Use of *Escherichia coli* for Microbial Source Tracking in a Mixed Use Watershed in Northern Virginia

Timothy R. Wade II

Thesis submitted to the faculty of Virginia Polytechnic Institute and State University in partial fulfillment of the requirements of the degree of

Master of Science
In
Crop and Soil Environmental Sciences

Dr. Charles Hagedorn III
Dr. Carl Zipper
Dr. Golde Holtzman

August 27th, 2007
Blacksburg, VA

Keywords: Microbial source tracking, Antibiotic resistance analysis, *E. coli*, Prince William County, Water quality, Fluorometry, Optical Brighteners

Use of *Escherichia coli* for Microbial Source Tracking in a Mixed Use Watershed in Northern Virginia

Timothy R. Wade II

Abstract

Prince William County, located in the rapidly developing Northern Virginia region, contains watersheds of mixed rural and urban/suburban uses. The project goal was to monitor and evaluate 21 stream locations, over 13 months, in the Occoquan Basin identified as impaired due to high *E. coli* densities. One site on each of eight streams, two sites on each of five streams, and three sites on the remaining stream were chosen for *E. coli* monitoring and microbial source tracking (MST). MST was performed using antibiotic resistance analysis (ARA) and fluorometric analysis. *Escherichia coli* was chosen as the indicator bacterium for purposes of comparison with previous project data and because a large body of evidence supports its use in freshwater systems.

This study involved the only known MST project to incorporate data from five or more consecutive years. A total of 2854 environmental isolates were collected for analysis with ARA. These isolates were classified using a known source library (KSL) that consisted of 1003 unique resistance patterns. The resistance patterns of the KSL came from known fecal sources (human, pets, livestock, wildlife) in Prince William County. The KSL included isolates from previous years but was also updated with fresh isolates. The accuracy of the KSL was assessed through the use of a challenge set. The challenge set was classified against the KSL using discriminant analysis, verified by logistic regression. The average rate of correct classification was 93% for discriminant analysis and 96% for logistic regression.

Results indicated that multiple sources of contamination were present at all sampling locations and that the major source(s) (human, pets, livestock, wildlife) of contamination were generally related to the land-use patterns and human activities at each location. Although no major or minor human signatures were found, all but two locations had either pet or livestock as the major signature, suggesting that human-related activities are playing a key role in contamination of the streams. Pets were the single most frequent major signature and wildlife was the most common minor signature.

Fluorometric analysis was used to corroborate human-derived contamination. Fluorometric analysis has the ability to detect the presence of optical brighteners, synthetic compounds added to such household items as laundry detergent, dishwashing detergent and other washing agents. Despite having an undesirably high rate of false negatives (negative fluorometry readings not supported by ARA), fluorometric analysis maintained a low rate of false positives (positive fluorometry readings not supported by ARA) and continued to demonstrate its potential for source tracking.

This project represented one of the first attempts at applying a full suite of performance criteria now recommended by the source tracking community for all MST projects. Such concepts as experimental design, toolbox approach, minimum detectable percentage, quantification, accuracy, specificity, robustness, range of applicability, and practicality were successfully incorporated. These performance criteria have in effect set a new standard to which all subsequent MST projects should adhere.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank, first and foremost, Dr. Charles Hagedorn III for his incalculable help and guidance throughout the entire course of this project. His expert advice, along with the suggestions of fellow committee members Dr. Carl Zipper and Dr. Golde Holtzman, made this project a reality. The other members of the lab team, Annie Hassall, Mike Saluta, Jay Dickerson, Paul Youmans, Justin Evanylo and James Deykes assisted with the daily grind. Their unique senses of humor and perspectives on life served to lighten the mood.

The consistent, loving support of my friends and family provided a foundation I could depend on. I would especially like to thank my parents, Sarah and Sarah's parents, Steve and Sue, for all their assistance and encouragement along the way. I am truly blessed.

Finally, I would like to express sincere gratitude for the funding provided by the Prince William County Department of Public Works. I am particularly indebted to Mrs. Patty Dietz and Ms. Amelia Warren for their assistance with sampling and extensive knowledge of Prince William County.

Chapter 1. Literature Review

I. Project History and Justification

The Clean Water Act of 1972 states that it is the responsibility of the federal government to identify impaired waters if state governments do not. This was the principal argument of several grassroots organizations when they sued the EPA beginning in the 1980s (22). These organizations argued that the EPA must set total maximum daily load (TMDL) restrictions. TMDL is the maximum amount of a pollutant that a body of water can receive in one day and still meet the water quality standards (22). TMDLs must include a plan to lower the amount of pollutants for bodies of water that do not currently meet water quality standards. The grassroots organizations won the lawsuit and the court-ordered consent decree mandated several changes take place. First, the EPA now ensures TMDL amounts are set and reviews Virginia's plans for each year. Secondly, the federal government must use all available resources to identify the most heavily polluted waters of a state (22). The Virginia Department of Environmental Quality (DEQ) develops and maintains lists of impaired waters, including the pollutant responsible for the violations, and the cause and source of the pollutant (67).

An impaired body of water is defined as one that does not satisfy one or more of the five uses of water designated by the EPA: aquatic-life habitat, drinking-water supply, fish consumption, swimming, and shellfishing. The water quality standards set by the EPA determine whether a body of water is impaired or not (33). As of March 2004, 6301 miles (10,138 km) in the Commonwealth were classified as impaired, and most of the impairments result from violations of fecal bacterial water quality standards. Specifically, these streams do not meet the standard set for *E. coli* (33). The instantaneous, or single-sample, standard for *E. coli* obtained from freshwater is 235 colony forming units per 100 milliliters (CFU/100mL). If

more than 10% of the samples collected during a given period exceed the standard, the body of water is designated as impaired and therefore listed in Section 303(d) of the Clean Water Act (22).

Once a body of water has been classified as fecal-impaired, appropriate measures must be taken to identify the major source(s) of contamination. Pollution from fecal matter is the most pervasive problem in the rivers and streams of Virginia (22). Fecal matter can enter a watershed from point or non-point sources. A point source can be defined as any cause of contamination that can be identified to an exact or near-exact location, such as a failing sewage treatment plant. A non-point source, on the other hand, is an origin of contamination that can not be as easily pinpointed. The area may be much broader and more difficult to define. Among many other examples, non-point sources of pollution can include urban or agricultural runoff, or combined sewer overflows (22).

The waters of this project are officially listed as “designated use” and the standards are maintained in order to protect humans while swimming, fishing, or performing other water-related activities. However, in addition to their designated uses, these waters are important for several reasons. Prince William County is part of the Chesapeake Bay watershed and many of its waters flow into a reservoir used for drinking water. Both of these reasons further the mandate for the monitoring of fecal contamination.

This project in Prince William County began in 2003 and has continued into its fifth consecutive year. Monitoring began in July 2003 with eighteen sites sampled on a monthly basis for 12 months (34). In general, monthly monitoring occurs for the first twelve months for any new site, and is recommended, but not required, to continue on a quarterly basis for at least one additional year if more than 10% of the samples during the initial 12 month period exceeded the

DEQ standard. The only exceptions for the current project are sites Q4 (Cow Branch at Montgomery Avenue) and Q18 (North Fork of Lake Manassas), which were chosen to begin in September 2006 as quarterly sites despite not having prior monitoring data. This was a decision made by local officials and has no effect on any other site.

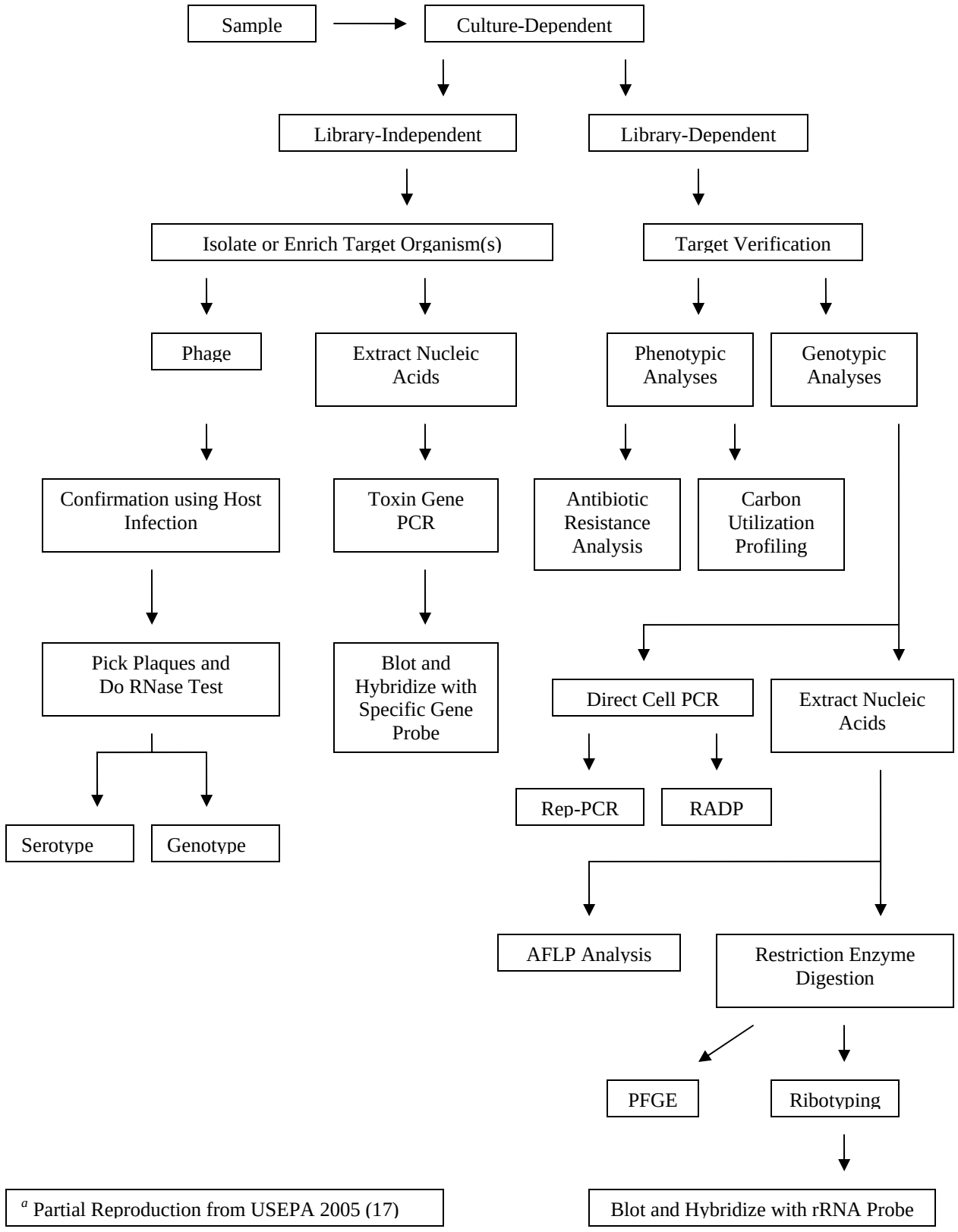
The results of the first year depicted a troubling scenario. All eighteen sites exceeded the DEQ standard more than 10% of the time. Interestingly, the source tracking results showed wildlife to be the largest contributor of fecal contamination, with birds, such as geese and gulls, being second (34). The sponsors elected to continue sampling at four locations, on a quarterly basis (31). Two of these locations, Cedar Run at Carriage Ford Road and Neabsco Creek at Benita Fitzgerald Road, are still monitored on a quarterly basis today. Meanwhile, ten new sites were added for monthly sampling during the second year (June 2004 – June 2005) of the project. Nine out of ten of these sites exceeded the standard and were therefore placed on the impaired waters list. Additionally, wildlife continued to dominate as the major contamination source and this was seen again during the third year of the project (30, 32). For the third year (September 2005 – June 2006), the sponsors chose to continue quarterly monitoring of six of the nine sites from year two. All six of these sites are still being monitored on a quarterly basis today. The remainder of the third year project consisted of quarterly monitoring of fourteen sites from the first year (32). Of these fourteen, all but four are still monitored on a quarterly basis today. This amounts to a total of eighteen sites that were monitored on a quarterly basis during the latest year, June 2006 to June 2007. Furthermore, the sponsors added three new sites that were monitored monthly. Both the monitoring and source tracking results argued for continued monitoring and investigation of the sources of contamination in Prince William County. The methods used in this project have been consistently demonstrated as reliable and reproducible

(35, 38, 74). Significant questions remain concerning whether wildlife will continue to predominate or if new sites will show evidence of other sources of pollution, especially as landscape use changes in a county that is undergoing rapid suburban development.

