Bacterial Source Tracking to Support Adaptive Management of the Arroyo Colorado Watershed Protection Plan: Final Report

Texas Water Resources Institute TR-494
January 2016

Texas Water Resources Institute
make every drop count

Allen Berthold, Texas Water Resources Institute
Jude A. Benavides, University of Texas Rio Grande Valley
Elizabeth Casarez, George Di Giovanni,
Joy Archuleta-Truesdale, University of Texas School of Public Health

Bacterial Source Tracking to Support Adaptive Management of the Arroyo Colorado Watershed Protection Plan

Final Report

By
Allen Berthold, Texas Water Resources Institute

Jude A. Benavides, University of Texas Rio Grande Valley

Elizabeth Casarez, George Di Giovanni, Joy Archuleta-Truesdale,
University of Texas School of Public Health

TSSWCB Project 12-10
Texas Water Resources Institute Technical Report TR-494
College Station, Texas

January 2016

Partners:

Texas Water Resources Institute
Brownsville Public Utilities Board
University of Texas Rio Grande Valley
University of Texas, School of Public Health El Paso

The Texas Water Resources Institute is part of the Texas A&M AgriLife Extension Service, Texas A&M AgriLife Research, and the College of Agriculture and Life Sciences at Texas A&M University.

Funding provided 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.

Contents

List of Acronyms

BMPs – Best Management Practices
ERIC-PRC - Enterobacterial Repetitive Intergenic Consensus – Polymerase Chain Reaction
IRNR – Institute of Renewable Natural Resources
TCEQ – Texas Commission on Environmental Quality
TSSWCB – Texas State Soil & Water Conservation Board
UTRGV – University of Texas Rio Grande Valley
UTSPH-EP – University of Texas School of Public Health – El Paso
WPP – Watershed Protection Plan

Executive Summary

The Arroyo Colorado River, located in the Lower Rio Grande Valley of Texas, is a major tributary to the Lower Laguna Madre. Currently, it is impaired for a variety of pollutants, bacteria being one of those. Watershed planning efforts are ongoing and in order to identify what the primary causes and sources of bacteria are, watershed managers developed a monitoring strategy to conduct bacterial source tracking. Through this project, the Texas A&M AgriLife, Institute of Renewable Natural Resources visited with local agency personnel and decided that birds and wildlife were the key sources of interest. As a result, 254 fecal samples were collected from 27 known sources. Samples were shipped to the University of Texas School of Public Health – El Paso, isolated, and archived (409 isolates). Additionally, the University of Texas, Rio Grande Valley collected monthly water samples for one year at 10 different sites. Periodically samples were unable to be collected due to dangerous conditions but 113 samples were collected. Collected samples were taken to the Brownsville Public Utilities Board Analytical Laboratory for enumeration of *E. coli* and *Enterococcus*. Samples were also processed and shipped to University of Texas School of Public Health – El Paso for source tracking analysis. Results of source tracking indicated that 52% of the bacteria resulted from non-avian wildlife, 16% from avian wildlife, 10% from cattle, 9% from human, 10% unidentified, and 1% from pets, avian livestock, and other non-avian livestock. Results of this analysis will be used to guide the development of the Arroyo Colorado Watershed Protection Plan.

Introduction

The Arroyo Colorado is located in the Lower Rio Grande Valley of South Texas and flows through the middle of Hidalgo and Cameron counties. The Lower 16 miles of the Arroyo Colorado make up the boundary between Cameron and Willacy counties. The Arroyo Colorado drainage area is a subwatersehd of the Nueces-Rio Grande Costal Basin, also known as the Lower Laguna Madre watershed. The streams of the Nueces- Rio Grande Coastal Basin, including the Arroyo Colorado, drain to the Laguna Madre, which is considered one of the most productive hypersaline lagoon systems in the world. The Lower Rio Grande Valley comprises the northern part of the Rio Grande Delta, a broad fluviodeltaic plain laid down over tens of thousands of years by the ancestral Rio Grande. Just as the Rio Grande is the major source of freshwater for the Lower Rio Grande Valley, the Arroyo Colorado serves as the main drainage stream for this area of Texas.

The Arroyo Colorado currently has low dissolved oxygen levels within the tidal segment, not meeting the aquatic life use designated by the State of Texas and described in the Water Quality Standards. This has been the case for every 303(d) list prepared by the State since 1996. In addition, bacteria has always been a parameter of concern and as of 2006, the Arroyo Colorado became impaired due to high levels of bacteria. There are many challenges associated with restoring water quality in the Arroyo Colorado watershed. The watershed is one of the most productive agricultural areas in the State; however, it also has one of the fastest growing populations of any region in the State. This causes significant land fragmentation which increases the threat for bacterial impairments. The bacteria impairment in the Arroyo Colorado not only poses a human health threat through contact recreation, but also potentially through consuming food that is grown with this water.

The Texas Commission on Environmental Quality (TCEQ) is funding a project to revise the Arroyo Colorado WPP. Initially written with an implementation schedule through 2015, the Arroyo Colorado WPP focused primarily on nutrients but through adaptive management, the bacteria impairment has become an issue stakeholders are prepared to address in the next phase of the plan.

Previous work conducted in the watershed has laid the ground work for updating the watershed protection plan and this project provided additional information that will be incorporated into that effort. Additionally, other projects are ongoing with a range of focus from sustaining the partnership to educating landowners on available incentive programs to implement agricultural Best Management Practices (BMPs).

Although data collection through past projects tend to further justify the currently listed impairment, this data remains limited and additional data was needed to accurately calculate bacteria loading rates and identify the most likely sources of bacterial contamination. As planning continued, the need to bolster datasets and comprehensive data analysis arose as management options were and continue to be considered. Without adequate data, uncertainty increases in properly identifying the source of contamination in the watershed and this project attempts to address some of that uncertainty.

Objectives

Key objectives for this project included: 1) Evaluating bacterial concentrations and sources in the Arroyo Colorado watershed; 2) Conducting a detailed assessment of bacterial levels throughout the Arroyo Colorado watershed; and 3)Identification of major bacterial sources.

Methods

Through this project, a water quality monitoring regime was employed in an attempt to help make appropriate recommendations for addressing the bacteria impairment in the revision of the Arroyo Colorado WPP. Monthly sampling was conducted by University of Texas Rio Grande Valley (UTRGV) at 10 sampling stations identified in Table 1 below.

Table 1. Monitoring Stations and Descriptions

Station ID	Description
13086	Arroyo Colorado at FM 336 South of McAllen
13084	Arroyo Colorado at US 281 South of Pharr
13082	Arroyo Colorado at FM 493 South of Donna
13080	Arroyo Colorado at FM 506 South of La Feria
13079	Arroyo Colorado at U.S. 77 in Southwest Harlingen
13074	Arroyo Colorado at Low Water Bridge at Port Harlingen
13072	Arroyo Colorado Tidal FM 106 Bridge at Rio Hondo
13073	Arroyo Colorado Tidal at Camp Perry North of Rio Hondo
13559	Arroyo Colorado Tidal at Marker 27 (Mile 15) 0.5 Mile North of the Point Where Channel Becomes Boundary Between Willacy and Cameron Counties
13782	Arroyo Colorado Tidal Near CM 16 at Arroyo City, KM 10.9

Figure 1. Locations of Monitoring Stations in the Arroyo Colorado Watershed

Field parameters were measured (pH, temperature, conductivity, and dissolved oxygen), and water samples were delivered to the Brownsville Public Utilities Board Analytical Laboratory for enumeration of *E. coli* and *Enterococcus*. Also, samples were processed using the EPA 1603 method and shipped to University of Texas School of Public Health-El Paso (UTSPH-EP) for bacterial source tracking (BST) analysis. The UTRGV maintained a database containing field and laboratory analysis and TWRI has prepared data for submission to TSSWCB and subsequent transmittal to TCEQ for inclusion in the Surface Water Quality Monitoring Information System.

The Institute of Renewable Natural Resources (IRNR) and TWRI consulted with local experts on possible primary sources of bacteria and attempted to collect known source samples based on this information. Samples were packaged and mailed to the UTSPH-EP laboratory for processing and inclusion in the Texas *E. coli* BST Library.

The UTSPH-EP assessed and identified different sources contributing to bacterial loadings, using the library-dependent BST methods and analyzing *E. coli* isolates using the ERIC-PCR and RiboPrinting combination method.

Results

Known source sampling in the Arroyo Colorado watershed resulted in a total of 254 samples being collected between October 2013 and October 2014. An emphasis was placed on seabirds and small mammals because this was expected to be a contributing source, and the Texas *E. coli* BST library lacked in these categories. Samples collected were shipped to UTSPH-EP for processing and Table 2 describes the number of samples collected per source, the number of

samples testing positive for *E. coli*, screened, validated, archived and added to the Texas *E. coli* BST library.

Table 2. Number of Samples Collected per Source, the Number of Samples Testing Positive for *E. Coli*, Screened, Validated, Archived and Added to the Texas *E. Coli* BST Library

Source	Samples collected	Samples (+) for E. coli	Isolates archived	Isolates screened by ERIC	Isolates RP in local library	Self-validated (isolate/sample)	TXSV 5-15 (isolate/sample)
Human	**21**	**21**	**104**	**63**	**49**	**27/18**	**23/16**
Sewage	6	6	30	18	16	9/5	6/4
Septage	15	15	74	45	33	18/13	17/12
Cattle	**6**	**5**	**25**	**15**	**5**	**2/2**	**1/1**
Other non-avian livestock	**17**	**15**	**72**	**45**	**19**	**4/4**	**1/1**
Goat	3	3	15	9	4	1/1	0/0
Pig	5	5	24	15	7	2/2	0/0
Rabbit	5	3	13	9	4	1/1	1/1
sheep	4	4	20	12	4	0/0	0/0
Other avian livestock--chicken	**4**	**4**	**14**	**10**	**7**	**2/2**	**2/2**
Pets--dog	**1**	**0**	**0**	**0**	**0**	**0/0**	**0/0**

continued next page

Source	Samples collected	Samples (+) for E. coli	Isolates archived	Isolates screened by ERIC	Isolates RP in local library	Self-validated (isolate/sample)	TXSV 5-15 (isolate/sample)
Avian Wildlife	**155**	**31**	**111**	**78**	**37**	**22/20**	**15/14**
Black tern	61	3	13	9	4	3/3	2/2
Ruddy turnstone	1	1	5	3	1	0/0	0/0
Bird	33	12	21	21	15	11/10	9/8
Western snowy plover	6	0	0	0	0	0/0	0/0
Herring gull	2	0	0	0	0	0/0	0/0
Laughing gull	7	2	10	6	3	2/1	1/1
Owl	1	1	5	3	1	0/0	0/0
Pelican	17	7	35	21	8	4/4	2/2
Pigeon	1	0	0	0	0	0/0	0/0
Sparrow	1	0	0	0	0	0/0	0/0
Royal tern	22	5	22	15	5	2/2	1/1
Wren	2	0	0	0	0	0/0	0/0
Unmarked species	1	0	0	0	0	0/0	0/0
Non-avian Wildlife	**50**	**23**	**83**	**63**	**27**	**18/15**	**17/14**
White-footed mouse	39	18	64	49	21	15/12	14/11
Norton pygmy mouse	2	0	0	0	0	0/0	0/0
White-tailed deer	2	2	8	6	2	1/1	1/1
Hispid cotton rat	2	1	4	3	1	1/1	1/1
Wild rabbit	1	1	5	3	1	0/0	0/0
Unknown mammal	4	1	2	2	2	1/1	1/1
Total	254	99	409	274	144	75/61	59/48

Water quality monitoring was conducted on a monthly basis at 10 sites along the Arroyo Colorado. Periodically, data was unable to be collected at some sites due to conditions of the waterbody. Parameters collected were salinity, dissolved oxygen, temperature, pH, specific conductance, Enterococcus, and *E. coli*. Tables 3 and 4 contain measured levels from both Enterococcus and *E.coli*, respectively, and Appendix A contains results from the remaining parameters.

