isolates were assayed using ERIC-RP (Table 2). *Bacteroidales* results are shown as a percentage of positive samples for the watershed for all of the stream samples (n=225) as well as the two storm sampling locations (n=22).

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 at each site individually. Instead, the ERIC-RP results were summarized across the entire study area with stream samples from both base and storm flow conditions. Each water sample processed and having archived *E. coli* had at least one isolate identified. 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. In this study, one isolate's best match identification in the library classified as the same percentage match to two separate source categories and was identified as half an isolate to both categories.

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.

Table 1. Water samples processed for BST analysis

	2010	2011	2012	Total Collected
Parameter (# sites)	**Aug - Dec**	**Jan - Dec**	**Jan - June**	
Bacteroidales				
Stream (10)	44	131	50	225
WWTFs (4)	3	13	4	20
Storm - Stream (2)	0	4	18	22
Bacteroidales Total	**47**	**148**	**72**	**267**
***E. coli* (ERIC-RP)**				
Stream (10)	19	37	30	86
Storm - Stream (2)	0	0	18	18
E. coli Total	**19**	**37**	**48**	**104**

Library-Independent BST Results

Overall *Bacteroidales* Results from Stream Samples

Bacteroidales PCR marker occurrence for all base flow, stream samples (n=225) is shown in Figure 2. The general marker was detected in 96% of samples (n=216), the ruminant marker was detected in 47% (n=105) of samples, the hog marker was detected 28% (n=63) of samples, the human marker was detected in 5% (n=11) of samples, and the horse markers was not detected in any of the 225 total samples.

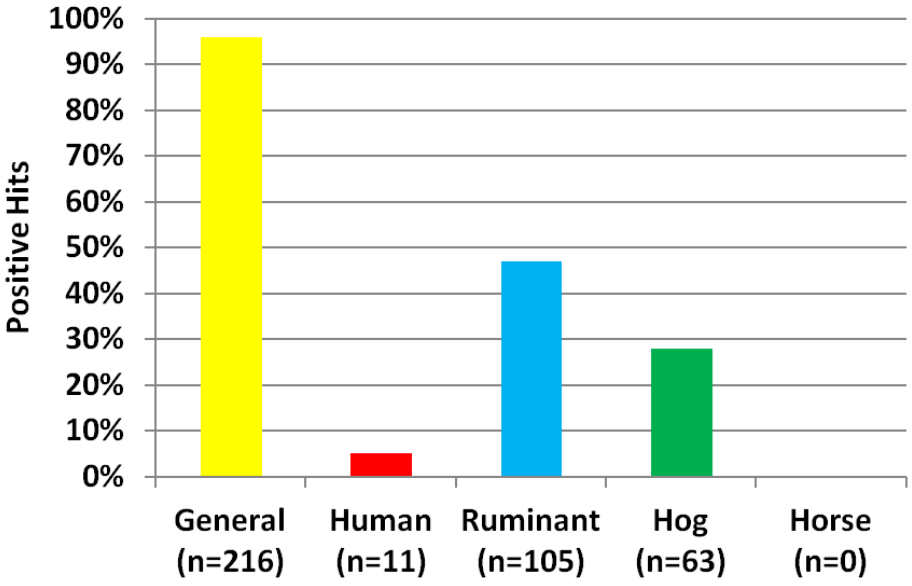

Figure 2. *Bacteroidales* PCR marker occurrence in stream samples (n=225).

Comparison of Bacteroidales Results from Samples Collected under Base-Flow and Storm-Flow

Two sites were selected for further analysis during storm flow events and included Attoyac Bayou at SH 21 and Big Iron Ore Creek at FM 354. Bacteroidales results for each site each with 24 base flow samples and 11 storm flow samples are shown in Figure 3.

For base flow samples at Attoyac Bayou at SH 21, the general marker was detected in 96% (n=23) of samples (data not shown), the ruminant marker was detected in 50% (n=12) of samples, the hog marker was detected in 29% (n=7) of samples, the human marker was detected in 4% (n=1) of samples, and the horse marker was not detected. For storm samples from Attoyac Bayou at SH 21, the general marker was detected in 100% (n=11) of samples (data not shown), the ruminant marker was detected in 91% of samples (n=10), the hog marker was detected in 82% (n=9) of samples, and the human and horse markers were not detected.