II. History and Development of Antibiotic Resistance Analysis

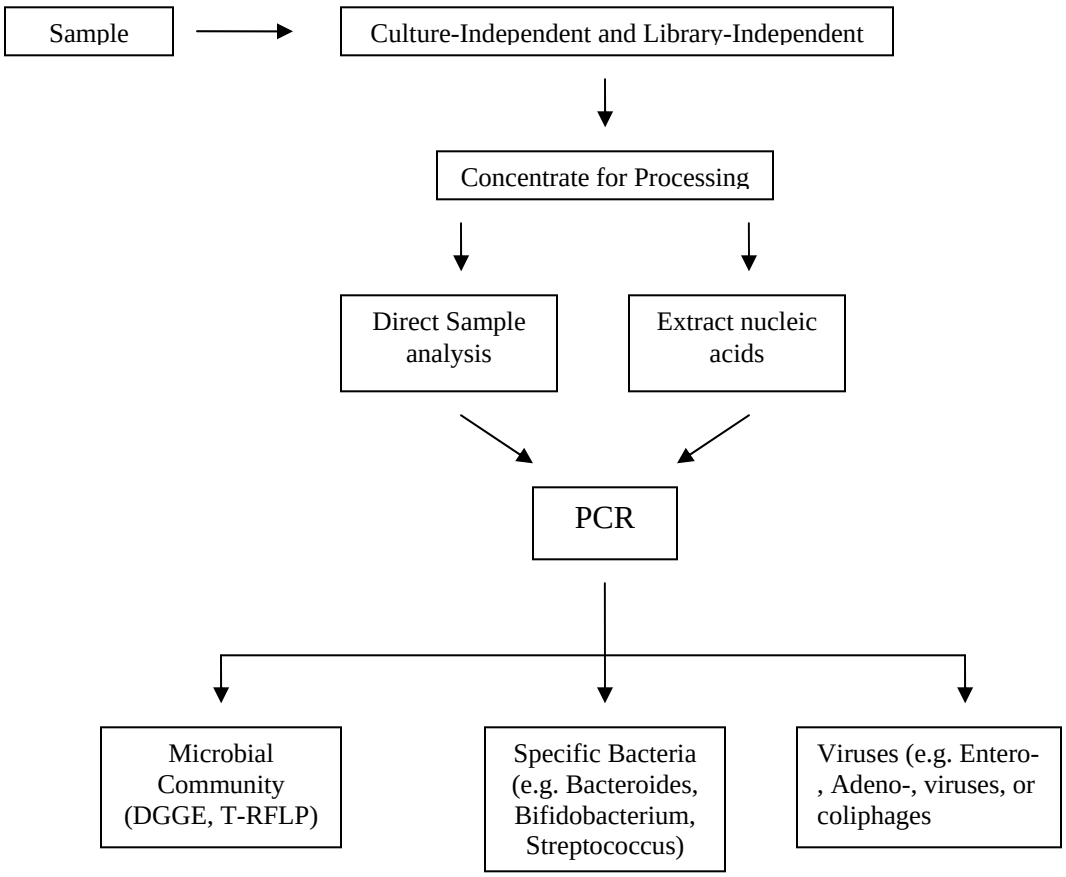
As a means for confronting the increasing problem of fecal contamination in aquatic systems, microbial source tracking (MST) has made great strides over a relatively short period of time. Also known as bacterial source tracking, MST can be conducted by a wide range of methods. In general, these methods can be classified as library-independent or library-dependent. An even broader means of categorization is to begin with whether the technique is culture-dependent or culture-independent. Examples of library-dependent genotypic-based methods include ribotyping, pulsed-field gel electrophoresis, and several PCR-based techniques (20, 48, 62). Figures one and two illustrate an organized breakdown of the current and most popular MST techniques (17). The majority of these techniques are relatively new and continue to demonstrate a clear need for improvement. Phenotypic techniques, such as ARA, can trace their origin to an era pre-dating the first genotypic methods (43). As a pioneer in the field of MST, ARA has garnered considerable merit. Despite limitations such as the need for a large known source library (KSL), or the inability to provide real-time results, ARA continues to substantiate itself as a highly capable source tracking tool (2, 15, 61). The goal of this section is to provide a comprehensive assessment of ARA: its history and development, how it has been applied in the field, and a forecast of how it may evolve.

Figure 1. Culture-Dependent Source Tracking Methods[a]



[a] Partial Reproduction from USEPA 2005 (17)

Figure 2. Culture-Independent Source Tracking Methods[a]



[a] Partial Reproduction from USEPA 2005 (17)

A. Selecting an Appropriate Indicator Organism

The purpose of any MST method is to provide accurate clues as to the potential sources of pollution in a body of water. The majority of methods require selecting for specific indicator organisms. The indicator organism must be ubiquitous among all possible sources of contamination, and is generally chosen because of its relative ease of detection, persistence in the environment, and association with more potentially harmful microbes (59). The earliest attempts used fecal coliforms to compare potential patterns of antibiotic resistance between Gram negative bacteria. Kelch and Lee (43) concluded that a generally high correlation, with respect to patterns of antibiotic resistance, existed among fecal coliforms from different sources as well as among fecal coliforms of different genera.

As knowledge about enteric bacteria grew, researchers began to look at the fecal streptococci group. Dr. Bruce Wiggins, the first to coin the term "antibiotic resistance analysis," made several groundbreaking conclusions in his earliest source tracking effort (72). Wiggins was the first to recognize the importance of a substantially large library of known source isolates. Despite average rates of correct classification (ARCC) as high as 98%, he acknowledged that more samples would improve the statistical analysis (72). In a follow-up investigation (73), Wiggins increased the number of known source fecal samples from 17 to 236 and nearly doubled the number of antibiotics. Encouragingly, comparable results were obtained. Depending upon the combination of antibiotics analyzed, the ARCCs ranged from 64 to 78%, when classified into one of four possible categories (cattle, human, poultry, or wild).

Scientific procedures and results are only as strong as the ability to reproduce them. As part of a watershed improvement project, Hagedorn et al. (35) constructed a KSL of 7,058 isolates, using 147 samples and thirteen antibiotics, the highest number to date. When tested

against itself, the library produced ARCCs ranging from 78 to 93%, with 10% being the highest misclassification percentage of any category. This study was one of the first to take the next step: to apply a source library as a means to identify unknown sources of fecal contamination in a body of water. Moreover, when cattle were found to be the predominant source of fecal contamination, fencing of the streams resulted in drastically reduced fecal coliform densities and an almost two-fold decrease in the percentage of fecal streptococci isolates classified as cattle. As the first test of ARA in subtropical waters, Harwood et al. (38) duplicated the procedure designed by Wiggins (72). The most notable difference was an increase in the number of antibiotics (from 5 to 9), and this was attributed as the reason for higher ARCCs. A theme of "more antibiotics equals better classification" was beginning to come forth. When tested in the field, the library from this study produced results consistent with expectation. The independent verification by multiple laboratories offers strong evidence for the reliability and effectiveness of ARA.

As a subset of the fecal streptococci group, the enterococci are an even more specific indicator of fecal contamination. Butaye et al. (8) compared farm animals to pet animals by determining the antibiotic resistance patterns of *Enterococcus faecalis* and *Enterococcus faecium* strains isolated from fecal matter. Overall rates of resistance were higher for farm animals, especially for those known to have been subjected to growth-promoting and/or therapeutic antibiotics. A separate study suggested that communal bathing may expose individuals to high levels of resistant bacteria (3). Graves et al. (27) were one of the first to apply this knowledge to the source identification of fecal pollution in a body of water. As a follow-up to the studies involving fecal streptococci, stream samples were collected from various locations surrounding a rural, nonsewered community. Results were consistent with expectations. A human signature

(10%) was found, but livestock predominated as the principal source of contamination in more than two-thirds of the samples. Evidence continues to demonstrate, therefore, that in terms of antibiotic resistance patterns, inherent differences exist between humans and animals (26).

Perhaps the most heavily studied fecal indicator is *Escherichia coli*. One of the earliest source tracking papers (42) used resistance patterns of *E. coli* to construct multiple antibiotic resistance (MAR) index. It was suggested that isolates from the same sampling area that expressed identical or very similar MAR indices may have originated from the same source of contamination. Although the total isolate number of 202 isolates was small, these results offered new clues for future research. A larger-scale study (50) concluded that human *E. coli* contamination is more strongly associated with point source isolates, such as those from a failing onsite septic system, while animal *E. coli* contamination correlates more strongly with non-point source isolates, such as those originating from a large pasture. In general, it was shown that human isolates have a wider variation in resistance than animals. Fogarty et al. (21) were the first to analyze densities and MAR profiles of both *E. coli* and enterococci in the same study. In accordance with the conclusions of earlier efforts, sufficient representation of the population characteristics of *E. coli* and enterococci is best achieved by the development of a large library of isolates.

The majority of later ARA projects have elected to use either *E. coli* or *Enterococcus* as the fecal indicator of choice (2, 6, 9, 10, 13, 15, 16, 18, 23, 28, 29, 39-41, 44, 49, 52, 55, 60, 62, 68, 70, 74). Both require relatively simple and inexpensive culturing techniques (59), are ubiquitous in the guts of warm-blooded animals, and have been shown to be correlated with the presence of deadlier pathogens (66). *Enterococcus* may become the standard indicator for

marine environments nationwide, while *E. coli* continues to prove effective for freshwater systems (56).

B. Antibiotic Resistance as a Measurement Tool

As the underlying mechanism behind ARA, antibiotic resistance must be an effective and verifiable measurement tool. However, a recurring criticism of ARA is that antibiotic resistance is a fleeting characteristic of bacteria, that it changes too quickly over time and therefore cannot offer a valid means of pollution detection. Nevertheless, there is a growing suite of evidence that argues in favor of the persistence of resistant bacteria. Several studies have been performed that show antibiotic resistance to be a naturally occurring, universal phenomenon able to survive within its host organism long enough to be accurately detected in water. Díaz-Mejía (14), for example, presented new evidence that resistance patterns do not change even if antibiotic usage changes. Other mechanisms need to be explored in order to explain why resistance can be maintained in countries such as Mexico that does not impose strict regulations on the use of antibiotics, and Cuba, where the United States embargo has severely limited the availability of antibiotics. Resistance levels were found to be the same in both countries (14). Amábile-Cuevas (1) also argued that indiscriminate use can select for novel, tenacious resistance strains, such as the troublesome vancomycin-resistant enterococci discovered in 2002. One possible explanation is that virulence and resistance may result from the same bacterial mechanism. The same processes that allow bacteria to survive free-radicals emitted by an immune system response may also shield them from antibiotics. Additionally, evidence is mounting in favor of the horizontal transfer of resistance genes between bacteria of different species and habitats (53).

Another criticism is that contamination from livestock raised without antibiotics would go undetected. However, even on organic farms, where strict regulation of antibiotic use is a major tenet, resistant bacteria have been shown to persist. In one three-year study, ampicillin resistance ranged from 27.3% to 40.7%, with multiple resistance found in more than 44% of isolates tested (40). Another study compared *E. coli* resistance patterns among swine of different ages and level of antibiotic use. It was concluded that, while resistance differed depending upon age and farm type, resistance patterns remained relatively constant within each individual farm and animal over the course of the study (46). Scientists may not have the knowledge to adequately explain these observations, but that does not mean they are not occurring. While the resiliency of antibiotic-resistant bacteria may be alarming to certain aspects of society, it offers an effective device for the analysis of polluted bodies of water.

III. ARA in the Field

Construction of a library of known source isolates is an essential first step for ARA, and the library can be applied in the field as a means to classify the unknown sources of fecal contamination that may exist in a body of water. Booth et al. (6), for example, implemented a KSL in four sub-watersheds of a major Virginia watershed encompassing approximately 72,000 ha. In a rural region dominated by dairy and beef cattle, the results were consistent with expectations. A KSL of 1,451 *Enterococcus* isolates was constructed, and upon discriminant analysis, was shown to be representative of the major sources of potential fecal contamination: humans, livestock, and native wildlife. The ARCCs for the library were 82.3%, 86.2%, and 87.4% for human, livestock, and wildlife, respectively. When used to assess a contaminated

stream, the library showed livestock to be responsible for almost half of the pollution. All three categories, however, were high enough to warrant further investigation.

A 2,491 enterococci isolate library was applied to marine water with an established fecal contamination problem (10). The known sources consisted of bird feces, urban runoff, coastal marsh sediment, and sewage effluent. Only the ARCC for sewage (64.5%) fell below 70%, and was attributed to the fact that the sewage may not have been exclusively human. When used to classify the environmental water column samples, the library assigned 30% to bird feces, 24% to sediment, 6.5% to urban runoff, and 39.3% to sewage. A single source of contamination did not predominate. However, when reclassified after re-dividing by date collected, different sources predominated on different dates. This was one of the first papers to suggest the concept of temporal variability. This paper is supported by the results of an earlier study (71) that found fecal coliform resistance patterns to be highly variable over a period of only seven months. Within the urban watershed, wild animals were classified as the dominant source during high bacteria loads, but humans were the major source during lower bacteria levels. Fecal coliform levels can undergo significant fluctuation depending on the season. The latest attempt by Wiggins (74) offered additional insight into the temporal and geographical stability of ARA libraries. With a total of 6,587 isolates, this was the biggest library to date, and the first to examine whether libraries built from different watersheds can be merged to create a single multi-watershed library. The merged library effectively represented each of the six contributing watersheds, and maintained comparable ARCCs for a period of one year.

ARA has been successfully implemented by the international source tracking community. For example, Carroll et al. (9) applied ARA to two mixed land use locations in the Gold Coast region of Australia. The objective was to determine whether the high number of onsite

wastewater treatment systems (OWTS) was the primary contributor of fecal contamination. It was concluded that nonhuman contamination predominated in rural areas, but more urbanized regions that used OWTS were marked by significantly higher percentages of human *E. coli* isolates. Interestingly, the largest human signatures occurred during drier sampling conditions.

A project in the United Kingdom designed an automated ARA method capable of digitally recording antibiotic resistance results (16). Using 21 antibiotics, perhaps the most of any project, this was the first ARA paper to be published in the UK. With 2,195 isolates, ARCC results comparable to other projects (35, 38, 72) were obtained. Similarly, the highest ARCCs were obtained when fewer categories were used. Importantly, a blind challenge set was also successfully implemented, providing further evidence that the KSL was representative. A challenge set tests the ability of a KSL to classify known source isolates not already included in the library. For example, a challenge set might involve using a KSL to classify deer isolates as if they were unknown isolates. Assuming the challenge isolates are from the same watershed, the KSL should produce ARCCs comparable to running the library against itself. The UK project used 425 challenge isolates and correctly assigned 85% of them (16). Further work will investigate the potential of this library to identify unknown sources of fecal contamination.

Canada is emerging as a leader in the source tracking community. One of the most recent ARA-related projects (36) attempted to integrate DNA micro-array technology with antimicrobial resistance genes in order to provide additional clues about the distribution of *E. coli* in surface waters. The highest percentage of *E. coli* resistance genes were found at a site directly downstream from a municipal wastewater treatment plant, suggesting a correlation between human antibiotic usage and *E. coli* resistance in water. Edge and Hill (18) designed a study to compare antibiotic resistance patterns of various waterfowl species to humans. When

used to discriminate between possible sources of pollution, the KSL showed *E. coli* from the waterfowl to be the major contaminant. A separate Canadian study demonstrated the usefulness of *C. perfringens* as a fecal indicator, concluding that it would be effective for general fecal contamination events (11). How well ARA may work with *C. perfringens* remains to be seen.

IV. The Future of ARA and MST

A. Understanding the Limits of ARA

Like any scientific technique, ARA requires meticulous method development and thorough understanding of how it can be applied. The papers mentioned previously had certain key elements in common. These included a large KSL relative to the size of the watershed, adequate sampling of all possible known sources, and an understanding of the inherent limitations of ARA. One methods comparison study, for example, concluded that a technique known as amplified fragment length polymorphism (AFLP) was superior to both ARA and sequence analysis of bacterial 16S rRNA genes (29). The application of ARA in this paper, however, was critically flawed. The most egregious failing was considering representative the collection of only 319 *E. coli* isolates, over a widespread region. Similarly, Kelsey et al. (44) concluded that several surface water locations had human-source contamination, but failed to build a library representative of all potential sources. Experimental bias can become an overwhelming factor if an MAR index does not account for the various contamination sources in a watershed study.