Table 3. Measured Enterococcus Levels and Geometric Means by Station and by Date

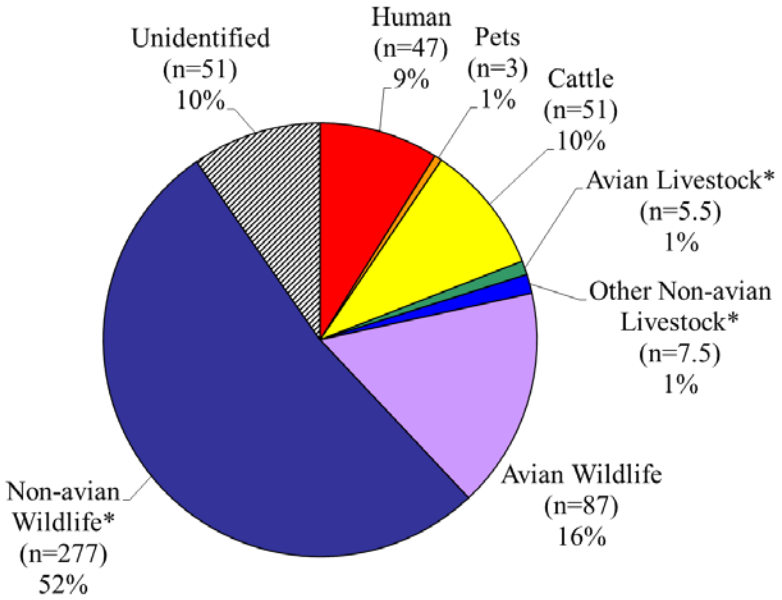
Station ID#	13086	13084	13082	13080	13079	13074	13072	13073	13559	13782	Geometric Mean
6/16-17/2014	548	50.4	2420	2420	816	579	31.1	64.5	6.3	56.0	187.6
7/7-8/2014	344	151	830	437	866	525	102	328	33.6	2420	344.1
8/11-12/2014	2420	2420	2420	2420	2420	2420	248.1	98.4	2420	2420	1399.0
9/22-23/2014	2420	2420	------	2420	2420	2420	201.2	2420	------	196.8	1295.9
10/20-21/2014	2420	2420	816.4	866.4	579.4	2420	63.1	129.6	------	2420	786.3
11/17-18/2014	1553.1	1119.9	1119.9	613.1	2419.6	980.4	547.5	648.8	------	161.6	817.2
12/15-16/2014	------	2420	1732.9	1553.1	547.5	866.4	365.4	435.2	42.0	114.5	510.7
1/26-27/2015	648.8	365.4	184.2	183.5	162.4	190.4	160.7	110.0	62.2	59.4	166.3
2/9-10/2015	770.1	478.6	387.3	686.7	360.9	344.8	154.1	112.4	43.1	41.4	227.8
3/2-3/2015	1046.2	1119.9	1553.1	866.4	1203.3	290.9	178.5	172.5	------	547.5	594.7
4/6-7/2015	2420	2420	2420	2420	2420	2420	1119.9	387.3	------	387.3	1478.4
5/4-5/2015	2420	980.4	2420	2420	1732.9	2420	152.9	161.6	64.5	2420	861.3
Geometric Mean	1259.9	846.7	1157.5	1103.6	994.8	928.2	184.0	232.2	62.4	352.2	

Table 4. Measured E. coli Levels and Geometric Means by Station and by Date

Station ID#	13086	13084	13082	13080	13079	13074	13072	13073	13559	13782	Geometric Mean
6/16-17/2014	308	152	308	435	155	186	17.5	19.3	44.1	192	117.9
7/7-8/2014	770	276	866	345	326	166	17.6	19.5	137	816	205.1
8/11-12/2014	648.8	547.5	727.0	547.5	387	1300	116	51.2	70.8	2420	385.4
9/22-23/2014	1413.6	1203.3	------	727.0	328.2	2419.6	101.7	344.8	------	23.8	411.3
10/20-21/2014	1203.3	816.4	143.9	111.2	90.8	613.1	28.1	52	------	344.8	196.7
11/17-18/2014	325.5	185.0	410.6	248.9	209.8	193.5	261.3	155.2	------	410.6	252.4
12/15-16/2014	------	365.4	488.4	461.1	461.1	344.8	198.9	298.7	53.8	130.1	260.1
1/26-27/2015	517.2	135.4	64.5	125.0	79.4	108.1	78.5	62.0	51.2	56.3	96.2
2/9-10/2015	416.0	365.4	185.0	148.3	113.0	172.3	145.0	108.1	116.9	238.2	179.9
3/2-3/2015	727.0	365.4	275.5	224.7	125.9	101.7	48.0	56.3	------	517.2	188.1
4/6-7/2015	866.4	193.5	260.3	248.1	307.6	866.4	613.1	461.1	------	261.3	390.9
5/4-5/2015	290.9	224.7	275.5	248.1	142.1	155.3	16.0	20.1	6.3	95.9	84.8
Geometric Mean	595.8	321.0	292.1	276.0	193.7	316.6	74.3	78.6	50.8	235.4	

From the 12 sampling events, 113 water samples were sent to UTSPH-EP. From these water samples, 774 isolates (up to eight per sample) were archived and a total of 529 isolates were analyzed with Enterobacterial Repetitive Intergenic Consensus – Polymerase Chain Reaction (ERIC-PCR) and RiboPrint composite (ERIC-RP) fingerprinting. 90% of the water isolates were identified using the Texas *E. coli* BST Library (ver. 5-15) and results of the analysis indicated that wildlife was the leading contributor of bacteria in the Arroyo Colorado, followed by cattle and humans. Figure 2 displays the BST results and more discussion regarding BST can be found in Appendix B.

Figure 2. BST Results Identified by Source



Discussion and Continued Work

Work under this project was aimed at providing the Arroyo Colorado Watershed Partnership with detailed information about sources of bacteria. This information is to be used for identifying sources of bacteria as well as choosing the best management practices that would be most effective. While wildlife was found to be the largest contributor of bacteria to the Arroyo Colorado, wildlife has also been found to be the primary bacteria contributor in many other watersheds as well. Since there is a mixed land use of agriculture and urban within the Arroyo Colorado watershed, both agricultural and urban best management practices should be considered for adoption, especially those located closest to the riparian areas. Also, wildlife found within city boundaries will likely congregate in or near riparian areas due to the lack of other habitat leading to higher probability of contributing bacteria to the watershed. When planning to implement wildlife BMPs, it is important to also consider mid-sized and small mammals in addition to larger, more commonly thought of mammals.

Currently, no further BST work is planned to be conducted in the Arroyo Colorado as results will be incorporated into the updated Arroyo Colorado Watershed Protection Plan.

Appendix A – Arroyo Colorado Monitoring Data and Analysis

Bacterial Source Tracking to Support Adaptive Management of the Arroyo Colorado Watershed Protection Plan

UTB (UTRGV): Final Report on Routine Monitoring and Monitoring Data Analysis

Jude A. Benavides, Ph.D.

February, 2016

Summary of Field Data Collection Efforts for BST on the Arroyo Colorado

In support of a Bacterial Source Tracking (BST) effort on the Arroyo Colorado (Arroyo), the University of Texas Rio Grande Valley (UTRGV) research team conducted field sampling at ten locations along the Arroyo over a period of twelve months (June 2014 – May 2015). The primary focus of the sampling effort was to collect and transport grab samples to the Brownsville Public Utilities Board (BPUB) Analytical Laboratory for E. coli and Enterococcus isolation and enumeration, with further source tracking analysis completed by the UT School of Public Health El Paso (UTSPH EP). Grab samples were taken at each station with the exception of rare cases where field conditions prohibited access. Additionally, streamflow measurements were conducted at two of the non-tidally influenced sampling stations. Streamflow data was gathered at two other locations from stream gages operated by the International Boundary and Water Commission (IBWC). Also, vertical profiles of pH, dissolved oxygen, specific conductance, temperature and total depth of the water column (midstream) were conducted at each sampling location where grab samples for bacteria were collected. Additional data collected included noting the days since last significant rainfall to assist with determining possible correlations between bacterial counts and flow in the stream.

Table 1: Number of grab samples, vertical profiles, and streamflow measurements taken over the year (June 2014 – May 2015) at each of the approved 10 sampling stations along the Arroyo Colorado.

	13086 McAllen	13084 Pharr	13082 Donna	13080 La Feria	13079 Harlingen	13074 Cemetery Rd	13072 Rio Hondo	13073 Camp Perry	13559 County Line	13782 Arroyo City
Grab Samples	11	12	11	12	12	12	12	12	7	12
Vertical Profiles	10	12	11	12	12	12	12	12	7	12
Stream Flow	10	0	IBWC Gage 08-4703.01	IBWC Gage 08-4703.01	IBWC Gage 08-4704.00	10	N/A - Tidal	N/A - Tidal	N/A - Tidal	N/A - Tidal

Table 1 illustrates the number of samples/measurements taken at each station over the year-long sampling period. Grab samples were collected unless field conditions prohibited access. Station 13559 was only accessible seven months out of the 12. Stations 13086 (McAllen) and 13082 (Donna) were inaccessible one month out of the year, and the McAllen station was not accessible for vertical profile and streamflow for two out of the 12 months. Streamflow measurements were taken at the most downstream, non-tidal station (13074 – Cemetery Rd) and the most upstream station (13086 – McAllen) for all but two of the twelve months. Gages downstream of 13074 are in the tidally influenced segment and were not sampled for streamflow. Streamflow measurements at stations between McAllen and Cemetery Rd proved to be extremely difficult to measure for streamflow due to accessibility constraints caused by excessive vegetation and/or debris in the channel and/or banks. Fortunately, two IBWC gages (Harlingen and Mercedes)

were available and data provided by them was used in place of field measurements. Note, the USGS gage at Rio Hondo was not operational until after May 2015.

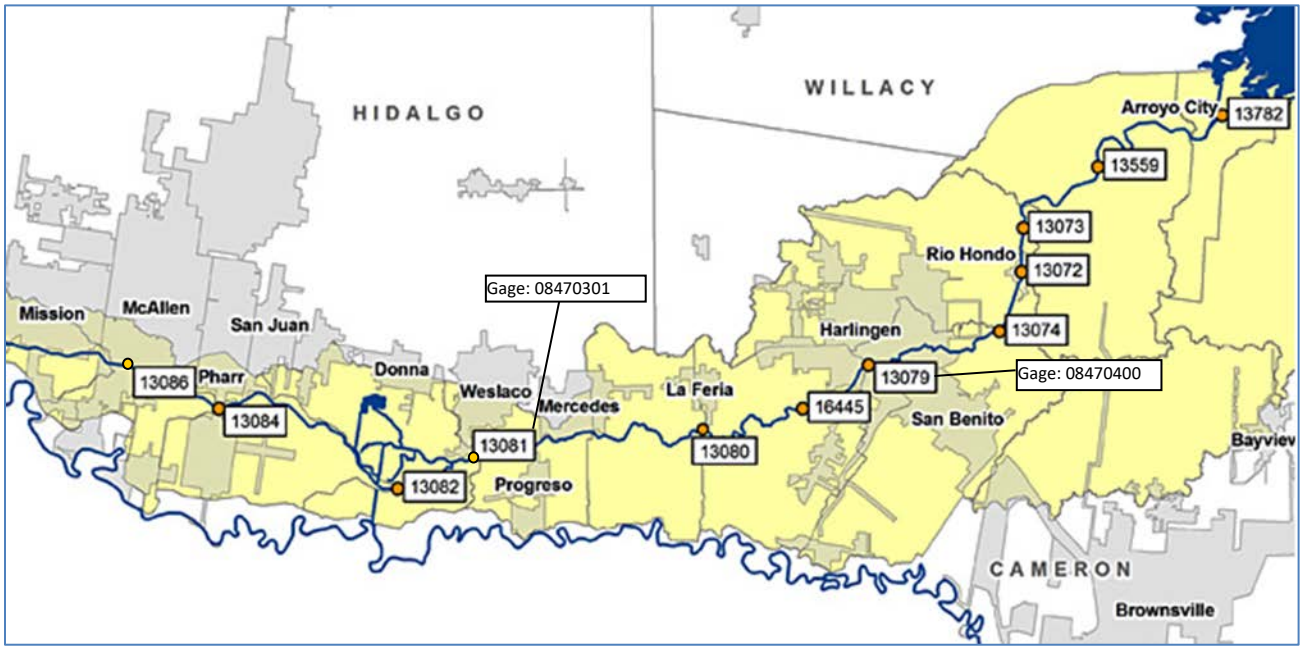


Figure 1: Arroyo Colorado Watershed Map, Monitoring Stations Sampled, and IBWC Streamflow gage locations at Mercedes and Harlingen

This report will summarize the methodologies used in collecting grab samples and field measurements, as well as provide a concise summary of the resulting data. Specifically, the report will show and briefly discuss the bacterial enumerations (counts) for E. coli and Enterococcus; a summary of vertical profile data; and streamflow measurements for each month/station sampled or measured.