For base flow samples at Big Iron Ore Creek at FM 354, the general marker was detected in 92% (n=22) of samples (data not shown), the ruminant marker was detected in 58% of samples (n=14), the hog marker was detected in 54% (n=13) of samples, the human marker was detected in 8% (n=2) of samples, and the horse marker was not detected. For the storm samples at Big Iron Ore Creek at FM 354, the general marker was detected in 100% (n=11) of samples (data not shown), the hog marker was detected in 82% (n=9) of samples, the ruminant marker was detected in 55% (n=6) of samples, while the human and horse markers were not detected.

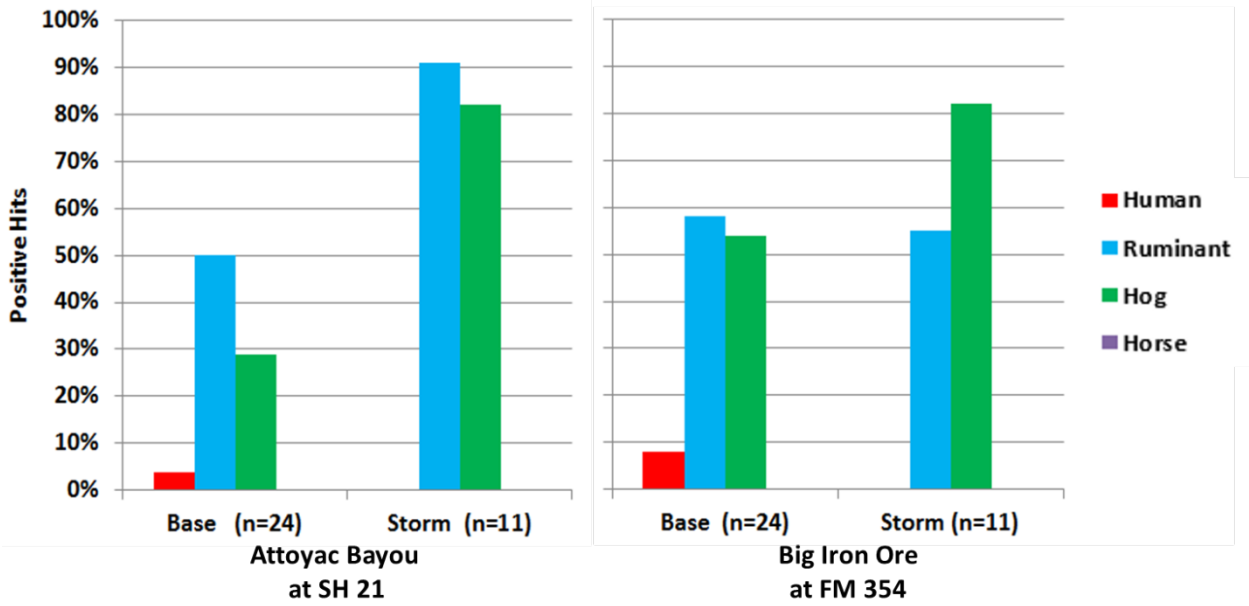

Figure 3. *Bacteroidales* PCR marker occurrence in storm sampling sites, Attoyac Bayou at SH 21 and Big Iron Ore Creek at FM 354. Results are shown for both base flow samples (n=24) on the left and storm flow samples (n=11) on the right for each location.

Library-Dependent BST Results

Overall Identification of *E. coli* from Steam Samples

In total, 104 stream sample water *E. coli* isolates, including both base and storm flow conditions, were classified using the Texas *E. coli* BST library. Using a 3-way split, 61% of the isolates (n=63.5) classified as originating from wildlife sources, followed by 21% (n=21.5) from livestock and domesticated animals, and 5% (n=5) from human sources. The originating source could not be identified for 13% (n=12) of the isolates (Figure 4). In the more detailed 7 –way split, 46% (n=47.5) of the isolates were identified as non-avian wildlife followed by 15% (n=16) from non-avian wildlife, 10% (n=10.5) from cattle, 6% (n=6) from humans, 5% (n=5) from pets, and finally 3% each from non-avian and avian livestock, respectively. For one of the tested isolates, the best match in the Texas *E. coli* BST Library was to two different isolates – one from a livestock and domesticated animal source and one from a wildlife source. Since this isolate matched each category equally, the isolate was classified as originating from either a wildlife and domesticated animal source (cattle for the 7-way split) or a wildlife source (non-avian wildlife for the 7-way split) and thus split (0.5:0.5) between these two categories.