Another criticism is that culture-dependent methods are unable to analyze watershed quality in a timely manner (59). Acquiring rapid results has never been the goal of ARA. Rather, ARA has been successfully applied for multiple projects aimed at assessing the long-

term anthropogenic effects on a watershed (15, 27, 68, 74). It is not always necessary to immediately understand the sources of contamination, and ARA offers a relatively quick method at a fraction of the cost of the more expensive genotypic methods. A stronger appreciation of all elements required for ARA will better serve the scientific community and the public at large.

B. ARA vs. Other MST Methods

ARA has been included as a part of several review papers and methods comparison studies. Scott et al. and Seurinck et al., for example, agree that ARA requires the construction of large known source databases consisting of the most significant sources that fall within the scope of the project (56, 57). This requirement applies to virtually all library-based MST methods. Additionally, Meays et al. argued that isolate-level analysis is preferable to sample-level analysis, as it is doubtful that an environmental water column sample was contaminated by only one major source (47).

The United States Geological Survey (USGS) conducted a large-scale methods comparison study (62), hoping to evaluate the current status of seven different techniques. The comparison, however, was plagued by more than one significant shortcoming. Perhaps the single biggest failing was that ARA did not fare as well as other papers have demonstrated, but neither did any of the other six source tracking techniques. Not a single technique produced results consistent with expectations. Furthermore, the study only involved library-dependent techniques. A better design would have included a method such as community terminal restriction fragment length polymorphism (T-RFLP), which is culture-independent and does not require a known source library (17, 20). T-RFLP was one of several techniques tested by a California collaborative study. Included in this study was an assortment of phenotypic methods,

such as ARA using various indicator organisms, MAR, the Kirby-Bauer antibiotic susceptibility test, and carbon source utilization profiling. Interestingly, the best phenotypic-based results came from ARA when fecal streptococci were used as the indicator. ARA using either *E. coli* or *Enterococcus* also performed well, with true-positive percentages of 86.7% and 80.0%, respectively (39). Overall, the study argues that the future of MST lies in improvements in method optimization, development, and evaluation (60). Specifically, with regard to ARA, the need for a challenge set is becoming absolute (61). A challenge set avoids internal bias by classifying the KSL based on isolates from outside the library, or isolates from the library that are held out and treated as unknowns. Only when a KSL can be cross-validated with a challenge set will it be acceptable for classifying environmental isolates.

C. ARA Remains in the Primary Literature

2003 was a banner year for ARA (Table 1). The number of publications increased somewhat steadily until 2003, and ARA continues to surface in the primary literature. Moreover, the diversity of ARA-related projects with respect to scope and application has expanded. One possible niche for ARA is with relatively smaller watersheds. For example, Graves et al. (26) used ARA to show that cattle were the major source of pollution in a rural Virginia watershed. Although their KSL was comparatively small (562 unique isolates), 65% of the challenge set isolates were correctly classified. Webster et al. (70) used cluster analysis of *E. coli* resistance patterns to show that two different watersheds were contaminated with effluent from various wastewater treatment plants. Although specific sources of pollution were not identified, there was a positive correlation between a higher degree of urbanization of a watershed and higher fecal coliform loads. Greater urbanization can result in higher fecal coliform loads due to higher

volumes of wastewater discharged, a higher percentage of impervious surfaces, and a potentially higher percentage of resistant bacteria. Earlier studies produced similar results with regard to the difference between urban and rural watersheds (42, 50).

The latest efforts have delved deeper into the structure of fecal coliforms and *E. coli* populations among different host species. Shehane et al. (58) investigated the impact of varying rainfall conditions on microbial indicators (*Clostridium perfringens*, enterococci, coliphage, and fecal coliforms) in an urban watershed. Specifically, ARA was performed using resistance patterns of fecal coliforms from various known sources, including humans, chickens, cattle, pigs, dogs, and wild animals (rabbits, raccoons, birds). When compared to the other microbial indicators, fecal coliforms did not vary significantly over time. A separate study, however, concluded temporal stability may be a questionable characteristic of an ARA project, suggesting that a project should be designed in order that known source fecal samples are collected simultaneously with environmental water column samples (2). Further investigation is required to determine if this phenomenon readily occurs in more than the three host-species (human, cow, horse) studied. Orosz-Coghlan et al. investigated *E. coli* blooms in an Arizona wetland to assess the impact of various avian species (49). A KSL consisting of several avian species, including both passerine and waterfowl, was used to classify environmental samples from four locations. Overall, the passerine species were determined to be the major source for the majority of the locations.

Table 1. Number of ARA Publications, by Year[a]

Year	N	Use of ARA[b,c,d,e]
1970s	3	MD (5, 12, 43)
1980s	2	MD (4, 45) MD & FA (45)
1990	1	MD & FA (42)
1996	1	MD & FA (72)
1997	1	MD & FA (50)
1999	3	MD & FA (35, 46, 73)
2000	1	MD & FA (38)
2001	2	MD (8), MD & FA (3)
2002	6	MD (29), MD & FA (13, 27, 71), RP (56, 59)
2003	14	MD (21), MD & FA (6, 7, 10, 24, 25, 44, 51, 74), MC (20, 28, 39, 52, 60)
2004	4	MC (62), MD (16), MD & FA (70), RP (47)
2005	7	MC (11), MD & FA (9, 18, 23, 55, 58), RP (57)
2006	4	MD & FA (2, 19, 49, 68)
2007	5	MD & FA (15, 26, 36, 41), RP (61)

[a] As of 8/01/07
[b] FA = Field application study
[c] MC = Method comparison study
[d] MD = Method development study
[e] RP = Review paper

The most recent projects are also learning to apply ARA in a manner agreeable with its inherent limitations. Kaneene et al., for example, used antibiotic resistance patterns of *E. coli* to classify environmental isolates (41). Their KSL was sufficiently large relative to the size of the watershed, and it included isolates from all the major potential sources of pollution. Similar to other projects, it was found that reducing the number of species classifications and/or antibiotics can increase the ARCC.

D. Performance-based Criteria for MST Projects

Recent literature (17, 61, 65) has proposed a comprehensive list of performance criteria that every MST project should fulfill, including such characteristics as **experimental design, minimum detectable percentage, quantification, repeatability, accuracy, specificity, robustness, range of applicability, and practicality**. These criteria should be considered before applying any MST project in the field.

Experimental Design. Experimental design deals largely with the construction of the KSL, and includes such characterisitics as composition, size, continuity, and sensitivity (17, 61). Proper library composition involves selecting an appropriate indicator bacteria (*E. coli* for freshwater or enterococci for saltwater) and then collecting fecal samples from host species in order to isolate the chosen indicator. The host species should consist of all the animals in the watershed that have realistic potential to impact water quality. It is generally not necessary to classify environmental isolates to the host species level. Instead, reproducible results have been obtained by categorizing host species into broader categories, such as wildlife or birds (35, 38, 74). The next characteristic, library size, also concerns the representativeness of the KSL. Although no standard size is currently agreed upon, it is recommended that a library contain at

least 1,000 isolates per host species (17). After establishing the size, the representativeness can be estimated by comparing the ARCC from a resubstitution analysis with the ARCC from a cross-validation analysis. Library continuity involves updating the library at least once a year, in order to reduce the impact of such factors as temporal stability and geographic stability. This also means that the library should be used only in the watershed for which it was designed, unless other evidence shows it can be used elsewhere (61). Finally, library sensitivity defined as the proportion of samples that are positive. Also known as the rate of correct classification, sensitivity should be reported for each host species and an average of all individual rates of correct classification should also be calculated (17).

Other important considerations with experimental design include representative sampling and using the "toolbox" or "tiered" approach. Representative sampling requires a thorough understanding of the project watershed and the temporal and spatial variability of the indicator bacterium. For example, *E. coli* concentrations have been shown to be up to three times higher in the morning than later in the day (61). The toolbox approach refers to the use of multiple methods to identify the sources of fecal contamination. Although this may increase upfront costs, it may help to prevent more expensive infrastructure changes from being improperly mandated. Fluorometric analysis combined with ARA is an example of the toolbox approach. Similar to the toolbox approach, the tiered approach allows the researcher to use one method to identify the problem areas, and then a different technique to identify the source of the problem.

Minimum Detectable Percentage. Every library-dependent MST project should include a calculation of the Minimum Detectable Percentage (MDP), a measure of the lower limit for considering that a source is present in a sample. The MDP is an estimation of the likelihood that an isolate that is not from a given source will be classified into that source. It is also known as

the sensitivity of a test (65). Knowing the MDP gives the researcher a significance cut-off when it comes time to classify the sources of isolates in environmental samples (17). One way to calculate the MDP is to use the observed frequency of isolate misclassification (71). Wiggins et al. (74) used an MDP of 25% for a large (6,587) enterococci library, while the SCCWRP comparison study applied an MDP of 15% (39).

Quantification. The ultimate objective of any MST method is to be able to quantitatively assess the amount of fecal contamination in an environmental sample. The purpose of developing a KSL is to be able to apply it to the sources present in an environmental sample. In order to assure proper quantification, the library must first be cross-validated using challenge isolates. Performing a challenge set reduces the error associated with classification of environmental isolates. One of the original methods used to perform a challenge test was discriminant analysis. Accuracy of the KSL was evaluated by classifying each library isolate into a source category based on its similarity to the other library isolates, while remaining in the library. The accuracy value was termed the rate of correct classification (RCC). Recent projects have demonstrated, however, that this method does not perform as well for isolates outside the library (39, 52). A newer method is a form of jackknife analysis known as the pulled-sample test, in which all the isolates from a given sample are held out of the library and treated as unknowns. Recent projects have successfully incorporated many of the performance criteria. For example, Dickerson et al. (15) used pulsed-field gel electrophoresis as part of a toolbox approach, validated their KSL using a challenge set, and calculated an MDP for each year that known source samples were collected. This was also the first project to present results based on the classification method known as logistic regression. Graves et al. (26) used both jackknife

analysis and the artificial clustering procedure to validate the KSL of a relatively small, rural watershed.

Repeatability. Also known as precision, repeatability concerns whether the results will be the same if a test is conducted under the same or very similar conditions. Repeatability is different from accuracy (see below). Repeatability is a measure of how many times a person hits the same place on a dartboard, not how many times the bull's eye is hit. Repeatability is generally expressed by the calculation of standard deviation, but can also be shown using relative standard deviation (65).

Accuracy. Accuracy is a measure of the degree to which a method identifies its target. It is the degree of agreement between an observed value and an accepted reference value. This criterion includes random error and systematic error that are natural consequences of sampling and analysis (65).

Specificity. Specificity is the rate of false positives and false negatives for a given test. A false positive asks whether the method is significantly more or less likely to detect non-target organisms that would be reported as the target organism. The determination that the samples do not contain the target organism (or other such parameter) should be based on a second, independent standard method. For example, if antibiotic resistance analysis suggests a human signature, then fluorometric analysis could be used to confirm or deny. A false negative asks whether a method is more or less likely to not detect target organism when the target organism is indeed present. As with false positives, false negatives should be confirmed using a separate, independent standard method (65).

Robustness. Robustness is a term to describe the degree to which a method can perform in the presence of incorrect inputs or stressed conditions. In other words, how poorly can a

method perform and still produce useful results? For example, if a test is for cultured microorganisms, can it detect stressed organisms in ambient waters? It is difficult to quantify robustness but must be kept in mind when considering method development and application (65).

Range of Applicability. Range of applicability refers to how reliable a method is on a nationwide basis. For example, how well does ARA perform in subtropical vs. temperate climates, with the Great Lakes vs. marine waters? It does not, however, apply to parameters other than the one(s) for which the method was designed; in other words, a recreational water quality standard should not be applied to sewage sludge. Similar to robustness, this criterion is hard to calculate but should nevertheless be considered with overall method performance.

Practicality. Practicality focuses on four main issues: capital cost, training cost, per sample cost, and additional sampling requirements. Capital costs are the upfront costs such as equipment purchase and space required to perform the method. Training costs are the expenses incurred prior to routine testing so that the user can perform the method within the performance criteria. Examples include workshops or training modules. High per sample costs and additional sampling requirements can become prohibitive if large volumes of tests have to be performed on a routine basis (65).

V. Summary of Study Design

A. Antibiotic Resistance Analysis (ARA)

The argument has been laid out for the use of ARA as an effective phenotypic, species-specific, MST technique. Many bacterial indicators have been attempted, but few have performed as consistently well as *E. coli* and the enterococci. The method of ARA relies on different antibiotic resistance patterns in fecal bacteria that can be related to specific sources of

fecal pollution, and is predicated on the rationale that antibiotics exert selective pressure on the fecal flora of the animals that ingest or are treated with the antibiotic(s), and that different types of animals receive differential exposure to antibiotics. Benefits of ARA include the use of simple laboratory techniques that require only basic equipment, and can be performed at a relatively low cost compared to most other MST methods.

Two methods of ARA are used in research. In the first method, fecal bacteria are isolated from fecal samples and challenged with antibiotics and scored for growth or no growth. This provides a library of resistance patterns. A second, less widely used method, uses antibiotic zones of inhibition (55) rather than growth or no growth. Environmental isolates are then challenged with the same antibiotics and compared to the library set. The environmental isolates are then categorized through some form of discriminant analysis. Unlike molecular techniques, this high throughput, low cost method allows 10-fold or more increase in the number of isolates tested.

The antibiotic resistance variation between isolates of different sources tends to follow certain trends. Humans are typically found to have the highest rates of resistance, followed by livestock and pets. Wildlife generally have the lowest resistance. There are exceptions to this (55), as pets may share microbes with humans. Wildlife such as gulls feed in human sewage and may share fecal bacteria. Areas such as farms and hospitals, where antibiotic use is more common, tend to spread antibiotic resistance into the neighboring fauna in a manner similar to a chemostat. Lateral transfer of resistance genes also creates a source of variation in the data (55).