METHODOLOGY

Sample Collection, Storage and Delivery

Water quality grab samples for bacteria analysis were collected on monthly intervals per the approved Quality Assurance Project Plan (QAPP). The number of successfully collected samples per station is shown in Table 1. Field conditions frequently prohibited access at Station 13559, which was only accessible seven months out of the 12. Stations 13086 (McAllen) and 13082 (Donna) were inaccessible one month out of the 12. In all cases when the sampling site was accessible, water samples were collected directly from the stream (midway in the stream channel) via kayak, with appropriate care to avoid surface microlayer of water and bottom sediment to ensure the sample was representative of water in the stream. Grab samples for analysis of *E. coli* and Enterococcus were collected in 500 ml sterile bottles provided by Brownsville Public Utilities Board (BPUB). Samples were collected following procedures in the most recent version of the *TCEQ SWQM Procedures, Volume 1 (RG-415),* including proper storage in ice chests and delivery to the BPUB analytical laboratory within the required holding

time. Normally, samples were delivered well under four hours to permit sufficient time for laboratory analysis within the required six hour holding time window. In order to meet this time frame of delivery, a field sampling plan was developed that split the stations into two separate sampling days. Day one included sample collection and all field data measurements (flow, vertical profiles, field conditions, etc.) for the stations located at and downstream of Harlingen. These included stations: 13079, 13074, 13072, 13073, 13559 and 13782. Day two included sample collection and all field data measurements for the stations upstream of Harlingen commencing with the upstream locations. These included stations: 13086, 13084, 13082 and 13080.

Vertical Profiles of Salinity, DO, pH, Water Temperature and Total Depth of Water

Vertical profiles of salinity, DO, pH, water temperature along with total depth of water were conducted each time a grab sample was collected, with the exception of one sample at Station 13086 (McAllen) – see Table 1. These parameters were measured in situ with an EXO1 (Xylem / YSI) multiparameter 4-port water quality sonde with depth sensor. Data were recorded in field notes and transferred immediately to electronic format after returning to University of Texas – Brownsville (UTB).

Streamflow Measurements/Supplemental Streamflow Data from IBWC Gages

Streamflow measurements for all non-tidally influenced stations were not possible due to access restrictions resulting from excessive vegetation along the banks and/or excessive debris in the channel and/or banks. Navigation via kayak or boat to these stations was also not possible due to debris, navigation hazards and inaccessibility for miles both upstream and downstream of the sampling station. The extremely wet, deep clay along the banks of the Arroyo Colorado (often more than 5 feet deep) proved to be a detriment as well as a potential safety hazard. The two stations that permitted streamflow measurements included Station 13086 (McAllen) and Station 13074 (Cemetery Rd). These two stations represented the most upstream station (McAllen) as well as the most downstream, non-tidally influenced station (Cemetery Rd.). Streamflow measurements were taken in accordance with U.S. Geological Survey's method as described in *TCEQ SWQM Procedures, Volume 1 (RG-415).* Alternative methods for streamflow measurement including acoustic Doppler sensor deployment from kayak, boat or bridges was considered impractical due to the fact that all bridges near the sampling stations between McAllen and Cemetery road coincided with areas of extreme debris including rip-rap, submerged concrete blocks with protruding rebar, or had bridge piers that altered the flow regime significantly.

Data from two IBWC operated streamflow gages were obtained in order to compensate for the lack of streamflow data between McAllen and Cemetery Road. Fortunately, the location of these gages (See Figure 1) were ideally located in order to provide a streamflow dataset at locations nearly equally distributed across the non-tidal segment of the Arroyo. The IBWC gages used included: IBWC Gage #08470301 located south of Mercedes and IBWC Gage #08470400 located in Harlingen.

Data and Discussion – *E. Coli* and Enterococci Results

Results for *E. coli* and Enterococci analysis conducted by the BPUB are shown in Table 2a and 2b, respectively. The results include the 12 sampled months plus a dry run sample collected on June 11, 2014. This dry run data was not included in any analysis, but served as an excellent opportunity for the sample collection and analysis teams to make sure holding times could be met and to insure there were no logistical issues with the planned sampling. Tables 2a and 2b show the results by collection date and parameter and station ID. Dashed lines illustrate no sample was collected for that date and station. Units are in Most Probable Number per 100 mL (MPN/100mL).

Table 2a: Results for E. coli sampling shown in MPN/100mL.

Staition ID#	Blank	13086	13084	13082	13080	13079	13074	13072	13073	13559	13782	UNITS
Parameter	Results	Results	Results	Results	Results	Results	Results	Results	Results	Results	Results	
E. Coli, (6/11/2014)	< 1.0	145	------	------	365	------	435	137	------	------	49	MPN/100mL
E. Coli, (6/16-17/2014)	< 1.0	308	152	308	435	155	186	17.5	19.3	44.1	192.0	MPN/100mL
E. Coli, (7/7-8/2014)	< 1.0	770	276	866	345	326	166	17.6	19.5	137	816	MPN/100mL
E. Coli, (8/11-12/2014)	< 1.0	648.8	547.5	727	547.5	387	1300	116	51.2	70.8	>2420	MPN/100mL
E. Coli, (9/22-23/2014)	< 1.0	1413.6	1203.3	------	727	328.2	2419.6	101.7	344.8	------	23.8	MPN/100mL
E. Coli, (10/20-21/2014)	< 1.0	1203.3	816.4	143.9	111.2	90.8	613.1	28.1	52	------	344.8	MPN/100mL
E. Coli, (11/17-18/2014)	< 1.0	325.5	185	410.6	248.9	209.8	193.5	261.3	155.2	------	410.6	MPN/100mL
E. Coli, (12/15-16/2014)	< 1.0	------	365.4	488.4	461.1	461.1	344.8	198.9	298.7	53.8	130.1	MPN/100mL

Table 2b: Results for Enterococci sampling shown in MPN/100mL.

E. Coli, (4/6-7/2015)	< 1.0	866.4	193.5	260.3	248.1	307.6	866.4	613.1	461.1	------	261.3	MPN/100mL
E. Coli, (5/4-5/2015)	< 1.0	290.9	224.7	275.5	248.1	142.1	155.3	16	20.1	6.3	95.9	MPN/100mL
Station	Blank	13086	13084	13082	13080	13079	13074	13072	13073	13559	13782	UNITS
Parameter	Results	Results	Results	Results	Results	Results	Results	Results	Results	Results	Results	
Enterococci, (6/11/2014)	< 1.0	80.9	------	------	> 2420	------	722	130	------	------	328	MPN/100mL
Enterococci, (6/16-17/2014)	< 1.0	548	50.4	>2420	>2420	816	579	31.1	64.5	6.3	56.0	MPN/100mL
Enterococci, (7/7-8/2014)	< 1.0	344	151	830	437	866	525	102	328	33.6	>2420	MPN/100mL
Enterococci, (8/11-12/2014)	< 1.0	2420	>2420	>2420	>2420	>2419.6	>2419.6	248.1	98.4	>2419.6	>2419.6	MPN/100mL
Enterococci, (9/22-23/2014)	< 1.0	>2419.6	>2419.6	------	>2419.6	>2419.6	>2419.6	201.2	>2419.6	------	196.8	MPN/100mL
Enterococci, (10/20-21/2014)	< 1.0	>2419.6	>2419.6	816.4	866.4	579.4	>2419.6	63.1	129.6	------	>2419.6	MPN/100mL
Enterococci, (11/17-18/2014)	< 1.0	1553.1	1119.9	1119.9	613.1	2419.6	980.4	547.5	648.8	------	161.6	MPN/100mL
Enterococci, (12/15-16/2014)	< 1.0	------	>2419.6	1732.9	1553.1	547.5	866.4	365.4	435.2	42.0	114.5	MPN/100mL
Enterococci, (1/26-27/2015)	< 1.0	648.8	365.4	184.2	183.5	162.4	190.4	160.7	110.0	62.2	59.4	MPN/100mL
Enterococci, (2/9-10/2015)	< 1.0	770.1	478.6	387.3	686.7	360.9	344.8	154.1	112.4	43.1	41.4	MPN/100mL
Enterococci, (3/2-3/2015)	< 1.0	1046.2	1119.9	1553.1	866.4	1203.3	290.9	178.5	172.5	------	547.5	MPN/100mL
Enterococci, (4/6-7/2015)	< 1.0	>2419.6	2419.6	>2419.6	>2419.6	>2419.6	>2419.6	1119.9	387.3	------	387.3	MPN/100mL
Enterococci, (5/4-5/2015)	< 1.0	>2419.6	980.4	>2419.6	>2419.6	1732.9	>2419.6	152.9	161.6	64.5	>2419.6	MPN/100mL

Figures 2 and 3 illustrate the geometric mean of *E. coli* data collected from the Arroyo Colorado. Figure 2 illustrates the geometric mean of all sampled station data by month. As an example, the Jun-14 (June 2014) column in Figure 2 shows that the geometric mean of *E. coli* for all stations sampled that month was 118 MPN/100 mL. Monthly averaged *E. coli* levels range from a low of 85 MPN/100 mL in May of 2015 to a high of 411 MPN/100 mL during the month of September 2014. Despite the wet months of August and September having the highest *E. coli* values, there was no readily discernible correlation between *E. coli* counts and the month of the year, temperature or rainfall. It remains that possible further study may yield correlations between monthly *E. coli* values and time of year, temperature, or rainfall. This is discussed in greater detail in following sections.

Figure 3 illustrates the geometric mean of all monthly *E. coli* data samples by station. For example, the Station 13086 column shows that the geometric mean of E. coli for all monthly samples collected at that station was 596 MPN/100 mL. Station averaged E. coli levels exhibited a marked trend from higher levels upstream to lower levels downstream with the exception of Stations 13074 and 13782. It should be noted that station 13074 was the downstream most non-

tidally influenced location sampled. A low dissolved oxygen salt wedge was occasionally observed below 2-3m at this station; however, it is uncertain as to how this might influence *E. coli* concentrations. Station 13782 was the most downstream station sampled and the higher values observed here may indicate some influence from the Lower Laguna Madre, although there is insufficient evidence to state this with certainty.

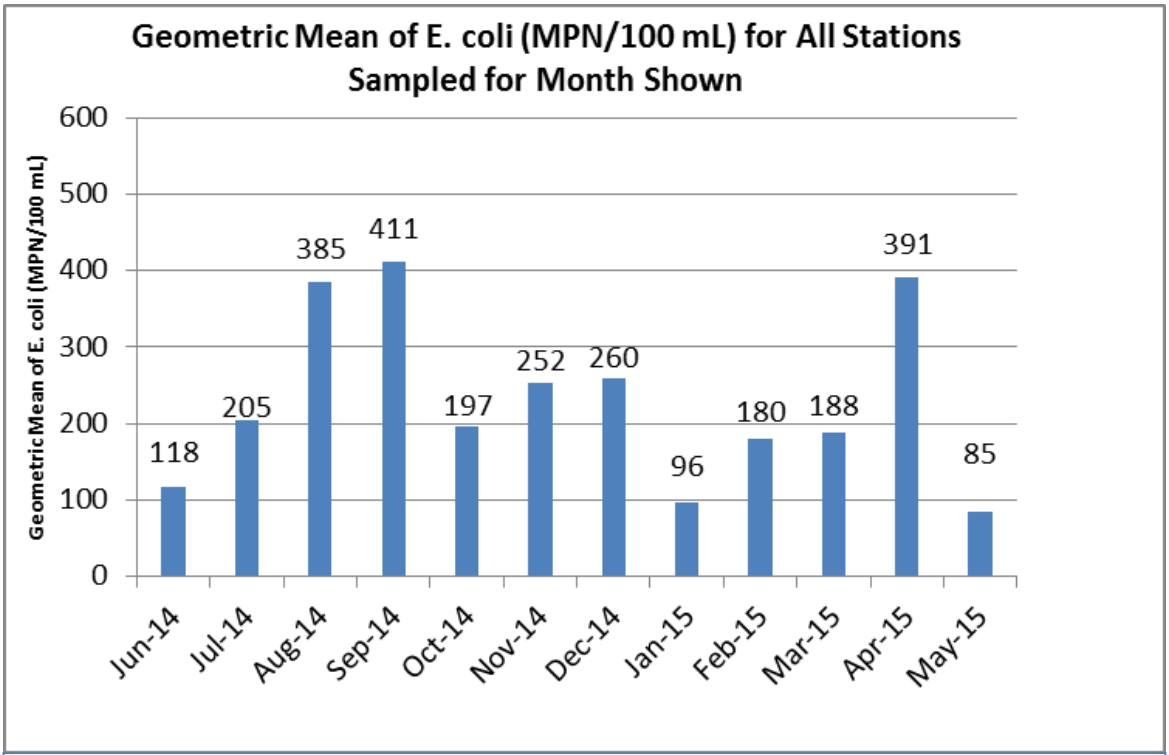


Figure 2: Geometric mean of *E. coli* for all sampled stations by month. (Example: Jun-14 shows that the geometric mean of E. coli for all stations sampled that month was 118 MPN/100 mL).