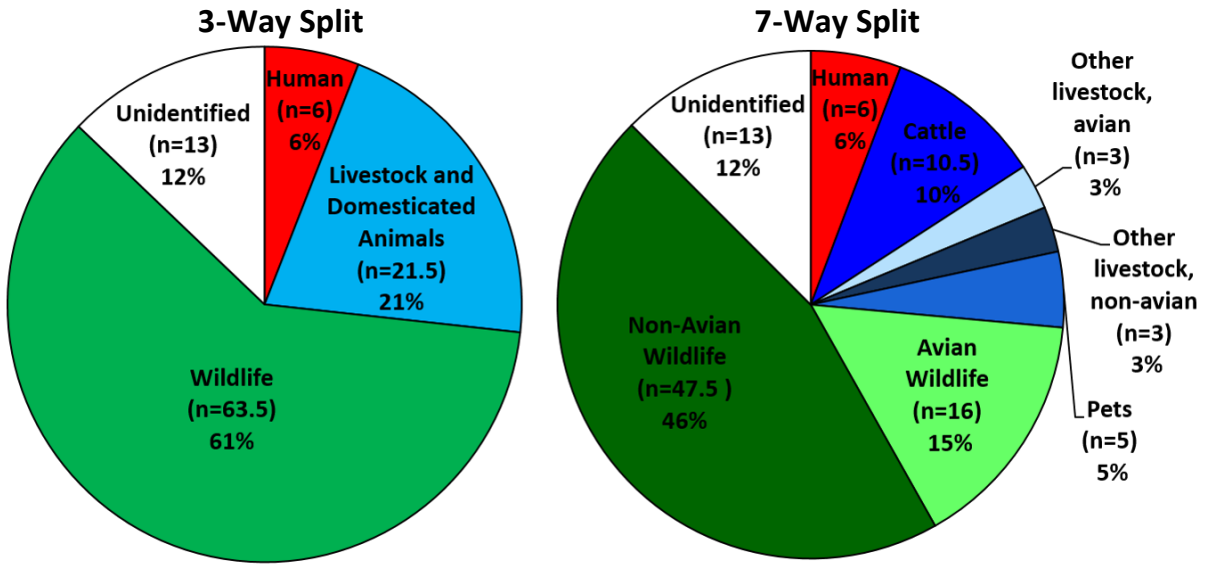

Figure 3. Identification of *E. coli* isolates (n=104) from stream samples (base and storm flow) using a 3-way split for source classification (L) and a 7-way split for source classification (R).

Comparison of *E. coli* Identifications from Samples Collected under Base-Flow and Storm-Flow

Results from the combined storm flow sites at Attoyac Bayou at SH 21 and Big Iron Ore Creek at FM 354 are shown in Figure 5 using a three-way split. Each site had 9 base flow isolates and 9 storm flow isolates for a total of 18 isolates from base flow and storm flow samples.

For the base flow samples using a 3-way split, 56% (n=10) of the isolates classified as originating from wildlife sources followed by 33% (n=6) from livestock and domesticated animals, and human 6% (n=1). The source could not be identified for 6% (n=1) of the isolates from the base flow samples. For the storm flow samples using a 3-way split, 56% (n=10) of the isolates classified as originating from wildlife sources followed by 33% (n=6) from livestock and domesticated animals. The source could not be identified for 11% (n=2) of the isolates.