All these sources of variation require thorough testing to assure that the library is representative of the project watershed. The known source library set must be sufficiently large,

1000 or more non-unique isolates depending on watershed size, to represent the local variation (61, 74). Due to the evolving variation, the known source library is constrained by time. Current data suggests that libraries are good for at least one year, and must therefore be updated in order to remain stable (74). Several additional criteria must be fulfilled in order to assure a representative known source library: performance of a challenge set, calculation of the minimum detectable percentage, and verification using a separate MST technique. In summary, the reasons for the use of ARA as a primary source tracking technique are as follows:

- Ease of method
- Well established in primary literature
- Low cost per isolate
- Results can be cross-validated using multiple MST methods
- Established regional library in Prince William County

B. Fluorometric Analysis

One of the project objectives is to use fluorometric analysis (FA), also known simply as "fluorometry," to complement the ability of ARA to detect human sources of pollution. Fluorometric analysis is an emerging source tracking technique that serves as a rapid presence/absence test for human fecal contamination (24). The technique is designed to detect optical brighteners (OBs), also known as fluorescent brighteners or fluorescent whiteners. OBs are synthetic compounds found ubiquitously in laundry detergent, soaps, and household cleaning products. At least two major sources of human contamination could contain optical brighteners: onsite wastewater systems and community wastewater treatment systems (54). In terms of cost-effectiveness, FA is relatively inexpensive compared to other chemical source tracking methods that typically require a mass spectrometer.

Fluorometry is gaining increasing notoriety in the source tracking community, especially as a secondary means of detection and confirmation (64, 69). For example, Dickerson et al. used ARA in conjunction with FA to inform public beach officials of sewage infrastructure problems. The problems were addressed and follow-up results confirmed that the issue was resolved (15). Hartel et al. used fluorometry to correctly identify two negative and three positive locations for human fecal contamination (37). The project in PWC has been using FA since 2003 (63) and it continues to show promise as a consistent, confirmatory measure of human fecal contamination.

VI. References

1. **Amábile-Cuevas, C. F.** 2003. New antibiotics and new resistance. Am. Sci. **91:**138-149.
2. **Anderson, M. A., J. E. Whitlock, and V. J. Harwood.** 2006. Diversity and distribution of *Esherichia coli* genotypes and antibiotic resistance phenotypes in feces of humans, cattle, and horses. Appl. Environ. Microbiol. **72:**6914-6922.
3. **Arvanitidou, M., V. Katsouyannopoulos, and A. Tsakris.** 2001. Antibiotic resistance patterns of enterococci isolated from coastal bathing waters. J. Med. Microbio. **50:**1001-1005.
4. **Bell, J. B., G. E. Elliott, and D. W. Smith.** 1983. Influence of sewage treatment and urbanization on selection of multiple resistance in fecal coliform populations. Appl. Environ. Microbiol. **46:**227-232.
5. **Bell, R. B.** 1978. Antibiotic resistance patterns of fecal coliforms isolated from domestic sewage before and after treatment in an aerobic lagoon. Can. J. Microbiol **24:**886-888.
6. **Booth, A. M., C. Hagedorn, A. K. Graves, S. C. Hagedorn, and K. H. Mentz.** 2003. Sources of fecal pollution in Virginia's Blackwater River. J Environ Eng **129:**547-552.
7. **Burnes, B. S.** 2003. Antibiotic resistance analysis of fecal coliforms to determine fecal pollution sources in a mixed-use watershed
   Environ. Monit. Assess. **85:**87-98.
8. **Butaye, P., L. A. Devriese, and F. Haesebrouck.** 2001. Differences in antibiotic resistance patterns of *Enterococcus faecalis* and *Enterococcus faecium* strains isolated from farm and pet animals. Antimicrob Agents Chemother **45:**1374-1378.
9. **Carroll, S., M. Hargreaves, and A. Goonetilleke.** 2005. Sourcing faecal pollution from onsite wastewater treatment systems in surface waters using antibiotic resistance analysis. J. Appl. Microbiol. **99:**471-482.
10. **Choi, S., W. Chu, J. Brown, S. J. Becker, V. J. Harwood, and S. C. Jiang.** 2003. Application of enterococci antibiotic resistance patterns for contamination source identification at Huntington Beach, California. Mar. Pollut. Bull. **46:**748-755.

11. **Cimenti, M., N. Biswas, J. K. Bewtra, and A. Hubberstey.** 2005. Evaluation of microbial indicators for the determination of bacterial groundwater contamination sources. Water Air Soil Pollut **168:**157-169.
12. **Cooke, M. D.** 1976. Antibiotic resistance among coliform and fecal coliform bacteria isolated from sewage, sewater, and marine shellfish. Antimicrob. Agents Chemother, **9**.
13. **Crozier, J. B., B. Clark, and H. Weber.** 2002. Identifying sources of fecal pollution in the Roanoke river, Roanoke County, Virginia. Va J Sci **53:**157-165.
14. **Díaz-Mejía, J. J., A. Carbajal-Saucedo, and C. F. Amábile-Cuevas.** 2002. Antibiotic resistance in oral commensal streptococci from healthy Mexicans and Cubans: resistance prevalence does not mirror antibiotic usage. FEMS Microbiol Lett **217:**173-176.
15. **Dickerson, J. W. J., C. Hagedorn, and A. Hassall.** 2007. Detection and remediation of human-origin pollution at two public beaches in Virginia using multiple source tracking methods. Water Res. **(accepted)**.
16. **Ebdon, J. E., J. L. Wallis, and H. D. Taylor.** 2004. A simplified low-cost approach to antibiotic resistance profiling for faecal source tracking. Water Sci. Tech. **50:**185-191.
17. **Edge, T. A., J. Griffith, J. Hansel, V. J. Harwood, M. Jenkins, A. Layton, M. Molina, C. Nakatsu, R. Oshiro, M. J. Sadowsky, J. W. Santo Domingo, O. C. Shanks, G. Stelma, J. R. Stewart, D. M. Stoeckel, B. A. Wiggins, and J. D. Wilbur.** 2005. Microbial source tracking guide document, p. 1-151. *In* USEPA (ed.).
18. **Edge, T. A., and S. Hill.** 2005. Occurence of antibiotic resistance in *Escherichia coli* from surface waters and fecal pollution sources near Hamilton, Ontario. Can. J. Microbiol. **51:**501-505.
19. **Evenson, C. J., and K. A. Strevett.** 2006. Discriminant analysis of fecal bacterial species composition for use as a phenotypic microbial source tracking method. Res. Microbiol. **157:**437-444.
20. **Field, K. G., E. C. Chern, L. K. Dick, J. Fuhrman, J. Griffith, P. Holden, M. G. LaMontagne, J. Le, B. Olson, and M. T. Simonitch.** 2003. A comparative study of culture-independent, library-independent genotypic methods of fecal source tracking. J. Wat. Health **1:**181-194.
21. **Fogarty, L. R., S. K. Haack, M. J. Wolcott, and R. L. Whitman.** 2003. Abundance and characteristics of the recreational water quality indicator bacteria *Escherichia coli* and enterococci in gull faeces. J. Appl. Microbiol. **94:**865-878.
22. **Garcia, L.** 1999. Revisiting that pesky acronym, TMDLs, p. 1-6, Virginia Water Central.
23. **Genthner, F. J., J. B. James, and D. F. Yates.** 2005. Use of composite data sets for source-tracking enterococci in the water column and shoreline interstitial waters on Pensacola Beach, Florida. Mar. Pollut. Bull. **50:**724-732.
24. **Gilpin, B., T. James, F. Nourozi, D. Saunders, P. Scholes, and M. Savill.** 2003. The use of chemical and microbial indicators for faecal source identification. Water Sci. Tech. **47:**39-43.
25. **Graves, A. K.** 2003. Identifying sources of fecal pollution in water as a function of sampling frequency under low and high stream flow conditions. Virginia Polytechnic Institute and State University, Blacksburg, VA.
26. **Graves, A. K., C. Hagedorn, A. Brooks, R. L. Hagedorn, and E. Martin.** Microbial source tracking in a ruralwatershed dominated by cattle,Water Research (2007), doi:10.1016/j.watres.2007.04.020.

27. **Graves, A. K., C. Hagedorn, A. Teetor, M. Mahal, A. Booth, and J. R. B. Reneau.** 2002. Antibiotic resistance profiles to determine sources of fecal contamination in a rural Virginia watershed. J. Environ. Qual. **31:**1300-1308.
28. **Griffith, J. F., S. B. Weisberg, and C. D. McGee.** 2003. Evaluation of microbial source tracking using mixed fecal sources in aqueous in aqueous test samples. J. Wat. Health **1:**141-151.
29. **Guan, S., R. Xu, S. Chen, J. Odumeru, and C. Gyles.** 2002. Development of a procedure for discriminating among *Esherichia coli* isolates from animal and human sources. Appl. Environ. Microbiol. **68:**2690-2696.
30. **Hagedorn, C.** 2005. Identifying sources of fecal pollution in impaired waters in Prince William County, Virginia, p. 1-45. Virginia Polytechnic Institute and State University.
31. **Hagedorn, C.** 2005. Identifying sources of fecal pollution in impaired waters in Prince William County, Virginia: sites monitored quarterly in 2004-2005, p. 1-4. Virginia Polytechnic Institute and State University.
32. **Hagedorn, C.** 2006. Identifying sources of fecal pollution in impaired waters in Prince William County, Virginia: sites monitored quarterly in 2005-2006, p. 1-8. Virginia Polytechnic Institute and State University.
33. **Hagedorn, C.** 1999. Tracking the wild-and domestic-bacteria., p. 8-10, Virginia Water Central.
34. **Hagedorn, C.** 2004. Updated final project report to Department of Public Works, Prince William County, VA, p. 1-29. Virginia Polytechnic Institute and State University.
35. **Hagedorn, C., S. L. Robinson, J. R. Filtz, S. B. Grubbs, T. A. Angier, and J. R. B. Reneau.** 1999. Determining sources of fecal pollution in a rural Virginia watershed with antibiotic resistance patterns in fecal streptococci. Appl. Environ. Microbiol. **65:**5522-5531.
36. **Hamelin, K., G. Bruant, A. El-Shaarawi, S. Hill, T. A. Edge, J. Fairbrother, J. Harel, C. Maynard, L. Masson, and R. Brousseau.** 2007. Occurence of virulence and antimicrobial resistance genes in *Escherichia coli* isolates from different aquatic ecosystems within the St. Clair River and Detroit River areas. Appl. Environ. Microbiol. **73:**477-484.
37. **Hartel, P. G., C. Hagedorn, J. L. McDonald, J. A. Fisher, M. A. Saluta, J. W. J. Dickerson, L. C. Gentit, S. L. Smith, N. S. Mantripragada, K. J. Ritter, and C. N. Belcher.** Exposing water samples to ultraviolet light improves fluorometry for detecting human fecal contamination Water Research (2007), doi:10.1016/j.watres.2007.03.034.
38. **Harwood, V. J., J. Whitlock, and V. Withington.** 2000. Classification of antibiotic resistance patterns of indicator bacteria by discriminant analysis: use in predicting the source of fecal contamination in subtropical waters. Appl. Environ. Microbiol. **66:**3698-3704.
39. **Harwood, V. J., B. A. Wiggins, C. Hagedorn, R. D. Ellender, J. Gooch, J. Kern, M. Samadpour, A. C. H. Chapman, B. J. Robinson, and B. C. Thompson.** 2003. Phenotypic library-based microbial source tracking methods: efficacy in the California collaborative study. J. Wat. Health **1:**153-166.
40. **Hoyle, D. V., H. C. Davison, H. I. Knight, C. M. Yates, O. Dobay, G. J. Gunn, S. G. B. Amyes, and M. E. J. Woolhouse.** 2006. Molecular characterisation of bovine faecal *Escherichia coli* shows persistence of defined ampicillin resistant strains and the presence of class I integrons on an organic beef farm. Vet Microbiol **115:**250-257.

41. **Kaneene, J. B., R. Miller, R. Sayah, Y. J. Johnson, D. Gilliland, and J. C. Gardiner.** 2007. Considerations when using discriminant function analysis of antimicrobial resistance profiles to identify sources of fecal contamination of surface waters in Michigan. Appl. Environ. Microbiol. **73:**2878-2890.
42. **Kaspar, C. W., J. L. Burgess, I. T. Knight, and R. R. Colwell.** 1990. Antibiotic resistance indexing of Escherichia coli to identify sources of fecal contamination in water. Can. J. Microbiol. **36:**891-894
43. **Kelch, W. J., and J. S. Lee.** 1978. Antibiotic resistance patterns of Gram-negative bacteria isolated from environmental sources. Appl. Environ. Microbiol. **36:**450-456.
44. **Kelsey, R. H., G. I. Scott, D. E. Porter, B. Thompson, and L. Webster.** 2003. Using multiple antibiotic resistance and land use characteristics to determine sources of fecal coliform bacterial pollution. Environ. Monit. Assess. **81:**337-348.
45. **Krumperman, P. H.** 1983. Multiple antibiotic resistance indexing of *Escherichia coli* to identify high-risk sources of fecal contamination of foods. Appl. Environ. Microbiol. **46:**165-170.
46. **Mathew, A. G., A. M. Saxton, W. G. Upchurch, and S. E. Chattin.** 1999. Muliple antibiotic resistance patterns of Escherichia coli isolates from swine farms. Appl. Environ. Microbiol. **65:**2770-2772.
47. **Meays, C. L., K. Broersma, R. Nordin, and A. Mazumder.** 2004. Source tracking fecal bacteria in water: a critical review of current methods. J. Environ. Manag. **73:**71-79.
48. **Myoda, S. P., C. A. Carson, J. J. Fuhrmann, B. Hahm, P. G. Hartel, H. Yampara-Iquise, L. A. Johnson, R. L. Kuntz, C. H. Nakatsu, M. J. Sadowsky, and M. Samadpour.** 2003. Comparison of genotypic-based microbial source tracking methods requiring a host origin database. J. Wat. Health **1:**167.
49. **Orosz-Coghlan, P. A., P. A. Rusin, M. M. Karpiscak, and C. P. Gerba.** 2006. Microbial source tracking of *Escherichia coli* in a constructed wetland. Water Environ. Res. **78:**227-232.
50. **Parveen, S., R. L. Murphree, L. Edmiston, C. W. Kaspar, K. M. Portier, and M. L. Tamplin.** 1997. Association of multple-antibiotic-resistance profiles with point and nonpoint sources of *Escherichia coli* in Apalachicola Bay. Appl. Environ. Microbiol. **63:**2607-2612.
51. **Porter, K. R.** 2003. Identifying sources of fecal pollution in Washington, DC waterways. M.S. Thesis. Virginia Polytechnic Institute and State University, Blacksburg, VA.
52. **Ritter, K. J., E. Carruthers, C. A. Carson, R. D. Ellender, V. J. Harwood, K. Kingsley, C. Nakatsu, M. J. Sadowsky, B. Shear, B. West, J. E. Whitlock, B. A. Wiggins, and J. D. Wilbur.** 2003. Assessment of statistical methods used in library-based approaches to microbial source tracking. J. Wat. Health **1:**209-223.
53. **Salyers, A. A., and C. F. Amábile-Cuevas.** 1997. Why are antibiotic resistance genes so resistant to elimination. Antimicrob Agents Chemother **41:**2321-2325.
54. **Sargent, D., and W. Castonguay.** 1998. An optical brightener handbook. Eight Towns and the Bay.
55. **Sayah, R. S., J. B. Kaneene, Y. Johnson, and R. Miller.** 2005. Patterns of antimicrobial resistance observed in *Escherichia coli* isolates obtained from domestic- and wild-animal fecal samples, human septage, and surface water. Appl. Environ. Microbiol. **71:**1394-1404.