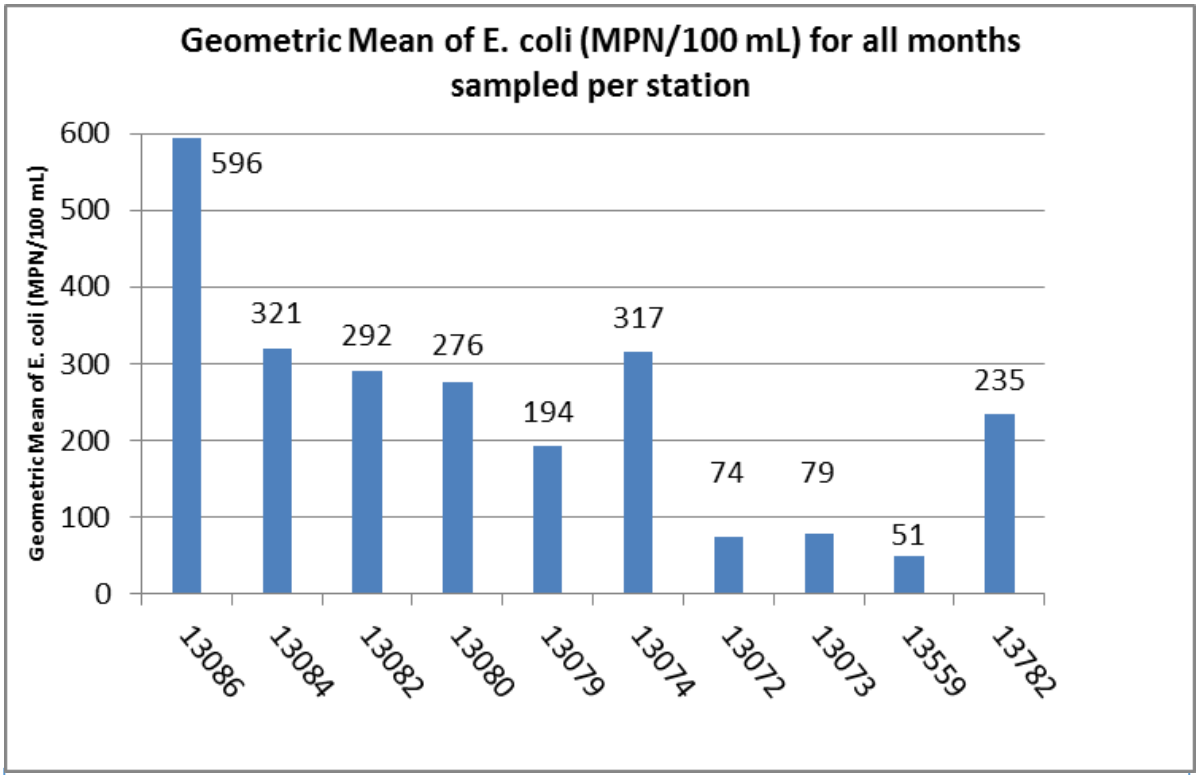


Figure 3: Geometric mean of *E. coli* for all monthly values per station. (Example: 13086 shows that the geometric mean of E. coli for all monthly samples for that station was 596 MPN/100 mL).

Figures 4 and 5 illustrate the geometric mean of Enterococci data collected from the Arroyo Colorado. Figure 4 illustrates the geometric mean of all sampled station data by month in the same manner as in Figure 2. Monthly averaged Enterococci levels ranged from a low of 166 MPN/100 mL in January of 2015 to a high of 1,478 MPN/100 mL during the month of April 2015. Again, the extremely wet months of August and September were among the highest levels, but no correlation between Enterococci levels and temperature, time of year or rainfall were obvious. Again, further study is recommended to determine if there are correlations between bacterial counts and any or all of precipitation, time of year or temperature.

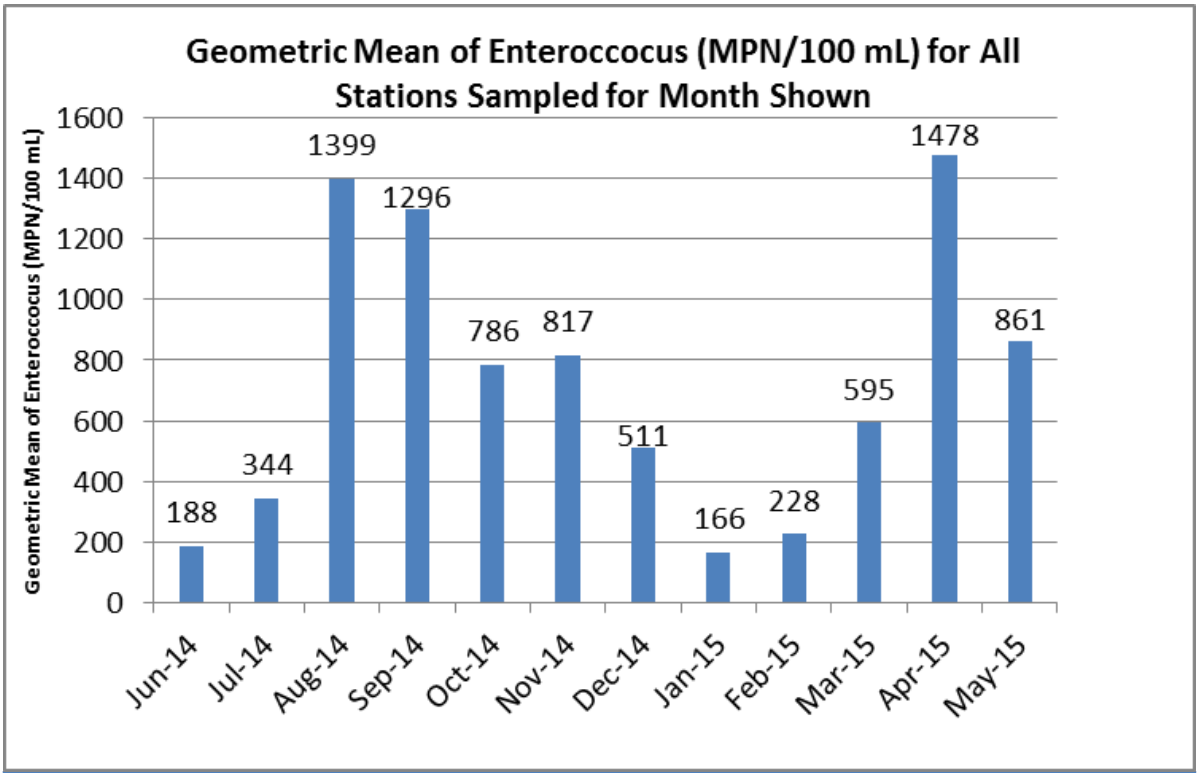


Figure 4: Geometric mean of Enterococci for all sampled stations during month shown. (Example: Jun-14 shows that the geometric mean of Enterococci for all stations sampled that month was 188 MPN/100 mL).

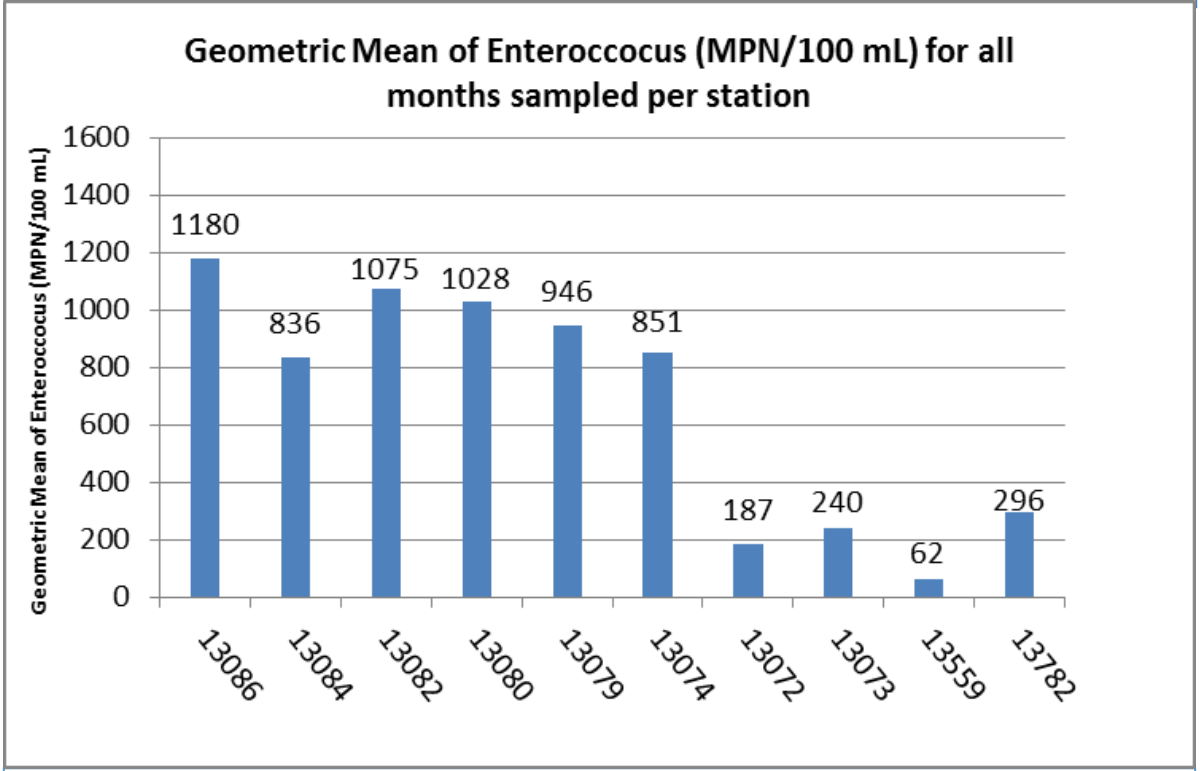


Figure 5: Geometric mean of Enterococci for all monthly values per station. (Example: 13086 shows that the geometric mean of Enteroccoci for all monthly samples for that station was 1180 MPN/100 mL).

Figure 5 illustrates the geometric mean of all monthly Enterococci data samples by station. As with *E. coli*, averaged Enterococci levels exhibited a marked trend from higher levels upstream to lower levels downstream with a marked reduction after Station 13074.

E. coli and Enterococcus data were evaluated for correlation as seen in Figures 6a and 6b, a moderate, positive correlation was present between geometric averaged *E. coli* and Enterococcus data for both monthly averaged and station averaged data.