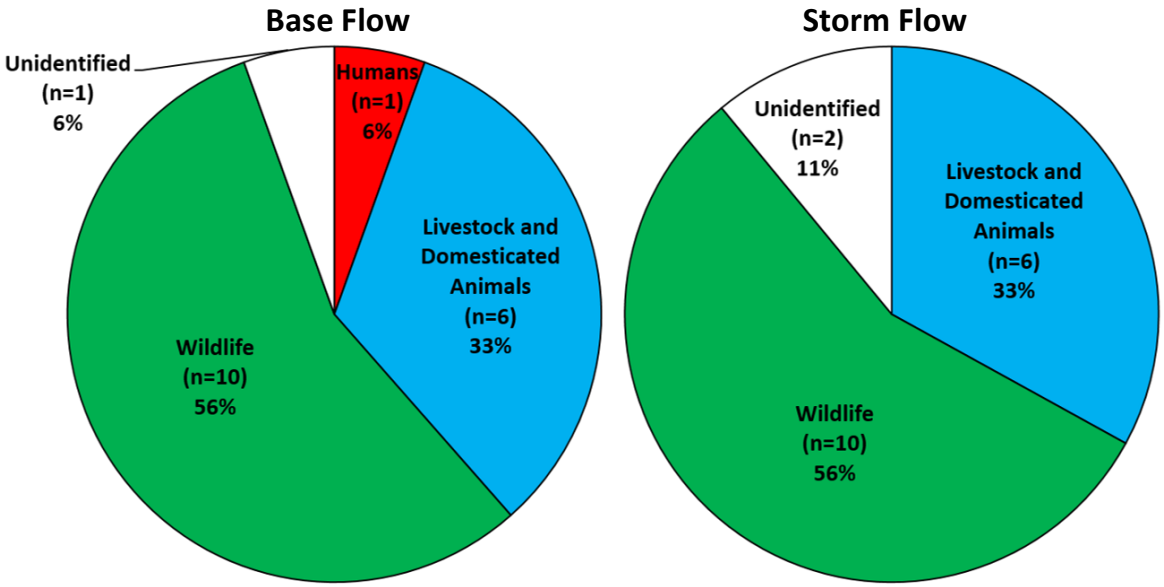

Figure 5. Comparison of identification of *E. coli* isolates (n=18) from storm sampling sites, Attoyac Bayou at SH 21 and Big Iron Ore Creek at FM 354, using a 3-way split for source classification for base flow (L) and storm flow (R).

Wastewater Treatment Facilities

WWTF in the Attoyac Bayou were included quarterly in monitoring activities and a limited number of samples (n=20) from the four facilities were analyzed for the *Bacteroidales* PCR marker occurrence. The general marker was detected in only 65% (n=12) of the samples, and the human marker and hog markers were detected in 55% (n=11) of the samples, followed by 10% (n=2) of the ruminant marker. The horse marker was not detected in any of the samples. Total volume outflow at these locations is relatively low (3-year average ranged from 0.0048 to 0.3 MGD) and generally the *E. coli* levels are well below state standards.

Summary and Discussion

A combination of library-dependent and -independent BST was utilized to characterize sources of fecal contamination in the Attoyac Bayou. Both *E. coli* and *Bacteroidales* results indicated a combination of wildlife and domesticated animals are the major source contributors of bacterial contamination in the watershed. The majority of the *E. coli* isolates (61%) were classified as originating from wildlife, including feral hogs, raccoons, coyotes, and other small mammals. Cattle, other livestock including horses, sheep and goats, as well as pets made up another 21% of the total isolates identified. *Bacteroidales* analysis indicated ruminant and hog markers were the most frequently detected with an increase in incidence during storm events. The *E.coli* identifications when evaluated from base flow versus storm events were very similar with the majority of the isolates identified from wildlife (56%) and livestock and domesticated animals (33%). *Bacteroidales* analysis showed increased frequency of human source markers just downstream of one of the WWTFs, but not at other locations. This was corroborated by the low number of human-classified *E. coli* isolates from across the

watershed. However, any human contributions are important particularly under base-flow conditions as the likelihood that fecal contamination from human sources contains pathogens is 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 47% of the total samples and the ERIC-RP characterized 10% 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, and the *Bacteroidales* PCR approach used in this study 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 47% 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 10% 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.

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.

USEPA (2006). 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.