56. **Scott, T. M., J. B. Rose, T. M. Jenkins, S. R. Farrah, and L. J.** 2002. Microbial source tracking: current methodology and future directions. Appl. Environ. Microbiol. **68:**5796-5803.
57. **Seurinck, S., W. Verstraete, and S. D. Siciliano.** 2005. Microbial source tracking for identification of fecal pollution. Rev. Environ. Sci. Biotechnol. **4:**19-37.
58. **Shehane, S. D., V. J. Harwood, J. E. Whitlock, and J. B. Rose.** 2005. The influence of rainfall on the incidence of microbial faecal indicators and the dominant sources of faecal pollution in a Florida river. J. Appl. Microbiol. **98:**1127-1136.
59. **Simpson, J. M., J. W. Santo Domingo, and D. J. Reasoner.** 2002. Microbial source tracking: state of the science. Environ. Sci. Tech. **36:**5279-5288.
60. **Stewart, J. R., R. D. Ellender, J. A. Gooch, S. C. Jiang, S. P. Myoda, and S. B. Weisberg.** 2003. Recommendations for microbial source tracking: lessons from a methods comparison study. J. Wat. Health **1:**225-231.
61. **Stoeckel, D. M., and V. J. Harwood.** 2007. Performance, design, and analysis in microbial source tracking studies. Appl. Environ. Microbiol. **73:**2405-2415.
62. **Stoeckel, D. M., M. V. Mathes, K. E. Hyer, C. Hagedorn, H. Kator, J. Lukasik, T. L. O'Brien, T. W. Fenger, M. Samadpour, K. M. Strickler, and B. A. Wiggins.** 2004. Comparison of seven protocols to identify fecal contamination sources using *Escherichia coli*. Environ. Sci. Technol. **38:**6109-6117.
63. **Touchton, G. D.** 2005. Microbial source tracking in a mixed use watershed in Northern Virginia. M.S. Thesis. Virginia Polytechnic Institute and State University, Blacksburg, VA.
64. **USEPA.** 2000. Optical brightener to shed light on sewer and septic tank leaks, p. 11-12, Nonpoint source news-notes, vol. 63.
65. **USEPA.** 2007. Report of the experts scientific workshop on critical needs for the development of new or revised recreational water quality criteria, p. 1-203. *In* USEPA (ed.).
66. **USEPA.** 2003. Wastewater technology fact sheet. *In* USEPA (ed.).
67. **VADEQ** June 22, 2007, posting date. TMDLs in Virginia. [Online.]
68. **Vantarakis, A., D. Venieri, G. Komninou, and Papapetropoulou.** 2006. Differentiation of faecal *Escherichia coli* from humans and animals by multiple antibiotic resistance analysis. Lett Appl Microbiol **42:**71-77.
69. **Waye, D.** 2000. A new tool for tracing human sewage in waterbodies: optical brightener modeling, Virginia Water Resources Research Symposium.
70. **Webster, L. F., B. C. Thompson, M. H. Fulton, D. E. Chestnut, R. F. Van Dolah, A. K. Leight, and G. I. Scott.** 2004. Identification of sources of *Escherichia coli* in South Carolina estuaries using antibiotic resistance analysis. J. Exp. Mar. Biol. Ecol. **298:**179-195.
71. **Whitlock, J. E., D. T. Jones, and V. J. Harwood.** 2002. Identification of the sources of fecal coliforms in an urban watershed using antibiotic resistance analysis. Water Res. **36:**4273-4282.
72. **Wiggins, B. A.** 1996. Discriminant analysis of antibiotic resistance patterns in fecal streptococci, a method to differentiate human and animal sources of fecal pollution in natural waters. Appl. Environ. Microbiol. **62:**3997-4002.
73. **Wiggins, B. A., R. W. Andrews, R. A. Conway, C. L. Corr, E. J. Dobratz, D. P. Dougherty, J. R. Eppard, S. R. Knupp, M. C. Limjoco, J. M. Mettenberg, J. M.**

Rinehardt, J. Sonsino, R. L. Torrijos, and M. E. Zimmerman. 1999. Use of antibiotic resistance analysis to identify nonpoint sources of fecal pollution. Appl. Environ. Microbiol. **65:**3483-3486.

74. **Wiggins, B. A., P. W. Cash, W. S. Creamer, S. E. Dart, P. P. Garcia, T. M. Gerecke, J. Han, B. L. Henry, K. B. Hoover, E. L. Johnson, K. C. Jones, J. G. McCarthy, J. A. McDonough, S. A. Mercer, M. J. Noto, H. Park, M. S. Phillips, S. M. Purner, B. Smith, E. N. Stevens, and A. K. Varner.** 2003. Use of antibiotic resistance analysis for representativeness testing of multiwatershed libraries. Appl. Environ. Microbiol. **69:**3399-3405.

Chapter 2. Goals and Objectives

The overall project goal was to monitor and evaluate twenty one stream locations in the Occoquan Basin identified as impaired due to high *E. coli* densities. One site on each of eight streams, two sites on each of five streams, and three sites on the remaining stream were chosen for *E. coli* monitoring and microbial source tracking.

The project objectives are:

1. The categories of fecal sources that lead to bacterial impairment will be determined using antibiotic resistance analysis (ARA).
2. The known source library (KSL) to be used with ARA will be designed in a manner that best represents the major fecal sources (human, pets, livestock, wildlife) in the watershed in order to best identify the sources present in the environmental isolates. This will include performing two challenge sets against the KSL. This will serve as an assessment of the ability of the KSL to classify known-source isolates not already included in the library.
3. Fluorometric analysis (FA) of optical brighteners will be utilized as an indicator of human wastewater. Data will be collected during the same collection period as for ARA. The FA data will be compared with the ARA results in order to assess the capability of FA to serve as a secondary indicator of human contamination.

Chapter 3. Materials and Methods

I. Defining the Study Location

The Occoquan Basin (OcB) of the Middle Potomac-Anacostia-Occoquan watershed within Prince William County (PWC) (Appendix I) served as the region for the project. Encompassing an area of approximately 1528 km^2, the OcB serves as a headwater for a Potomac River tributary, which eventually discharges into the Chesapeake Bay. Over the past several decades the OcB has transitioned from a traditionally rural farming region to one of the fastest growing regions in America. This change has affected the relationship between humans and animals by increasing the proportion of shared living spaces. Fewer open spaces forces animals into parks, refuges, and neighborhoods, causing over-crowding and increased competition for resources. An additional consequence is an increase in the concentration of fecal waste deposited near waterways. An increased concentration of animal waste can become a serious issue for communities as the waste is carried into waters designated for recreational uses.

The vast majority of animal waste for this project was collected from recreational and state parks located throughout PWC, such as Leesylvannia State Park. The animal sources included deer (*Odocoileus virginianus*), Canada geese (*Branta Canadensis* and *Anser domesticus*), various gull species (*Larus sp.*), horses (*Equus caballus*), dogs (*Canis familiaris*), and cows (*Bos taurus* and *Bos indicus*). Human (*Homo sapiens)* samples were collected from the H.L. Mooney Water Reclamation Facility in Woodbridge, VA.

II. Environmental Water Sample Locations

Environmental water samples were collected from 18 quarterly (once every three months) sites and three monthly sites. With the exception of sites Q4 (Cow Branch at Montgomery Avenue) and Q18 (North Fork of Lake Manassas), the 18 quarterly sites were a continuation of earlier efforts. The three monthly sites were first-time locations. Sampling of the quarterly sites commenced in June 2006 and happened once every three months until the final sampling in June 2007. Sampling of the monthly sites began in July 2006 and took place once every month (simultaneously with quarterly sampling when applicable) until the final collection in June 2007. Additionally, sediment samples were collected along with environmental samples during the June 2007 collection period for purposes of comparison and quality control. In general, the sites were surrounded by wooded buffer zones and dense undergrowth in the summer months. Individual site descriptions are provided below.

Quarterly Site 1 (Q1): Neabsco Creek, Lindendale Road - N38°38.7273' W077°21.9542'

Although located in a commercial district of the county, this site is surrounded by a green buffer zone. The sampling occurred where the creek runs under a two lane bridge, over which moderate traffic patterns were commonly observed. People were observed walking their dogs in the immediate vicinity, and evidence of dog scat was apparent on numerous occasions. No immediate or nearby construction or development was observed over the course of the project, but the water was quite muddy on several occasions, most likely the result of a recent rain event and/or upstream construction.

Site Q1: Recent rain events resulted in muddy water on multiple occasions.

Quarterly Site 2 (Q2): Neabsco Creek, Benita Fitzgerald Road – N38°37.5141' W077°18.8082'

This location is typical of the many highly developed residential areas of PWC. An apartment complex has been constructed within the last couple years at this site, but a green buffer zone has helped to reduce the impact of the suburban sprawl. Sampling took place where the creek runs under a four lane bridge, and at least one large storm drain was observed leading directly from the road into the creek. No immediate or nearby construction or new development was observed.

Site Q2: The storm drain leads directly to the creek.

Quarterly Site 3 (Q3): Neabsco Creek, Neabsco Mills Road & Route 1 - N38°36.6421' W077°17.4307'

Sampling was conducted where the creek goes under a heavily-trafficked four lane highway. During the 2007 summer months, road construction was observed in the immediate vicinity. This site is classified as a non-residential, commercial area of the county.

Site Q3: Litter was a recurring issue at this Neabsco Creek location.

Site Q3: The four lane highway was typically marked by heavy traffic.

Quarterly Site 4 (Q4): Cow Branch, Montgomery Avenue - N38°38.1860' W077°16.6908'

Sampling in this relatively quiet residential area took place where the creek runs under a two lane bridge. On several occasions the water possessed a reddish tint. This was a result of a heavy iron presence and is characteristic of urban and suburban streams, due to runoff from roads, sidewalks, parking lots, rooftops, and other impervious surfaces. A golf course is located nearby. No immediate or nearby construction was observed.

Site Q4: A green buffer zone surrounds this location.

Site Q4: The nearby golf course.

Quarterly Site 5 (Q5): Cow Branch, Rippon Landing Park - N38°37.0715' W077°16.4505'

Environmental samples were collected at a four lane divided-highway bridge, marked by heavy traffic. During the summer months, multiple construction jobs were observed. Construction was being performed at the bridge to install concrete pipes, and development of a large apartment complex was occurring nearby. The water had a reddish tint during multiple collections.

Site Q5: The water possessed a reddish tint on several occasions.

Quarterly Site 6 (Q6): Powell's Creek, Fox Mills Apt. & Route 1 – N38°35.7780' W077°18.1155'

Samples were acquired behind an apartment complex, in a quiet wooded area. Although no immediate construction was observed, nearby zoning for a new townhouse complex was noted during the summer months.

Site Q6: A wooded buffer zone characterizes this location.

Quarterly Site 7 (Q7): Quantico Creek, South Fork, Joplin Road – N38°35.2507' W077°25.7305'

This location is in one of the more remote parts of the county, if only because it is surrounded by the Quantico Marine Base. The base makes this a heavily wooded area with absolutely no construction or development. Deer and other wildlife are frequent occurrences.

Site Q7: The pole sampler was used whenever access to the stream bank was possible.

Quarterly Site 8 (Q8): Quantico Creek, Main Stem, Mine Road & I-95 Overpass - N38°34.1153' W077°20.1667'

This heavily wooded site is also within the Quantico Marine base. It is not uncommon to spot various types of wildlife, their tracks, and/or their scat. The sampling location is at a single lane bridge situated adjacent to a major four lane divided bridge. No immediate construction or development was observed.

Site Q8: The single lane bridge is adjacent to a larger four lane bridge.

Quarterly Site 9 (Q9): Cedar Run, Carraige Ford Road (Heim property) - N38°38.5709' W077°35.1393'

The environmental samples were collected adjacent to a large grassy field that is part of private farm property. Many additional farms dot the immediate surroundings. On more than one occasion deer were spotted at the creek. Additionally, several dogs were observed off leash. Construction and development are unlikely to occur at this site.

Site Q9: One of the few remaining rural locations in the county.

Quarterly Site 10 (Q10): Cedar Run, Bristow Road – N38°41.2152' W077°29.4485'

This site could be described as a quiet, residential neighborhood. Sampling took place at a two lane bridge, with no nearby or immediate construction or development.

Site Q10: The two lane bridge featured generally low traffic.

Quarterly Site 11 (Q11): Slate Run, Old Church Road - N38°40.5688' W077°30.5324'

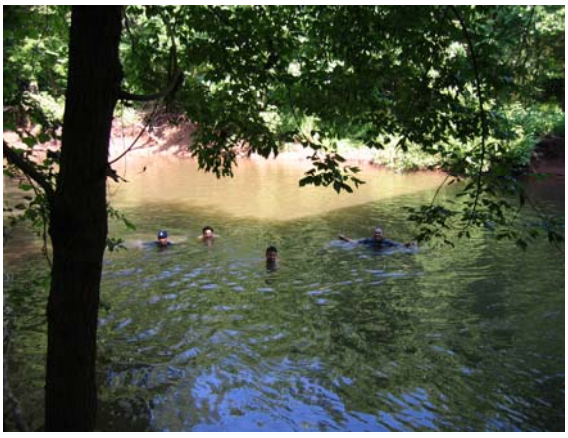
This relatively rural location is marked by dense forest and heavy undergrowth. Deer were spotted on more than one occasion. No construction or development was observed.