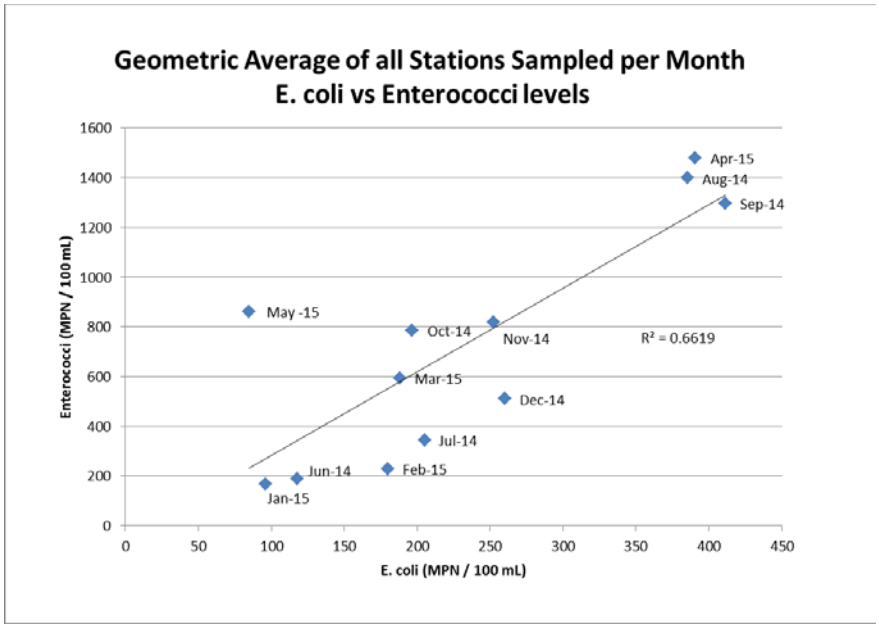


Figure 6a: Geometric Average of all Months Sampled per Station
***E. coli* vs Enterococci Levels (MPN / 100 mL)**

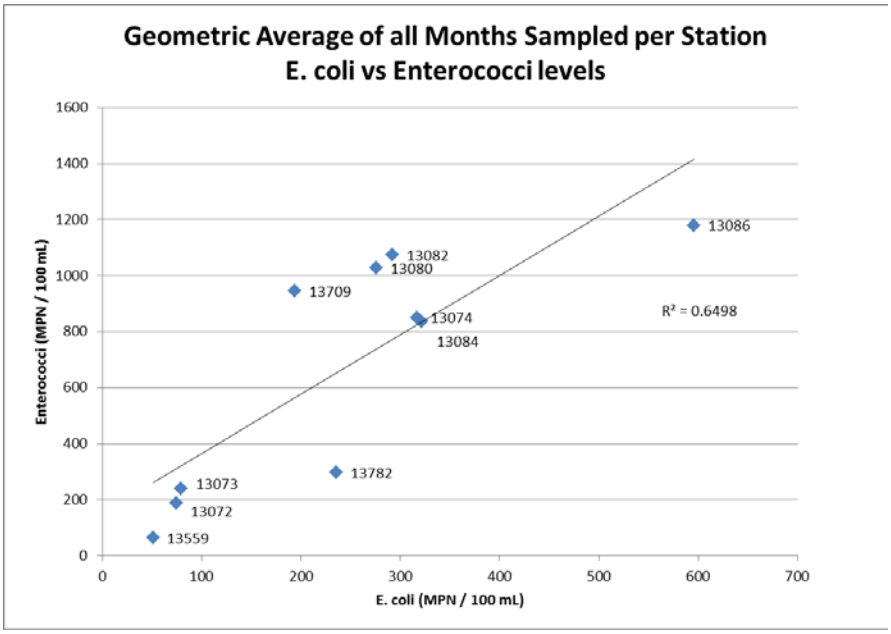


Figure 6b: Geometric Average of all Months Sampled per Station
***E. coli* vs Enterococci Levels (MPN / 100 mL)**

Data and Discussion – Vertical Profile data of Temperature, pH, Conductivity (Salinity), and Dissolved Oxygen

Vertical profile (VP) data for the upstream stations of 13086, 13084, 13082, 13080 and 13079 were frequently limited to only 1 or 2 vertical data points due to shallow depths. In fact, most were limited to one data point. In all cases, no stratification was observed within the shallow depths that ranged from 0.4 m to 1.7 m. Table 3 shows the values for these stations averaged across all months sampled. Tabular data for this dataset was limited to 0.3 m below the surface as this was the only common depth for all stations at all times. VP data for temperature and pH were as expected and of little notable interest. Salinity data showed the AC was consistently very mildly brackish with average values ranging from 2.0 ppt to 2.3 ppt. Salinity did not vary at all within the shallow depths of these stations. Dissolved oxygen (DO) values showed strong saturation values and concentration values well above desired thresholds, with DO averages varying from 6.1 mg/L for upper stations up to 8.4 mg/L for the lowest non-tidal station, 13079. Total depth in Table 3 shows the average total depth for the midline of the stream at each station for all months sampled. Monthly specific data was collected, but is not shown here.

Table 3: Values for non-tidally influenced stations averaged across all months sampled. Values shown are at 0.3 m below the surface.

	13086	13084	13082	13080	13079
Depth (m)	0.3	0.3	0.3	0.3	0.3
Temp (◦C)	22.0	22.9	22.7	23.5	23.6
pH (units)	7.3	7.4	7.5	7.8	7.8
SpCond (us/cm)	3729.3	3961.5	4293.3	4112.2	4083.5
Salinity (ppt)	2.0	2.2	2.3	2.2	2.2
DO (%)	70.7	72.5	82.5	91.0	93.6
DO (mg/L)	6.1	6.2	7.1	7.7	8.4
Total Depth (m)	0.6	0.6	0.6	0.7	1.5

Stations 13074, 13072, 13073, 13669 and 13782 had sufficient depth to reveal marked gradients with respect to salinity (conductivity) and DO. This was consistently true for all sampled months for the tidally influenced stations of 13072 and downstream. Station 13074 displayed notable salinity and DO gradients during the months of June-14, Oct-14, Dec-14, and Feb-15. During the other eight months, there was sufficient flow in the stream (due either to excessive runoff (Sep and Aug) or possibly an outgoing tide (other months) to make the depth profile uniform with respect to salinity and DO.

Tables 4a-4e show the VP parameters averaged across all months at depths of 0.3, 1, 2, 3, and the averaged greatest depth sampled (typically 0.3 m above the surface) for these stations. Reductions in DO and increases in salinity typically begin 1-2 m for all of the tidally influenced stations with marked near anoxic conditions present at 3 m and bottom sampled depths having strongly brackish to near

Tables 4a-4e: Values for stations 13074, 13072, 13073, 13559, and 13782 averaged across all months sampled at specified depths. Comparable data are limited to the 3m mark as this was the deepest common depth to all stations and months sampled. Greatest depth shown is average depth at that station minus 0.3 meters.

Depth (m)	Temp (◦C)	pH (units)	SpCond (us/cm)	Salinity (ppt)	DO (%)	DO (mg/L)
0.3	23.18	7.73	3991.38	2.22	80.56	6.87
1	23.17	7.70	4007.71	2.27	79.48	6.84
2	23.09	7.64	6467.09	3.77	70.57	6.08
3	22.82	7.49	10170.83	5.61	55.31	4.69
3.57	23.24	7.48	9149.83	5.52	57.21	4.86
Station 13074						

Depth (m)	Temp (◦C)	pH (units)	SpCond (us/cm)	Salinity (ppt)	DO (%)	DO (mg/L)
0.3	23.14	7.82	8807.78	5.28	82.43	6.94
1	23.12	7.73	10189.99	6.12	69.47	5.85
2	22.86	7.66	19314.53	12.19	35.66	3.05
3	22.47	7.63	29315.06	19.06	15.51	1.38
3.875	22.41	7.60	34621.66	22.59	6.03	0.52
Station 13072						

Depth (m)	Temp (◦C)	pH (units)	SpCond (us/cm)	Salinity (ppt)	DO (%)	DO (mg/L)
0.3	23.09	7.83	9635.41	5.81	82.77	6.87
1	23.13	7.79	10616.69	6.41	74.29	6.12
2	22.94	7.66	19072.67	12.17	39.54	3.31
3	23.15	7.65	32878.94	21.36	9.29	0.81
3.71	23.16	7.68	37211.08	24.32	5.72	0.50
Station 13073						

Depth (m)	Temp (◦C)	pH (units)	SpCond (us/cm)	Salinity (ppt)	DO (%)	DO (mg/L)
0.3	23.83	8.06	11022.66	6.56	116.84	9.30
1	24.26	7.94	13496.20	8.00	88.25	7.72
2	23.35	7.77	29125.76	19.35	22.31	1.82
3	22.63	7.81	43481.33	29.45	9.66	0.81
3.94	22.53	7.76	44918.06	30.47	6.80	0.59
Station 13559	**(Note: Data represent only 7 sampled months for this station.)**					

Depth (m)	Temp (◦C)	pH (units)	SpCond (us/cm)	Salinity (ppt)	DO (%)	DO (mg/L)
0.3	23.03	7.99	15052.49	9.09	107.26	8.70
1	23.11	8.02	18443.78	11.41	95.19	7.55
2	22.23	7.93	31031.46	20.99	59.08	4.81
3	21.87	7.93	43277.47	29.46	35.92	3.02
3.53	22.54	7.91	44314.61	29.73	26.83	2.31
Station 13782						

ocean water salinity. A similar trend held for Station 13074, but with the decrease in DO and increase in salinity beginning between 2-3 m. Additionally, DO levels did not go as low nor salinity levels as high at Station 13074 as compared to the more downstream stations.
Table 5 shows the monthly averaged vertical profile values for all sampled stations at 0.3 m for comparison purposes; furthermore, Table 6 shows the same data at 3.0 m depth for stations with sufficient depth. Figures 7a-7d illustrate the data and trends shown for the data in Tables 5 and 6 – highlighting the increased salinity and depressed DO levels at depth for stations 13074 and downstream.

Table 5: Monthly averaged vertical profile values for all sampled stations at 0.3 meters.

	13086	13084	13082	13080	13079	13074	13072	13073	13559	13782
Sample Depth (m)	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
Temp (◦C)	22.0	22.9	22.7	23.5	23.6	23.2	23.1	23.1	23.8	23.0
pH (units)	7.3	7.4	7.5	7.8	7.8	7.7	7.8	7.8	8.1	8.0
SpCond (us/cm)	3729.3	3961.5	4293.3	4112.2	4083.5	3991.4	8807.8	9635.4	11022.7	15052.5
Salinity (ppt)	2.0	2.2	2.3	2.2	2.2	2.2	5.3	5.8	6.6	9.1
DO (%)	70.7	72.5	82.5	91.0	93.6	80.6	82.4	82.8	116.8	107.3
DO (mg/L)	6.1	6.2	7.1	7.7	8.4	6.9	6.9	6.9	9.3	8.7
Avg Depth (m)	0.6	0.6	0.6	0.7	1.5	3.7	4.3	4.0	4.3	3.8
# of Months Sampled	10	12	11	12	12	12	12	12	7	12

Table 6: Monthly averaged vertical profile values at 3.0 meters for tidally influenced stations. (These stations were also the only stations with sufficient depth.)

	13074	13072	13073	13559	13782
Sample Depth (m)	3.0	3.0	3.0	3.0	3.0
Temp (◦C)	22.8	22.5	23.2	22.6	21.9
pH (units)	7.5	7.6	7.7	7.8	7.9
SpCond (us/cm)	10170.8	29315.1	32878.9	43481.3	43277.5
Salinity (ppt)	5.6	19.1	21.4	29.5	29.5
DO (%)	55.3	15.5	9.3	9.7	35.9
DO (mg/L)	4.7	1.4	0.8	0.8	3.0
Avg Depth (m)	3.7	4.3	4	4.3	3.8
# of Months Sampled	12	12	12	7	12

Data and Discussion – Instantaneous Streamflow measurement and recorded values along the Arroyo Colorado

Instantaneous streamflow data were measured at two locations -- Station 13086 (McAllen) and Station 13074 (Cemetery Rd) -- as discussed earlier. Measured flow data were augmented with data from two IBWC gages along the Arroyo – Gage #08470301 near Mercedes and Gage #08470400 near Harlingen. Measured and recorded flow data are shown in Table 7 along with an average flow value for all flow measurements or recordings collected for that location. In Table 7, stations are listed upstream to downstream (left to right). McAllen and Mercedes data are reflected from within two hours of each other. Harlingen and Cemetery Rd data reflect instantaneous data within four hours of each other. All collected data are within 24 hours of each other in order to be comparable. Averaged flow values increase as expected from upstream to downstream and with a few exceptions, all measured and recorded flow values by month

increase from upstream to downstream. It is important to again point out that Station 13074 is officially in the non-tidal segment of the AC; however, we did observe zero flow or even minor upstream flow at the 80% depth measurements near the midline of the stream at this station. The USGS small stream flow procedure calls for averaging the 20% and 80% flow depth values and this resulted in the lower than expected flow values for some months as compared to the recorded Harlingen gage flow. Figure 8 graphs the averaged flow data for each gage or station shown in Table 7.