Site Q11: Warmer months enhanced the green buffer zone.

Quarterly Site 12 (Q12): Bull Run, Route 28 - N38°48.1749' W077°26.9728'

Sampling was conducted at a four lane divided-highway bridge with moderate traffic. The surroundings could be characterized as a wooded buffer zone that is attractive to various types of wildlife. The wooded zone separates older houses from the stream. Samples were collected at Centreville Rd. (Rt 28), on the Fairfax side of the stream. Although nearby feeder streams drain from housing areas, their flow was relatively minor. Obvious evidence of ATV use was apparent just upstream on the Prince William side, and walking trails were adjacent to the stream on the Fairfax side, within Bull Run Regional Park. This section of the stream has become popular as a swimming location in the summer. No construction or development was observed.

Site Q12: Blackburn's Ford on Upper Bull Run.



Site Q12: Swimmers at Upper Bull Run (June 05), just downstream from Blackburn's Ford.

Quarterly Site 13 (Q13): Catharpin Run, Robin Drive - N38°50.6622' W077°32.8825'

Samples were collected at a two lane bridge marked by minimal to moderate traffic. A wooded buffer zone, frequented by wildlife, surrounds this location. Large populations of deer were observed in the area, and a substantial agricultural area was located upstream. No construction or development was observed.

Site Q13: Deer on the stream bank were a regular occurrence.

Quarterly Site 14 (Q14): Flat Branch, Lomond Drive - N38°46.9103' W077°29.2204'

This heavily-wooded, rocky area is located at a two lane bridge marked by moderate traffic, between older housing developments in Manassas Park. Samples were collected at the Lomond Drive Bridge. Signs of wildlife were noted, litter was a recurring problem, unleashed dogs were observed, and children were seen playing in the stream during warmer weather. No construction or development was observed.

Site Q14: One of the more rocky sampling locations.

Quarterly Site 15 (Q15): South Run, Buckland Mill Road - N38°46.1814' W077°39.9624'

The environmental samples were collected off a single lane gravel road, in a large wooded buffer zone, with a substantial deer population. This area also featured some small farms, a few houses on rural lots, and several large horse farms in the immediate vicinity. The sampling site was on the stream as it enters Lake Manassas, resulting in very little active flow. No construction or development was observed.

Site Q15: The stream widens considerably before entering Lake Manassas.

Site Q15: The South Run sampling site from the bridge on Buckland Mill Road. This view is looking upstream, showing the upper reach of Lake Manassas. South Run emerges from the woods in the upper center of the picture.

Quarterly Site 16 (Q16): Broad Run, Route 28 - N38°44.1842' W077°32.0215'

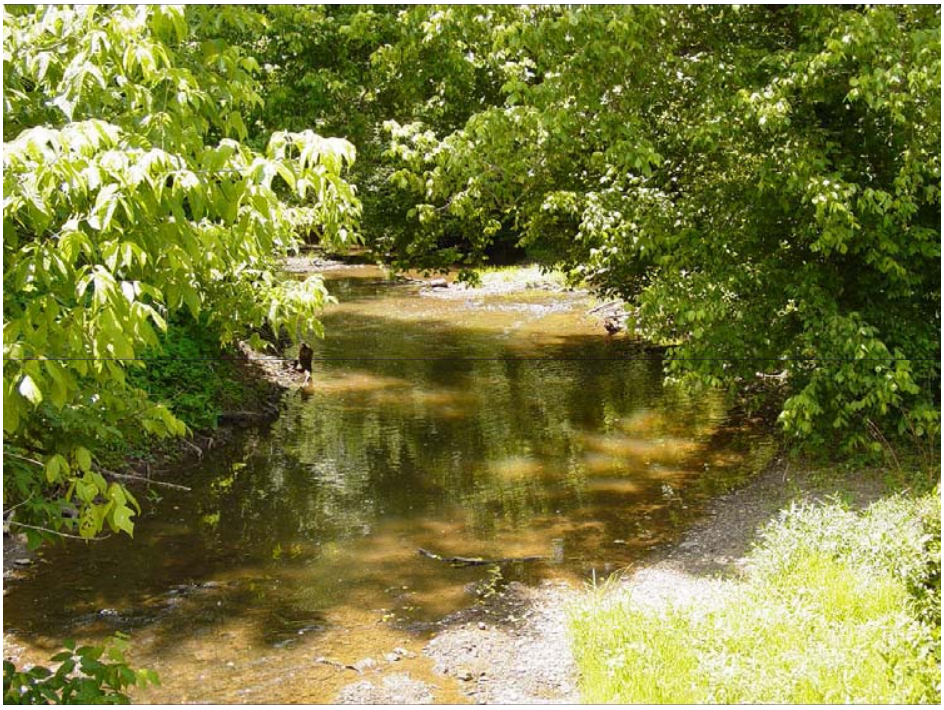
This location could be defined as a commercial area marked by heavy traffic. Sampling occurred at a two lane bridge where heavy bridge construction was ongoing during the summer months. There were no operational farms in the area and no livestock in the fields. People were observed on several occasions walking or exercising dogs. Several housing developments are under construction, in addition to the new developments recently completed.

Site Q16: Generally heavy traffic defined this commercial location.

Site Q16: An example of the ongoing bridge construction.

Quarterly Site 17 (Q17): Kettle Run, Valley View Road – N38°42.1771' W077°32.0032'

This site is situated in a rural setting where a country club and community park (Valley View Park) reside nearby. Samples were collected at a single lane bridge. There were still active hobby farms in the area, but none were adjacent to the sampling location. People were observed using Valley View Park to walk and exercise dogs. Water was backed up below the bridge due to debris collecting in the drains under the bridge on several occasions, but the flow was otherwise moderately strong. No immediate or nearby construction was observed.



Site Q17: A relatively rural setting.

Quarterly Site 18 (Q18): North Fork of Lake Manassas, Route 29 S - N38°47.5768' W077°37.5020'

This commercial area is marked by heavy traffic across a four lane divided highway. Although no human construction was observed, a beaver dam has been in place since at least March 2007. The dam is surrounded by a fairly dense wooded buffer zone, in which deer were spotted on more than one occasion.

Site Q18: At the left edge is the beginning of the large beaver dam.

Site Q18: A popular habitat for deer and waterfowl.

Monthly Site 1 (M1): Little Bull Run, Old Carolina Road - N38°49.2758' W077°37.8274'

A heavily forested, quiet residential neighborhood with a nearby golf course surrounds this location. Samples were obtained at a two lane bridge that is not sufficient for the local traffic and development that is occurring.

Site M1: The bridge may need to be expanded as development continues.

Monthly Site 2 (M2): Powell's Creek, Northgate Drive - N38°36.4787' W077°19.9019'

This location is nestled inside a town park, with a municipal golf course nearby. No construction or development was observed and is unlikely to occur in the immediate vicinity.

Site M2: View from the bridge.

Monthly Site 3 (M3): Broad Run, Route 55 – N38°49.3822' W077°42.3242'

This is a heavily wooded area where a busy two lane highway crosses the creek. No construction or development was observed.

Site M3: View looking upstream.

III. Sample Processing

A. Environmental Samples

For this project an environmental water sample, which may also be referred to as a water column sample, was defined as any sample collected from a body of water containing an aggregation of fecal bacteria for which the source was unknown. The fecal source possibilities were human, livestock, pets, or wildlife. All environmental samples were processed in accordance with EPA Method 1603 : *Escherichia coli* (*E. coli*) in water by membrane filtration using modified membrane-thermotolerant *Escherichia coli* agar (Modified mTEC) (12). After collection, the samples were transported on ice to the laboratory and processing was begun within 24 hours. The water temperature was measured using a "blank" water bottle consisting of tap water. This ensured that the sample temperature was below 1°C for every sample collection. Immediately upon arrival to the laboratory, each sample was subjected to an initial screening for *E. coli* via the IDEXX Colilert® test, within six hours of sample collection whenever possible. Samples that fluoresced under long wave ultraviolet radiation were deemed positive for *E. coli* and were subsequently filtered onto mTEC agar through sterile 47mm 0.45µm-pore-size filter paper, at volumes ranging from 1mL to 100mL (12). The filtered volume varied from sample to sample and often had to be adjusted to obtain densities between 20 and 200 CFU/100mL before proceeding to the next step. A minimum of 2 filtrations were completed for each sample (each individual water sample from each site during each month), allowing for more precise enumeration. The mTEC plates were dry-incubated at 35°C ± 0.5°C for 2 ± 0.5 hours and then incubated in a water bath at 44.5°C ± 0.2°C for 22 ± 2 hours (12).

Following incubation the dark, purple *E. coli* colonies were enumerated and the number of colony forming units (CFUs) per 100mL was calculated for each sample. These data were

recorded as the monitoring results, and included the fluorometry readings explained in Section C below. After enumeration the *E. coli* colonies were transferred using sterile toothpicks into a 96 microwell tray containing Colilert® broth. Ideally, 24 colonies were obtained per sample; however, not every sample provided enough colonies even when the entire volume of water was filtered. For these instances, a minimum of 12 colonies was required for the sample to be included for further analysis, in order to be consistent with the procedures used by the Virginia Department of Environmental Quality (DEQ). The microwell trays were dry-incubated at 35°C ± 0.5°C for 24 ± 2 hours (12).

Antibiotic resistance analysis (ARA) begins with the plating of the microwell cultures onto Petri dishes containing 1% Tryptic Soy Agar (TSA) mixed with various concentrations of antibiotics. Each Petri dish contained a unique concentration of one of seven antibiotics. Table 2 lists the name and concentration of the seven antibiotics. The antibiotics were made from fresh stock solutions summarized in Table 3. The type and concentration were selected based on previously published source tracking efforts (7, 14, 15), but were amended during the initial method development stage of this ongoing project to suit PWC and the OcB (6). The cultures from the microwell tray were aseptically transferred to the set of Petri dishes using a 48 prong replica-plater (Sigma, Inc.). The dishes were dry-incubated at 35°C ± 0.5°C for 24 ± 2 hours.

Incubation elicited the resistance patterns of the individual isolates. The isolates were scored for growth or no growth based on visual appearance. Isolates were considered resistant if enough growth occurred to create a solid, distinct, and complete ring around the perimeter of the colony. An isolate expressing anything less was considered susceptible to the particular antibiotic. This was an absolute requirement and allowed for clear distinction between resistant

Table 2. Antibiotic Concentrations

Antibiotic	**Concentrations (µg/L)**
Cepalothin	15, 25, 35
Erythromycin	60, 70, 90, 100
Neomycin	2.5, 5, 10
Oxytetracyline	2.5, 5, 7.5, 10, 15
Rifampicin	60, 75, 90
Streptomycin	2.5, 5, 7.5, 10, 15
Tetracycline	2.5, 5, 7.5, 10, 15

Table 3. Antibiotic Stock Solution Preparations

Antibiotic	Formulation	Solvent	Stock Concentration (mg/mL)
Cepalothin	Cepalothin	1:1 water:methanol	10
Erythromycin	Erythromycin	1:1 water:ethanol	10
Neomycin	Neomycin Sulfate	Distilled water	10
Oxytetracycline	Oxytetracycline HCL	1:1 water:methanol	10
Rifampicin	Rifampicin	Methanol	2.5
Streptomycin	Streptomycin Sulfate	Distilled water	10
Tetracyline	Tetracycline HCL	Methanol	10

and non-resistant isolates. The data were entered first into Excel 2000 (Microsoft Corp., Redmond, WA) and then into SAS-JMP (v. 5.0.1, SAS Inst., Cary, NC) as binary code, with "1" indicating resistance (growth) and "0" indicating susceptibility (no growth). Excel was used as the intermediary data-entry step due to the relative ease associated with entering such a large volume of data.

The decision to use binary code was based on the comparison of four different methods (Table 4): High, Last, Binary, and Combination (11). The High method uses only the highest concentration of antibiotic to which the isolate was resistant. All lower concentrations at which the isolate failed to grow are ignored. The Last method is based on the highest antibiotic concentration before isolates failed to grow. It ignores all higher concentrations. The Binary method records growth data from each concentration as a binary value, making the presence/absence of growth at a particular concentration a unique variable. The fourth method, Combination, brings together the High and Last methods. It uses the highest concentration where growth occurred and the highest concentration before isolates failed to grow. The Combination method ignores all other data. The primary literature supports use of the Binary method for known source library (KSL) design (7, 8, 13-15), and was therefore the method of choice for this project. Further analysis is explained in Section IV.

Table 4. Methods of Known Source Library Design

		Design Interpretation			
Tetracycline Concentration (µg/mL)	**Growth[a,b]**	**High**	**Last**	**Binary**	**Combination**
10	Yes	Ignored	10 µg/mL	Growth	10 µg/mL
15	No	Ignored	No growth	No growth	No growth
30	Yes	Ignored	Ignored	Growth	Ignored
50	No	Ignored	Ignored	No growth	Ignored
100	Yes	100 µg/mL	Ignored	Growth	100 µg/mL
Recorded Value		100 µg/mL	10 µg/mL	1,0,1,0,1	100 µg/mL, 10 µg/mL

[a] Indicated by at least a complete ring of cell growth at the edge of 5 µL inoculation
[b] Growth could be either Yes or No. This table is merely an example of possible results.

B. Source Samples and the Known Source Library

For this project a source sample (human, pets, livestock, wildlife) was defined as any aggregation of fecal bacteria of which the source was known. Among other possibilities, a source sample could be deer scat or human sewage effluent from a wastewater treatment facility. With a few key differences, a procedure similar to the processing of the environmental samples was followed for the handling of the known source samples. Source samples were collected using sterile containers and transported on ice to be processed in the laboratory within 24 hours. Two separate trips were taken to collect source samples, the first on February 1st, 2007 and the second on March 13th, 2007. It is critical to note the date because the KSL must be updated on at least a yearly basis.

A portion of each fecal sample was suspended in approximately 100mL of sterile DDI water. Volumes ranging from 10µl to 1000µl, dependent upon the number of bacteria, were filtered onto MFC agar plates. Following incubation, isolates presumed to be *E. coli* were transferred to a 96 microwell tray containing Colilert® broth, which allowed for appropriate confirmation. For statistical purposes, a minimum of 6 unique isolates had to be selected from a given source sample in order for it to be considered in subsequent analysis. After incubation of the microwell tray, the same ARA steps used for the environmental samples were followed for completion of the source samples. The only major difference was the use of an EMB control plate in addition to the standard TSA control plate. The EMB allowed for final confirmation of *E. coli* for each isolate.

The purpose of collecting source samples was to construct a library of known source isolate resistance patterns, in order to use the library to classify the sources of the environmental samples. The isolates collected for this project were in addition to those collected previously in

PWC (3-6, 11), and were considered the most up-to-date. The library also included a smaller percentage of isolates collected as part of other projects outside PWC but still within the Northern Virginia and Washington, DC region (2, 10). The statistical analysis is explained in Section IV.

C. Fluorometric Analysis

With the ability to detect synthetic compounds such as optical brighteners (OB), fluorometric analysis acts as a rapid presence/absence test for a human signature in an impaired body of water. The method established by Dickerson et al. was used for this project (1). A Turner Designs 10-AU Fluorometer was calibrated with known standards of a commercially available OB, Fluorescent Brightener 28 (FB-28, Sigma Chemical Co.), at concentrations ranging from 0 to 500.0 mg/L (1). The standards and blanks were assembled using sterile DDI water and FB-28. 250.0 mg/L FB-28 was used to set the basic sensitivity to between 25% and 35% of full scale. The fluorometer was blanked using 0.0 mg/L FB-28 in sterile DDI water from the same container used to assemble the standards.

Calibration was performed in order to subtract the value of the blank from all measurements, and the blanking percentage was set between 0% and 5%. The standard solution concentration was set at 125 and calibrated using the 125.0 mg/L FB-28 solution. After calibration, a standard curve was made using concentrations of 0.0, 15.0, 30.0, 60.0, 125.0, 250.0 and 500.0 mg/L FB-28. The fluorometer was re-calibrated as needed, based on whether readings fell outside 10% of the constructed standard values. Additionally, because of the potential for the FB-28 to degrade over time, new standards were made every two weeks and the fluorometer was re-calibrated using the fresh standards and updated standard curve (1).

In order to determine whether a fluorometer reading should be considered high or low, a range of values was established based on data gathered from samples of untreated sewage, ambient water samples containing virtually no fecal pollution, and attempts to match readings with high levels of sewage isolates, as determined by ARA. In total, 24 untreated sewage samples produced values ranging from 164.0 to 558.0 with an average of 259.5. The ambient water samples yielded background levels of fluorescence, ranging from 10.0 to 45.0. After taking into account an expected dilution factor when sewage mixes with a body of water, a concentration value of 100.0 was deemed the minimum threshold for an environmental sample to be recorded as positive for the presence of OBs (1). Fluorometric analysis was utilized over the course of the entire project. A fluorometer reading was recorded for each environmental sample during each collection period for an overall total of 124 readings.

IV. Statistical Analysis

Statistical analysis was conducted using two parametric classification methods, discriminant analysis and logistic regression, in SAS-JMP (v. 5.0.1, SAS Inst., Cary, NC). The classification tables generated by SAS-JMP were used to calculate the rate of correct classification (RCC) for each source of isolates (human, pet, livestock, wildlife), the average rate of correct classification (ARCC), average frequency of misclassification (AFM), and the minimum detectable percentage (MDP) for the KSL.

In addition to the two classification methods, two types of challenge sets were performed: repeatability test and accuracy test. Both challenge sets were conducted using 20 known source isolates (termed challenge isolates) from each source category, selected from the existing KSL. The repeatability test can be defined as a challenge set in which the challenge isolates are left in

the KSL and then reclassified using the same method for environmental isolates. The accuracy test is a challenge set in which the challenge isolates are removed from the KSL and then reclassified (9). For each challenge set, RCC, ARCC, and AFM values were calculated. Using the methods described by Harwood et al. (9) and Dickerson et al. (1), the MDP was calculated by averaging the AFM values from the KSL, the repeatability test, and the accuracy test. The MDP is a cut-off value for the environmental isolates. When classifying the environmental samples, the source categories identified at percentages below the MDP were considered a negligible contributing source (1).

V. References

1. **Dickerson, J. W. J., C. Hagedorn, and A. Hassall.** 2007. Detection and remediation of human-origin pollution at two public beaches in Virginia using multiple source tracking methods. Water Res. **(accepted)**.
2. **Graves, A. K.** 2003. Identifying sources of fecal pollution in water as a function of sampling frequency under low and high stream flow conditions. Virginia Polytechnic Institute and State University, Blacksburg, VA.
3. **Hagedorn, C.** 2005. Identifying sources of fecal pollution in impaired waters in Prince William County, Virginia, p. 1-45. Virginia Polytechnic Institute and State University.
4. **Hagedorn, C.** 2005. Identifying sources of fecal pollution in impaired waters in Prince William County, Virginia: sites monitored quarterly in 2004-2005, p. 1-4. Virginia Polytechnic Institute and State University.
5. **Hagedorn, C.** 2006. Identifying sources of fecal pollution in impaired waters in Prince William County, Virginia: sites monitored quarterly in 2005-2006, p. 1-8. Virginia Polytechnic Institute and State University.
6. **Hagedorn, C.** 2004. Updated final project report to Department of Public Works, Prince William County, VA, p. 1-29. Virginia Polytechnic Institute and State University.
7. **Hagedorn, C., S. L. Robinson, J. R. Filtz, S. B. Grubbs, T. A. Angier, and J. R. B. Reneau.** 1999. Determining sources of fecal pollution in a rural Virginia watershed with antibiotic resistance patterns in fecal streptococci. Appl. Environ. Microbiol. **65:**5522-5531.
8. **Harwood, V. J., J. Whitlock, and V. Withington.** 2000. Classification of antibiotic resistance patterns of indicator bacteria by discriminant analysis: use in predicting the source of fecal contamination in subtropical waters. Appl. Environ. Microbiol. **66:**3698-3704.
9. **Harwood, V. J., B. A. Wiggins, C. Hagedorn, R. D. Ellender, J. Gooch, J. Kern, M. Samadpour, A. C. H. Chapman, B. J. Robinson, and B. C. Thompson.** 2003.

Phenotypic library-based microbial source tracking methods: efficacy in the California collaborative study. J. Wat. Health **1:**153-166.

10. **Porter, K. R.** 2003. Identifying sources of fecal pollution in Washington, DC waterways. M.S. Thesis. Virginia Polytechnic Institute and State University, Blacksburg, VA.
11. **Touchton, G. D.** 2005. Microbial source tracking in a mixed use watershed in Northern Virginia. M.S. Thesis. Virginia Polytechnic Institute and State University, Blacksburg, VA.
12. **USEPA.** 2005. Method 1603: *Escherichia coli* (*E. coli*) in water by membrane filtration using modified membrane-thermotolerant *Escherichia coli* agar (modified mTEC), p. 1-42. *In* USEPA (ed.).
13. **Wiggins, B. A.** 1996. Discriminant analysis of antibiotic resistance patterns in fecal streptococci, a method to differentiate human and animal sources of fecal pollution in natural waters. Appl. Environ. Microbiol. **62:**3997-4002.
14. **Wiggins, B. A., R. W. Andrews, R. A. Conway, C. L. Corr, E. J. Dobratz, D. P. Dougherty, J. R. Eppard, S. R. Knupp, M. C. Limjoco, J. M. Mettenberg, J. M. Rinehardt, J. Sonsino, R. L. Torrijos, and M. E. Zimmerman.** 1999. Use of antibiotic resistance analysis to identify nonpoint sources of fecal pollution. Appl. Environ. Microbiol. **65:**3483-3486.
15. **Wiggins, B. A., P. W. Cash, W. S. Creamer, S. E. Dart, P. P. Garcia, T. M. Gerecke, J. Han, B. L. Henry, K. B. Hoover, E. L. Johnson, K. C. Jones, J. G. McCarthy, J. A. McDonough, S. A. Mercer, M. J. Noto, H. Park, M. S. Phillips, S. M. Purner, B. Smith, E. N. Stevens, and A. K. Varner.** 2003. Use of antibiotic resistance analysis for representativeness testing of multiwatershed libraries. Appl. Environ. Microbiol. **69:**3399-3405.

Chapter 4. Results and Discussion

I. Monitoring Results (Tables 5-6 and Figures 3-4)

Environmental monitoring of the sample locations consisted of enumerating the *E. coli* isolates present in the water and recording the fluorometry values. Assessment of the water quality is based upon the criteria set by the US EPA (11). For this project the Single Sample Maximum Allowable Density of 235 colony forming units (CFU) per 100mL of water was adopted (Table 5). In other words, the US EPA and the Virginia Department of Environmental Quality (VA DEQ) require that *E. coli* levels in a stream must not exceed 235 CFU/100mL in more than 10% of the samples collected over a given year. Additionally, the geometric mean for each site had to remain below 126 CFU/100mL (Table 5).

Table 5. 1986 US EPA Criteria for Indicators for Bacteriological Densities[a]

Freshwater	Acceptable swimming associated gastroenteritis rate per 1000 swimmers	Steady state geometric mean indicator density[c]	Single Sample Maximum Allowable Density			
			Designated beach area (upper 75% C.L.)[bc]	Moderate full body contact recreation (upper 82% C.L.)	Lightly used full body contact recreation (upper 90% C.L.)	Infrequently used full body contact recreation (upper 95% C.L.)
E. coli	8	126 CFU /100mL[d]	235	298	409	575

[a] Partial reproduction from the US EPA 1986 Ambient Water Quality Criteria for Bacteria

[b] Confidence limit

[c] Standard applied to sites tested in this project

[d] Colony Forming Units per 100mL

The *E. coli* monitoring indicated that sites Q7, Q8, Q10, Q12, Q13, Q15, Q16, and Q18 were the only quarterly sites that did not exceed 235 CFU/100mL at least once, while several locations exceeded the standard on more than one occasion (Table 6). Moreover, sites Q1, Q2, and Q14 exceeded the geometric mean standard. The monthly sites (M1, M2, M3) are of particular concern. They were monitored on a monthly basis because no prior data existed. The objective was to collect enough data to determine if the sites should be included on the official impaired waters list. All three sites exceeded the 235 CFU/100mL standard at least once, and site M2 was the only to not exceed more than once.

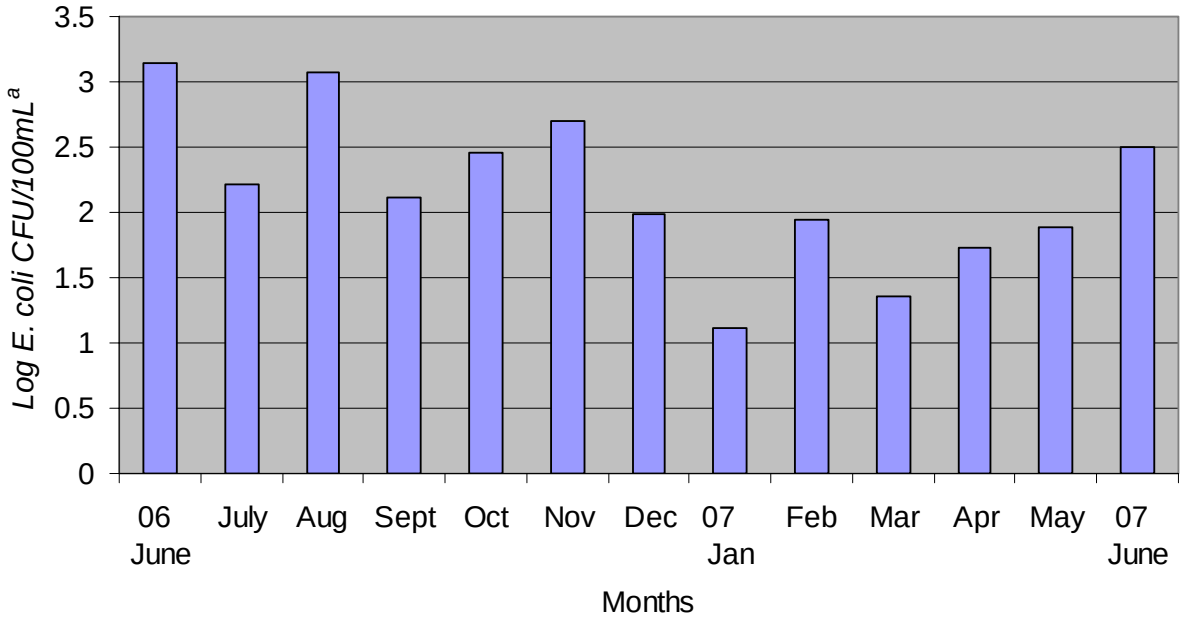
The quarterly sites showed some of the highest *E. coli* densities, and several locations possessed densities that exceeded the standard more than once. Sites Q5 and Q14 exceeded the standard on two occasions, while sites Q1 and Q2 showed densities higher than the standard in more than half the number of sampling events. The variability between sites, on a per month basis, was high enough that the standard deviation was greater than the mean in 7 of the 13 months (Table 6). Because monthly sites are compared with quarterly sites, it would be difficult to pinpoint a single site for the cause in variability; however, site Q14 was the only location for which, on more than one occasion, a density was more than twice the standard deviation for that month. June 2006 showed the highest average, and generally higher densities were recorded during warmer months. Nevertheless, the effects of month ($F = 1.10$, $p = 0.37$) and site ($F = 1.42$, $p = 0.14$) on *E. coli* density was assessed by two-way ANOVA, demonstrating no significant difference ($p < 0.05$). This is illustrated by Figures 3 and 4.