Figures 7a-d: Graphs of monthly averaged vertical profile data at 0.3 and 3.0 meters for stations with sufficient depth.

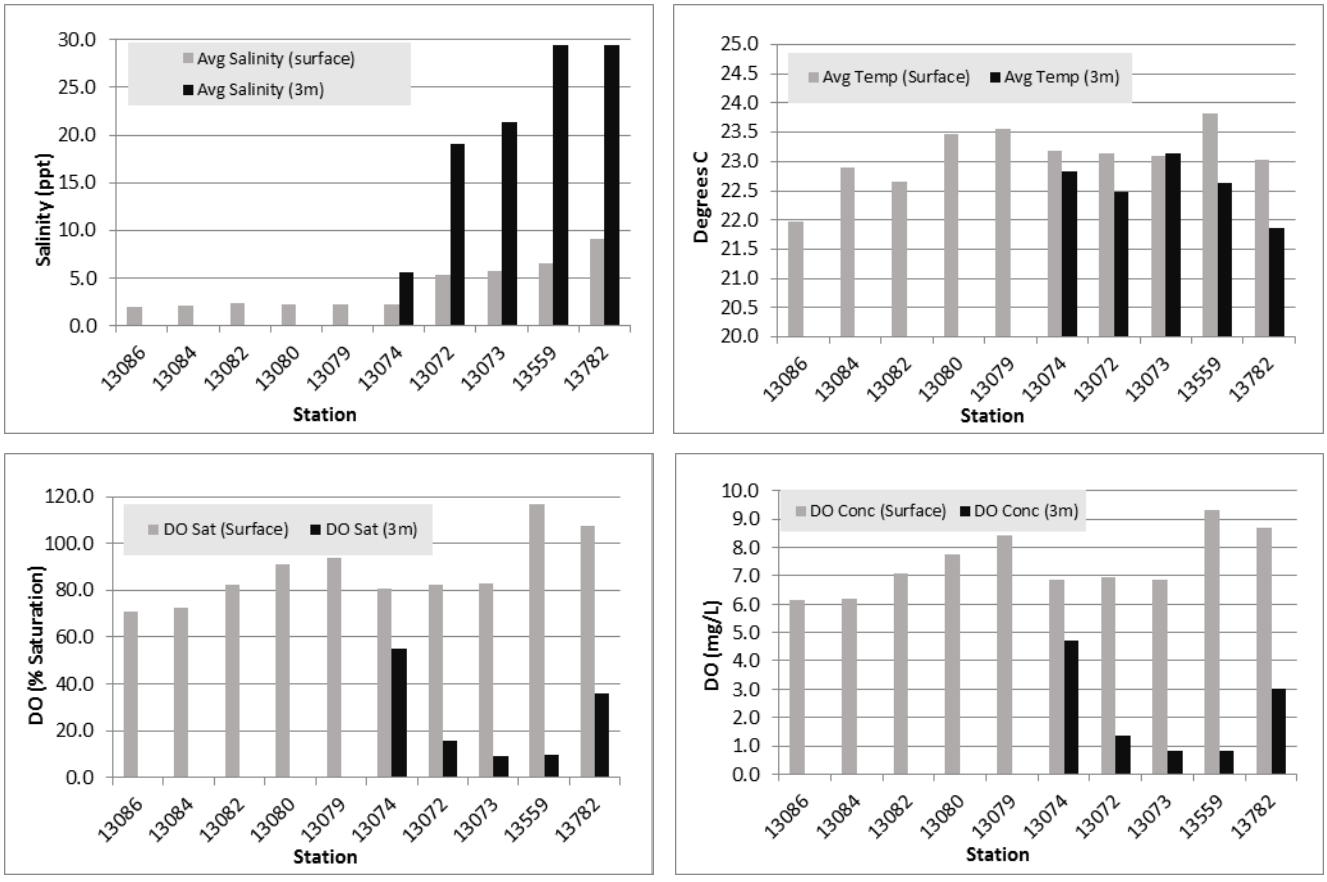


Table 7: Instantaneous streamflow (cfs) at four locations along the Arroyo Colorado. (Two measured by UTRGV and two recorded by IBWC gage.) Stations are listed upstream to downstream (left to right). McAllen and Mercedes data reflect data from within two hours of each other. Harlingen and Cemetery Rd data reflect instantaneous data within four hours of each other. These two pairs of data are within one day of each other.

Date	McAllen (Station 13086) Measured	IBWC Gage #08470301 @ Mercedes Recorded	IBWC Gage #08470400 @ Harlingen Recorded	Cemetery Rd (Station 13074) Measured
6/16/14	38	154	215	223
7/7/14	29	128	205	178
8/11/14	35	110	193	253
9/22/14		220	343	411
10/20/14	39	150	194	182
11/17/14	41	118	185	
12/15/14		107	158	
1/15/15	33	115	168	194
2/9/15	44	126	159	131
3/21/15	46	116	169	167
4/6/15	46	114	174	211
5/4/15	44	109	162	175
Average Flow Value	39	130	194	213

Appendix B – Arroyo Colorado BST Results

Bacterial Source Tracking to Support Adaptive Management of the Arroyo Colorado Watershed Protection Plan

UTSPH: Final Report on Bacterial Source Tracking

Elizabeth Casarez, Ph.D., George Di Giovanni, Ph.D., Joy Archuleta-Truesdale, Ph.D.

February, 2016

Bacterial Source Tracking

In waterbodies that exceed fecal indicator bacteria standards, a common approach to reducing monitored bacteria levels is to study the watershed and identify sources of fecal pollution and develop watershed protection plans. Laboratory tests are used by researchers to identify sources of fecal pollution, a process referred to as bacterial source tracking (BST). This process can identify different strains of *E. coli* that have adapted to conditions in the guts of their specific animal hosts, resulting in strains that are specifically associated with that species or class of animals (e.g. avian and non-avian wildlife, cattle, humans, etc.). As a result, BST laboratory tests allow the identification of likely human and animal sources of *E. coli* fecal pollution impacting a waterbody.

Two BST tests commonly used on *E. coli* are automated ribosomal ribonucleic acid genetic fingerprinting (RiboPrinting) and enterobacterial repetitive intergenic consensus sequence polymerase chain reaction (ERIC-PCR). These tests generate DNA fingerprints that resemble bar codes. The RiboPrinting and ERIC-PCR techniques are known as 'library-dependent' methods that require reference libraries of DNA fingerprints for *E. coli* isolated from known human, livestock, and wildlife fecal samples. The fingerprints of *E. coli* isolated from water samples are matched with the fingerprints in the identification library to identify the likely sources of fecal pollution.

Technical Approach

To identify the human and animal sources of fecal pollution impacting the Arroyo Colorado, ERIC-PCR and RiboPrinting composite DNA fingerprints (ERIC-RP) were generated for *E. coli* isolated from water samples. These were compared to the Texas *E. coli* BST Library, which was also supplemented with known source fecal *E. coli* isolates from the local Arroyo Colorado watershed.

Water Sample Processing

Ten locations in the Arroyo Colorado watershed were sampled monthly for one year during the study (120 scheduled measurement events). Grab samples were collected by University of Texas Rio Grande Valley team members and taken upstream of bridges when possible. Stream flow was measured within 50 meters of the sampling station depending on channel conditions. The presence of human activity, nesting birds or other wildlife was noted in the Field Log. Water samples for *E. coli* enumeration and BST were collected directly from the stream (channel midpoint or deepest accessible portion). Care was taken to avoid the surface and bottom micro-layers which may be enriched with bacteria and not representative of the water column. Immediately after collection, the sample was stored on ice for transport and delivered to the lab within 6 hours of collection.

Water samples for BST were processed by the Brownsville PUB Analytical Laboratory for *E. coli* isolation using USEPA Method 1603 with modified mTEC medium (USEPA 2006). After culturing, modified mTEC plates were shipped to UTSPH EP. Up to eight representative bacterial colonies (isolates) were isolated on Nutrient Agar with MUG (NA-MUG), confirmed as *E. coli*, and archived. Up to five isolates per water sample were then used to conduct BST analysis for identification.

Known Source Fecal Samples

Between October 2013 and October 2014, known source fecal samples were collected from the Arroyo Colorado watershed by UTRGV for the isolation of *E. coli*. Host sources were selected to supplement the Texas *E. coli* BST library. Special emphasis was placed on seabirds and small mammals. Septage and sewage samples were collected from local septage trucks and a waste water treatment plant. Livestock samples were collected during a livestock show. Mammalian wildlife samples were mostly collected from El Sauz Ranch habitat. Avian wildlife samples were mostly collected from South Padre Island. In total, 254 known source fecal samples were collected from humans and 23 subspecies of animals (see Appendix A).

Known source fecal samples were shipped to UTSPH EP within three days after collection. Samples were streaked onto modified mTEC medium. Up to five positive colonies were then reconfirmed to be *E. coli* by streaking onto NA-MUG medium. *E. coli* were successfully isolated from 99 fecal samples, and 409 isolates (up to 5 confirmed *E. coli* isolates per sample) were archived. Up to 3 isolates per sample, for a total of 274 isolates, were then screened for clones (identical isolates) using ERIC-PCR fingerprinting. The non-clonal isolates for each sample were selected for RiboPrinting and inclusion into the local watershed library yielding 144 isolates from the 99 known source samples. It should be noted that 80% of the bird samples received tested negative for cultural mobile *E. coli*.

ERIC-PCR and RiboPrinting of *E. coli*

E. coli isolates from water samples and known source fecal samples were DNA fingerprinted using a repetitive sequence polymerase chain reaction (rep-PCR) method known as enterobacterial repetitive intergenic consensus sequence PCR (ERIC-PCR) (Versalovic,

Schneider et al. 1994). Following ERIC-PCR analysis, *E. coli* water isolates and selected source isolates were RiboPrinted using the automated DuPont Qualicon RiboPrinter and the restriction enzyme *Hin*dIII. For RiboPrinting, all bacterial isolate sample processing was automated using standardized reagents and a robotic workstation, providing a high level of reproducibility. ERIC-PCR and RiboPrinting was performed as previously described (Casarez, Pillai et al. 2007).

Analysis of composite ERIC-RP DNA fingerprints was performed using Applied Maths BioNumerics software. Genetic fingerprints of *E. coli* from ambient water samples were compared to fingerprints of known source *E. coli* isolates in the Texas *E. coli* BST library (ver. 5-15) and the likely human and animal sources were identified. ERIC-RP composite patterns of water isolates were compared to the library using a best match approach and an 80% similarity cutoff (Casarez, Pillai et al. 2007). If a water isolate was not at least 80% similar to a library isolate, it was considered unidentified. Although fingerprint profiles were considered a match to a single entry, identification was to the source class, and not to the individual animal species represented by the best match. When analyzing data for the entire watershed, source classes were divided into seven groups, 1) human; 2) pets; 3) cattle; 4) avian livestock; 5) other non-avian livestock; 6) avian wildlife; and 7) non-avian wildlife, including feral hogs. When analyzing subset data (e.g. individual stations), source classes were divided into three groups: 1) human; 2) domestic animals (including cattle, other non-avian livestock, avian livestock, and pets); and 3) wildlife (avian and non-avian). The wildlife source class in this study included feral hogs since the DNA fingerprints of *E. coli* isolated from feral hog *E. coli* from wildlife rather than livestock.

Library Description

All de-cloned isolates from individual source samples (up to 3) were included in the local watershed library, independent of their similarity to other library isolates. The local Arroyo Colorado watershed library consists of 144 isolates from 99 known source fecal samples, representing 19 distinct species. Jackknife analysis of the local watershed library ERIC-RP fingerprints was used to identify the isolates that were 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. Self-validation Jackknife analysis (seven-way split) for source class specificity resulted in the selection of 75 isolates from 61 samples to form the group of self-validated Arroyo Colorado known source isolates.

The self-validated known source isolates from Arroyo Colorado were added to self-validated isolates from previous BST projects in Texas for further evaluation and possible inclusion in the current Texas *E. coli* BST Library. To increase its accuracy and utility, the Texas *E. coli* BST Library with combined self-validated local watershed libraries (2095 isolates) was refined through cross-validation. To remove cosmopolitan (non-specific) *E. coli* source isolates, repetitive Jackknife analyses of the combined self-validated libraries were performed to remove isolates that cross identified between human, domestic animals, and wildlife with the goal of 100% average rate of correct classification (ARCC) using a 3-way split of source classes. After 3 rounds, the Texas *E. coli* BST Library ver. 5-15 contains 1,765 isolates obtained from 1,554

individual fecal samples. The results of Jackknife analysis of Texas *E. coli* BST Library ver. 5-15 using a 7-way and 3-way split of source classes is included in Table 1.

Table 1: Texas *E. coli* BST Library (ver. 5-15, 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 3-way and 7-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	**384**	**330**	**22%**	**100**	**4.5**	**6**
DOMESTIC ANIMALS	**532**	**495**	**30%**	**100**	**3.3**	**19**
Pets	83	74	5%	84	16.8	41
Cattle	232	216	13%	93	7.2	11
Avian Livestock	95	88	5%	89	17.8	26
Other Non-Avian Livestock	122	117	7%	94	13.4	15
WILDLIFE	**849**	**729**	**48%**	**100**	**2.1**	**16**
Avian Wildlife	273	250	15%	79	5.3	19
Non-Avian Wildlife	576	479	33%	91	2.8	15
Overall	**1765**	**1554**		**ARCC** ** **= 3-way 100% 7-way 91%**		**18%**

*RARCC, expected random average rate of correct classification based on library composition
**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 4.5-fold greater than random chance based on library composition.

BST Results

UTRGV collected 113 water samples from the 10 sampling stations between June 2014 and May 2015. UTSPH EP successfully isolated *E. coli* from modified mTEC plates for all 113 water samples and a total of 774 isolates (up to eight per sample) were archived. Up to five isolates per sample, for a total of 529 isolates from the 113 water samples, were analyzed with ERIC-PCR and RiboPrint composite (ERIC-RP) fingerprinting and identified using the Texas *E. coli* BST Library (ver. 5-15).

E. coli BST results for all 529 watershed isolates are presented in Figure 1. Note that 90% of the water isolates were identified using the Texas *E. coli* BST Library (ver. 5-15). Given the rural nature of the watershed, it was not surprising that wildlife (both non-avian and avian) was the leading contributor of *E. coli* in the Arroyo Colorado. Approximately 9% of the isolates were identified as human and another 13% identified as domestic animals.

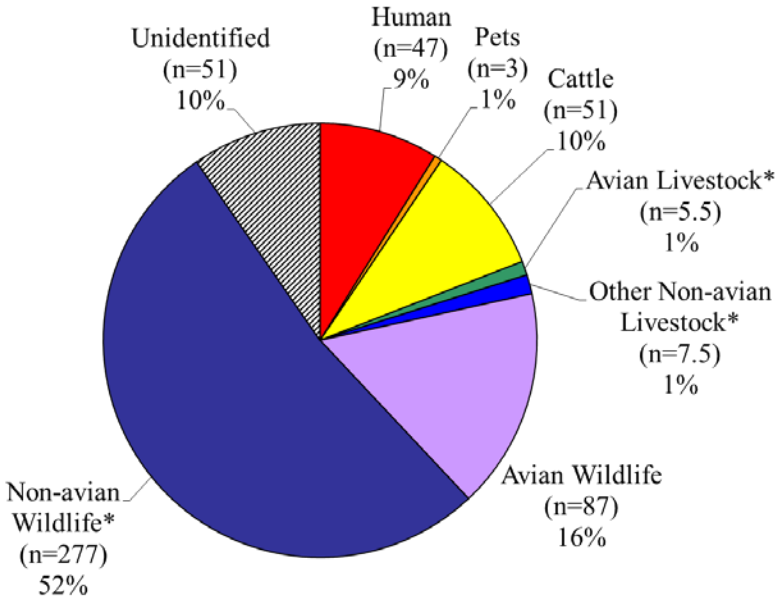


Figure 1. Identification of *E. coli* water isolates from the Arroyo Colorado watershed using a 7-way split of source classes and an 80% similarity cutoff (n = 529 isolates from 113 samples). One water isolate was equally similar to an “avian livestock” and a “non-avian wildlife” DNA fingerprint, while one other water isolate was equally similar to an “other non-avian livestock” and a “non-avian wildlife” DNA fingerprint. These were considered ties and split between the relevant source classes.

Although only the tidally influenced sampling stations must meet the enterococci water quality standard, *E. coli* and enterococci bacteria were enumerated at all ten stations along the Arroyo Colorado. During this study, all sampling stations exceeded the enterococci geometric mean standard of 35 MPN/100 mL. All freshwater stations (13086, 13084, 13082, 13080, 13079, and 13074) were above the *E. coli* geometric mean standard of 126 MPN/100 mL for the 12 months of the study. Figure 2 presents *E. coli* BST results for each station. Results are presented as a 3-

way split of sources (i.e., wildlife, domestic animals, and human), since 7-way splits typically need 80 or more *E. coli* isolates from each sampling station so that percent identifications are not greatly affected by very low numbers of isolates. Wildlife was the leading contributor at all stations. It should be noted that station 13559, one of the tidally influenced stations, was only accessible for 7 of the 12 sampling dates, and had low *E. coli* counts when sampled, and so is only represented by 22 isolates. A breakdown of the watershed by sampling station is given in Appendix B using a 3-way split of source classes, but generally all follow a similar pattern.

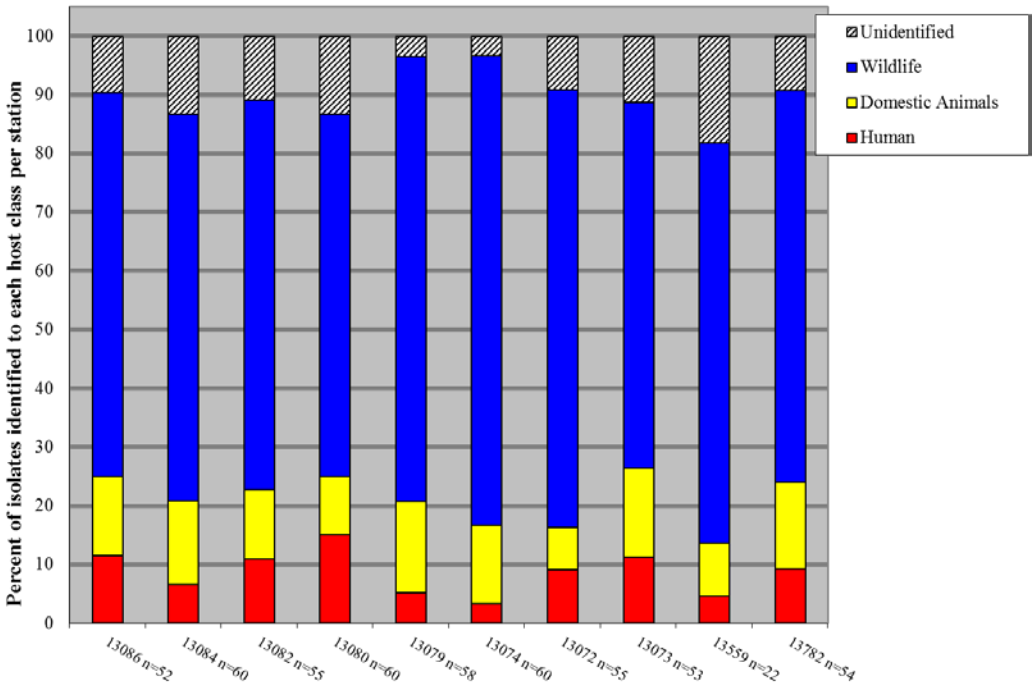


Figure 2. Three-way split of *E. coli* BST results for each station as percent of isolates per sampling station.

Since there are colonias in the area which depend upon septic systems, many of which are likely failing or not routinely maintained, the relatively low numbers of *E. coli* isolates identified as human derived was unexpected. Some known source *E. coli* isolates are considered "cosmopolitan" since they cross-identify with a known source isolate in another source class during self-validation or cross validation of watershed local libraries. However, in some cases these cosmopolitan isolates appear to be source specific during local watershed library self-validation, but do not pass cross validation between watershed libraries. Although they do not pass broader geographical and temporal scale specificity testing, at the local watershed scale they may be preferentially associated with a particular source class. Therefore, *E. coli* water isolates were also identified against the very small self-validated Arroyo Colorado local watershed library (75 isolates from 61 samples; see appendix A). Again, wildlife was still the major contributor at all stations. In contrast, the human contribution increased from 12% to 21% for station 13086 and from 11% to 24% for station 13082, providing at least some indication of more significant human fecal pollution. Since human fecal pollution poses the greatest public health risk, it is recommended that the areas surrounding the sampling stations be investigated further for potential human fecal pollution sources.

Changes in *E. coli* source identifications over the course of the study were also evaluated (Figure 3), although it should be noted that with only one year of data strong conclusions cannot be drawn. There were no significant changes in the source distribution profiles with wildlife the leading contributor and some minor fluctuations in domestic animal and human contributions.

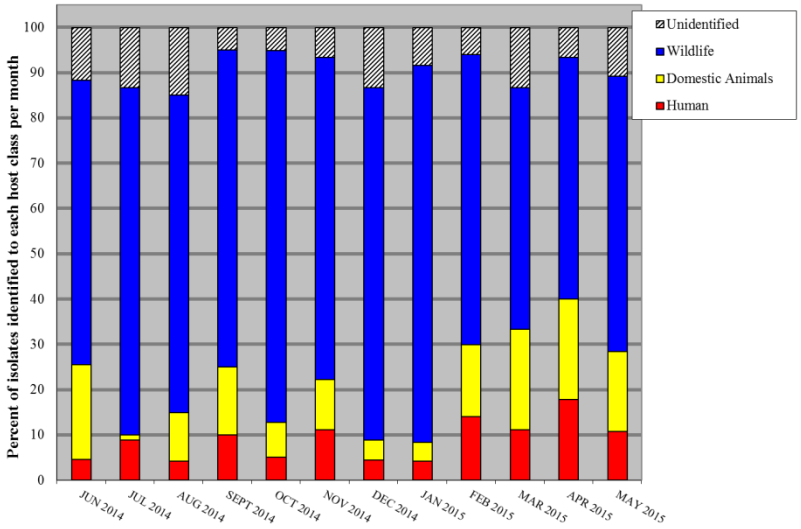


Figure 3. Three-way split of *E. coli* source class identifications by month for all stations combined.

Conclusions

Wildlife appears to be a major contributor of fecal pollution and *E. coli* bacteria at all sites. It is important to remember that wildlife can include small mammals such as rodents, raccoons, opossums, and skunks. Wildlife also includes waterfowl and other wild birds. Wildlife densities can be very high in riparian zones and are likely to have direct deposition of fecal material into waterways. These small animals may also contribute to fecal loading in urban runoff. Although rain events can greatly increase levels of *E. coli* in water, BST consistently identified wildlife as a major contributor for each month and station.

Human fecal pollution still poses the greatest health risk. Although only 9% of the total water isolates in this study were identified as human derived using the Texas *E. coli* BST Library, there is some evidence based on use of the local watershed library that stations 13086 and 13082 may have more significant human pollution impacts. Solutions to failing septic systems or sewage releases should be made a priority watershed management strategy.

Bacterial source tracking is one tool to help identify sources of fecal pollution. It is also important to incorporate *E. coli* and enterococci counts, knowledge of the local area, modeling, and common sense, in order to make sound recommendations for best management practices and an implementation plan.

REFERENCES

Casarez, E.A., S.D. Pillai, et al. (2007). "Direct comparison of four bacterial source tracking methods and a novel use of composite data sets." J Appl Microbiol **103**(2): 350–364.

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, et al. (1994). "Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction." Meth. Mol. Cell. Biol. **5**: 25-40.

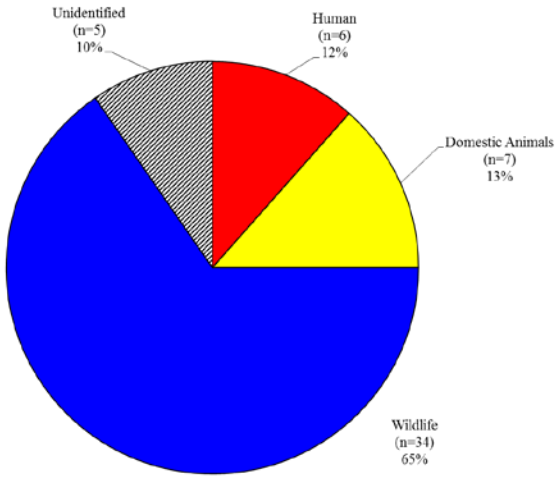
Appendix A: Known Source Samples from The Arroyo Colorado Watershed

Source	Samples collected	Samples (+) for E. coli	Isolates archived	Isolates screened by ERIC	Isolates RP in local library	Self-validated (isolate/sample)	TXSV 5-15 (isolate/sample)
Human	**21**	**21**	**104**	**63**	**49**	**27/18**	**23/16**
Sewage	6	6	30	18	16	9/5	6/4
Septage	15	15	74	45	33	18/13	17/12
Cattle	**6**	**5**	**25**	**15**	**5**	**2/2**	**1/1**
Other non-avian livestock	**17**	**15**	**72**	**45**	**19**	**4/4**	**1/1**
Goat	3	3	15	9	4	1/1	0/0
Pig	5	5	24	15	7	2/2	0/0
Rabbit	5	3	13	9	4	1/1	1/1
sheep	4	4	20	12	4	0/0	0/0
Other avian livestock--chicken	**4**	**4**	**14**	**10**	**7**	**2/2**	**2/2**
Pets--dog	**1**	**0**	**0**	**0**	**0**	**0/0**	**0/0**

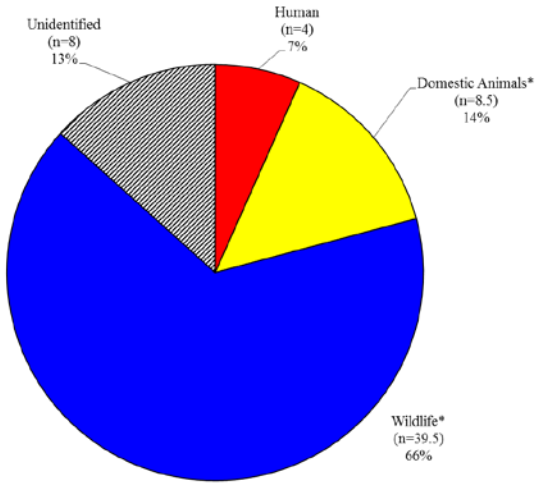
continued next page

Source	Samples collected	Samples (+) for E. coli	Isolates archived	Isolates screened by ERIC	Isolates RP in local library	Self-validated (isolate/sample)	TXSV 5-15 (isolate/sample)
Avian Wildlife	**155**	**31**	**111**	**78**	**37**	**22/20**	**15/14**
Black tern	61	3	13	9	4	3/3	2/2
Ruddy turnstone	1	1	5	3	1	0/0	0/0
Bird	33	12	21	21	15	11/10	9/8
Western snowy plover	6	0	0	0	0	0/0	0/0
Herring gull	2	0	0	0	0	0/0	0/0
Laughing gull	7	2	10	6	3	2/1	1/1
Owl	1	1	5	3	1	0/0	0/0
Pelican	17	7	35	21	8	4/4	2/2
Pigeon	1	0	0	0	0	0/0	0/0
Sparrow	1	0	0	0	0	0/0	0/0
Royal tern	22	5	22	15	5	2/2	1/1
Wren	2	0	0	0	0	0/0	0/0
Unmarked species	1	0	0	0	0	0/0	0/0
Non-avian Wildlife	**50**	**23**	**83**	**63**	**27**	**18/15**	**17/14**
White-footed mouse	39	18	64	49	21	15/12	14/11
Norton pygmy mouse	2	0	0	0	0	0/0	0/0
White-tailed deer	2	2	8	6	2	1/1	1/1
Hispid cotton rat	2	1	4	3	1	1/1	1/1
Wild rabbit	1	1	5	3	1	0/0	0/0
Unknown mammal	4	1	2	2	2	1/1	1/1
Total	254	99	409	274	144	75/61	59/48

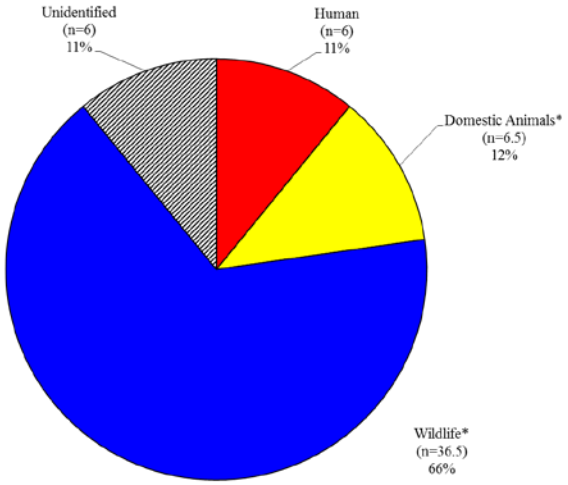
Appendix B: Source Identifications by Sampling Station



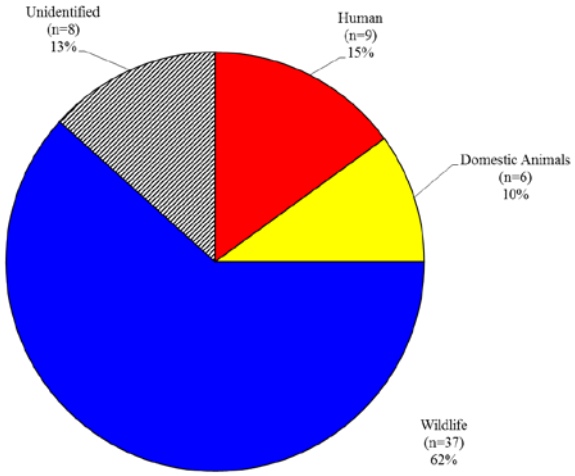
Station 13086: geometric mean = 596 MPN/100 mL (n=52 isolates; 11 samples)



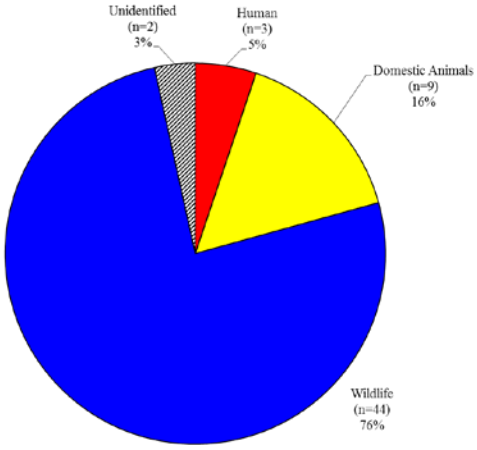
Station 13084: geometric mean = 321 MPN/100 mL (n=60 isolates; 12 samples)
One water isolate was equally similar to a “domestic animals” DNA fingerprint and a “wildlife” DNA fingerprint. This was considered a tie and split between the two source classes.



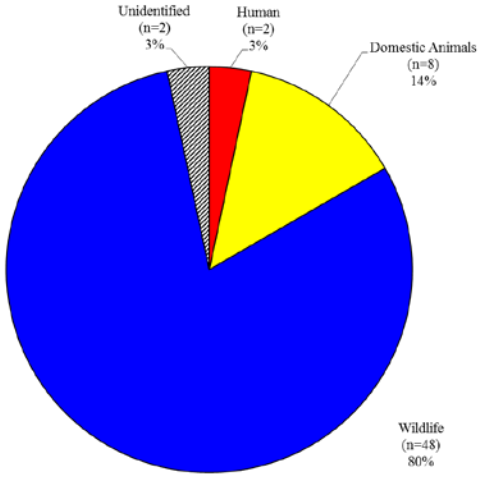
Station 13082: geometric mean = 292 MPN/100 mL (n=55 isolates; 11 samples)
One water isolate was equally similar to a “domestic animals” DNA fingerprint and a “wildlife” DNA fingerprint. This was considered a tie and split between the two source classes.



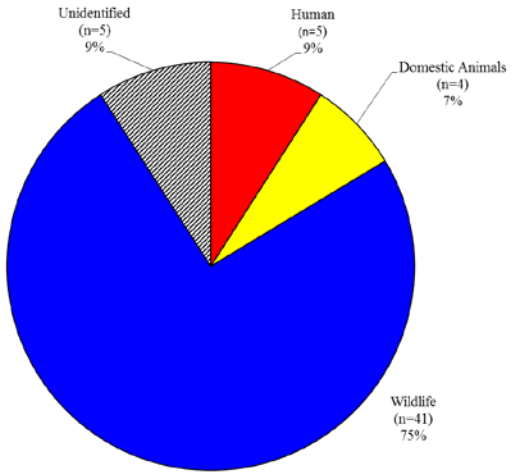
Station 13080: geometric mean = 276 MPN/100 mL (n=60 isolates; 12 samples)



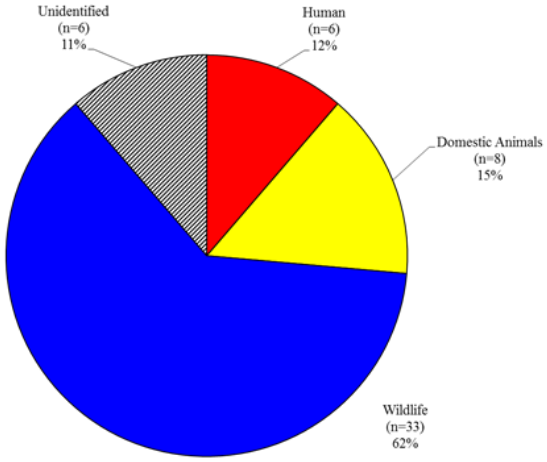
Station 13079: geometric mean = 194 MPN/100 mL (n=58 isolates; 12 samples)



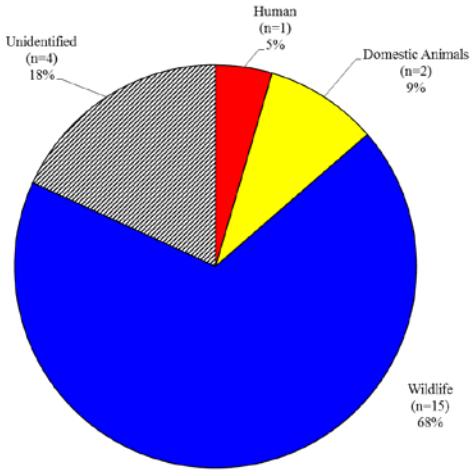
Station 13074: geometric mean = 317 MPN/100 mL (n=60 isolates; 12 samples)



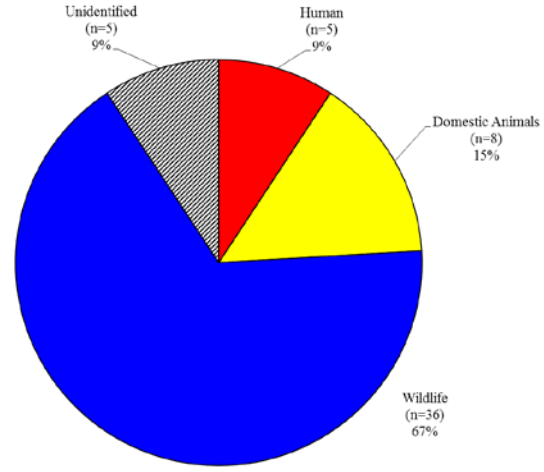
Station 13072: geometric mean = 74 MPN/100 mL (n=55 isolates; 12 samples)



Station 13073: geometric mean = 79 MPN/100 mL (n=53 isolates; 12 samples)



Station 13559: geometric mean = 51 MPN/100 mL (n=22 isolates; 7 samples)



Station 13782: geometric mean = 235 MPN/100 mL (n=54 isolates; 12 samples)