Table 6. *E. coli* Monthly Sampling Densities (CFU/100mL)[a]

Site[b]	06 June	July	Aug	Sept	Oct	Nov	Dec	07 Jan	Feb	Mar	Apr	May	07 June	Geometric Mean[c]
M1	---	315	3300	173	395	295	85	9	86	5	25	44	72	95
M2	---	78	160	143	232	90	**475**	19	154	23	14	64	210	88
M3	---	105	38	55	245	1100	53	10	27	**157**	123	123	94	87
Q1	1600	---	---	**353**	---	---	24	---	---	42	---	---	650	***206***
Q2	2960	---	---	275	---	---	103	---	---	36	---	---	1000	***313***
Q3	1140	---	---	60	---	---	33	---	---	38	---	---	140	104
Q4[d]	---	---	---	135	---	---	<10	---	---	0	---	---	800	31
Q5	320	---	---	45	---	---	<10	---	---	7	---	---	620	56
Q6	220	---	---	318	---	---	78	---	---	18	---	---	108	101
Q7	40	---	---	113	---	---	<10	---	---	19	---	---	134	40
Q8	160	---	---	128	---	---	<10	---	---	13	---	---	88	46
Q9	40	---	---	63	---	---	**430**	---	---	30	---	---	48	69
Q10	60	---	---	85	---	---	36	---	---	7	---	---	11	27
Q11	200	---	---	68	---	---	245	---	---	18	---	---	101	90
Q12	140	---	---	38	---	---	32	---	---	1	---	---	183	32
Q13	85	---	---	225	---	---	100	---	---	0	---	---	135	48
Q14	**15200**	---	---	65	---	---	23	---	---	23	---	---	**1840**	***249***
Q15	45	---	---	43	---	---	38	---	---	7	---	---	15	24
Q16	85	---	---	45	---	---	43	---	---	7	---	---	210	47
Q17	260	---	---	230	---	---	133	---	---	6	---	---	150	94
Q18[d]	---	---	---	100	---	---	73	---	---	23	---	---	105	65
x-bar[e]	1410	166	1166	131	291	495	97	13	89	23	54	77	320	***149***
s[f]	3761	130	1849	96	91	534	130	6	64	33	60	41	444	

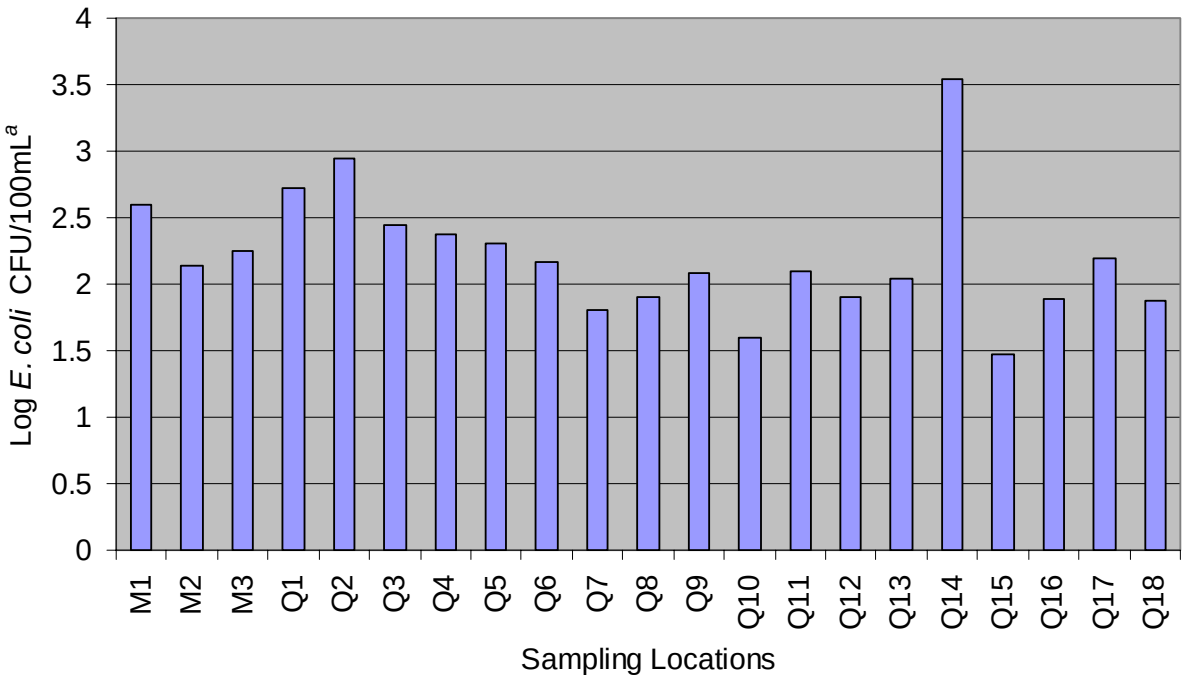
[a] Colony forming units per 100mL of water
[b] Monthly site monitoring began in July 2006
Quarterly sampling occurred in June, Sept, Dec 2006, and Mar and June 2007
[c] Mean of the monthly densities for each site as calculated by (June 06*July*Aug…*June 07)^1/13
Bold Italic indicates exceedance of water quality standard of 126 CFU/100mL (geometric mean standard)
[d] Sampling began in Sept 2006
[e] Average across all sites for the month
[f] Standard deviation across all sites for the month. Counts that are two or more standard deviations above the mean are in **Bold**.

Figure 3. Month to Month Comparison of *E. coli* Density



[a] Log transformed values based on monthly average across all sites sampled during the given month

Figure 4. Site to Site Comparison of *E. coli* Density



[a] Log transformed values based on yearly average for each location

As mentioned previously, monitoring and MST work in Prince William County dates back to July 2003. Ten of the eighteen quarterly sampling locations were part of the initial effort, and were therefore monitored on a monthly basis. For the sake of consistency, the cool season will be defined as the period from November to April and the warm season will be from May through October (3-6). Compared to the current project, many similarities could be seen in the monitoring results from the first year. Quarterly Sites 1 and 2 exceeded the 235 CFU/100mL standard the same number of times during the same months, and both exceeded the geometric mean standard, over the course of the current project. All exceedances occurred during the warm season. This compares quite similarly to the results of the first year, when 80% (4 of 5) of exceedances at Quarterly Site 1, and 100% (3 of 3) of exceedances at Quarterly Site 2 occurred during the warm season. Because of the high rate of exceedance during the first year, these two sites were selected for future monitoring, on a quarterly basis. The next round of monitoring data for Quarterly Sites 1 and 2 was collected from September 2005 to June 2006, where June 2006 was the first month of monitoring for the current project. Again, exceedances occurred at both locations during the warm season. Although the overall number of exceedances at all sites was down in Year 3, a pattern is beginning to emerge. Fecal bacterial populations can be "stored" on landscapes during drier, cooler months, and then be washed into receiving waters once precipitation occurs (5). In addition to the similarity in seasons, the lack of improvement over time suggests that Quarterly Sites 1 and 2 must continue to be monitored.

Quarterly Sites 3, 5, 6, 7, and 8 followed a similar initial pattern to Q1 and Q2. They were all monitored beginning in the first year and the initial diagnosis was bleak. All exceeded the 235 CFU/100mL standard more than 10% of the time, and the tendency toward warm season exceedance was seen as well. Eighty percent (5 of 6) of the exceedances for Quarterly Site 3

were in the warm season, while 100% (3 of 3) fell in the warm season for Quarterly Site 5. The two exceedances at Quarterly Site 6 were in the warm season, while 75% (3 of 4) and 100% (8 of 8) of the exceedances at Quarterly Sites 7 and 8, respectively, were in the warmer months. None of these sites were monitored again until the third year, and the results showed signs of improvement. Although not as much data was collected in Year 3 because they were monitored only quarterly, they each exceeded the standard only once. Quarterly Sites 3 and 5 each exceeded in June 2006 while sites 6-8 exceeded in September 2005. Quarterly Sites 3, 6 and 8 continued to demonstrate improved water quality through the current project, as they never again exceeded the standard. Quarterly Site 5 maintained its pattern of lower densities in the cooler months but the densities once again rose above the standard in June 2007. Similarly, Quarterly Site 6 exceeded the standard near the end of the warm season, in September 2006. In terms of *E. coli* monitoring and enumeration, Quarterly Sites 3, 7 and 8 appear to be headed in the right direction, but monitoring should continue in order to demonstrate consistently lower densities. Quarterly Sites 5 and 6 should continue to be monitored on a quarterly basis until the warm season densities fall consistently below the standard.

Quarterly Site 9 is one of only two sites that have been monitored every year of the ongoing project. It also has not followed the warm season exceedance trend as closely as other locations. The first year showed an exceedance rate of 27.3%, with two of the three occurring in the cool season. When monitored quarterly in Year 2, the exceedances shifted towards the warm season, but high densities were still recorded in December 2004. December proved again to be a troublesome month in Year 3, the only exceedance during that sampling period. Interestingly, the same pattern occurred yet again during the current project, with the only exceedance occurring during the December 2006 collection. The rate of exceedance over the course of any

given sampling year never fell below 25%. For this reason and because the cool season shows a definite pattern of exceedance, Quarterly Site 9 has consistently demonstrated the need for continued monitoring.

Quarterly Site 10 had one of the lowest rates of exceedance (25%) in Year 1, and it has shown steady improvement since. Although the rate of exceedance remained at 25% in Year 3, it had no exceedances during the current project year. These results should be confirmed with additional monitoring. In terms of *E. coli* monitoring, Quarterly Site 11 has made the biggest strides toward improvement. When monitored monthly in Year 1, 8 of the 12 (66.7%) samples exceeded the standard, with 62.5% (5 of 8) occurring in the cool season. Since monitoring began again in Year 3, it has only exceeded the standard twice, in September 2005 and December 2006. The yearly exceedance average had dropped from 66.7% to 25%. Despite these encouraging results, Quarterly Site 11 is still showing signs of impairment and should therefore continue to be monitored on a quarterly basis.

Monitoring of Quarterly Site 12 began in Year 2 and it was the only monthly site that year to never exceed the standard. It was scaled back to quarterly monitoring in Year 3 in order to confirm these encouraging results. Indeed, the densities remained below the exceedance level in Year 3 and the current year. Although the initial results in Year 2 were not as promising for Quarterly Site 13, subsequent monitoring has demonstrated similar improvement. Never again has the standard been exceeded at this site, and the densities during the initially troublesome cooler months were particularly low. Nevertheless, both of these sites should continue to be examined on a quarterly basis in order to account for any changes that may occur in such a rapidly developing, suburban watershed. Monitoring is inexpensive and requires minimal manpower and materials.

The warm season has proven particularly ominous for Quarterly Site 14. By the time sampling had concluded in Year 2, 5 of the 13 (38.4%) samples had exceeded the standard, with 4 occurring during the warm season. Things appeared to be improving until June 2006, when the highest density that month was recorded at this site. The densities decreased during the cooler months of 2006, in accordance with the pattern of earlier results, but rose again during the June 2007 collection. The highest densities that month were again seen at this site. Further monitoring should be done to determine if this pattern will continue to persist.

With a couple exceptions during the warm season, Quarterly Sites 15 and 16 have been relatively clean since monitoring began in Year 2. Two of the three exceedances at Quarterly Site 15 occurred in the warm season of Year 2, but all subsequent densities have been below the standard. Similarly, Quarterly Site 16 had two exceedances during the initial monthly monitoring, but has been clean ever since. As with the other sites that have been relatively clean in recent years, these sites should not be regarded as "fixed." The purpose of quarterly monitoring is to assess whether a site has improved, and to ensure no major changes occur at the sites that have.

Quarterly Site 17, the final location for which more than a year of data has been collected, had seven exceedances in its first year. Moreover, it did not show an easily discernible seasonal pattern, as three of the seven exceedances were in the cool season. Quarterly monitoring in Year 3 showed little change. Two of the four sample collections resulted in exceedances, split evenly between the warm and cool seasons. Although encouraging results were obtained during the current project, caution should be taken before considering this site improved. Additional monitoring would be needed to confirm the positive results of Year 4.

The remaining five sampling sites, consisting of two quarterly and three monthly, were all first-time locations. Three of the five sites exceeded the *E. coli* standard more than 10% of the time, with Monthly Site 1 the most egregious. Monthly Site 1 had a definite seasonality effect, with three of its four exceedances coming in the warm season. The sample collections in 2007 had much lower densities but only additional sampling will be able to determine if this will continue. Monthly Site 3 had two exceedances but no clear seasonality trend. Quarterly Site 4 exceeded the standard in June 2007, but additional monitoring would have to be completed to determine if warm weather plays a role with this location. The remaining two sites, Quarterly Site 18 and Monthly Site 2, showed encouraging results. Quarterly Site 18 had no exceedances and Monthly Site 2 had only one. Because these sites have only one year of data, it would be difficult to conclude that they are clean. Future monitoring results will offer more information and help develop a more complete picture.

II. The Known Source Library (Tables 7-9)

In addition to the steps taken previously by Touchton (10), several measures were incorporated to increase confidence in the correct classification rates of the KSL. The first step was to update the existing KSL with fresh known source isolates from Prince William County. As mentioned in Chapter III, the KSL consisted of both local and regional isolates. The two sampling trips in February and March 2007 added a total of 508 known source isolates to the existing library of 1092 isolates, for an overall total of 1600. Of the update isolates, 61 were human, 90 were pet, 93 were livestock, and 264 were wildlife (Table 7). Interestingly, after adding the update isolates and reclassifying, the rate of correct classification (RCC), based on discriminant analysis, decreased for each known source, with the average rate of correct

classification (ARCC) falling from 75% to 61% (Table 7). This was not completely unexpected, however, as other projects have demonstrated that old isolates can become obsolete, or temporally unstable (7, 8, 14).

Moreover, the library had not yet been analyzed for clonal resistance patterns. Recent literature has demonstrated that the ARCC will increase when only unique ARA patterns are used with the KSL (2). Using the sort function in SAS-JMP, a total of 597 clonal patterns (120 human. 26 pet, 123 livestock, and 328 wildlife) were identified and subsequently removed. The library now (only) consisted of 1003 unique resistance patterns. After clonal isolate removal, the RCC for each source category (human, pets, livestock, wildlife) increased to a level higher than that which existed before the update isolates were added. The ARCC increased from 75% to 93% (Table 7). This was in agreement with the expectation that an updated KSL should more accurately classify known source isolates from its own watershed (7, 8, 14).

Table 7. Summary of Known Source Library Isolates and RCC[a] Values Using Discriminant Analysis

	Human	Pet	Livestock	Wildlife	Total
Old Isolates	237	91	207	557	1092
Old RCC	74%	92%	78%	72%	75%[b]
Update Isolates	61	90	93	264	508
Updated RCC[c]	59%	76%	58%	60%	61%[b]
Total Isolates[d]	298 (178)	181 (155)	300 (177)	821 (493)	1600 (1003)
Final RCC	**89%**	**94%**	**89%**	**95%**	**93%**[b]

[a] Rate of Correct Classification

[b] Average Rate of Correct Classification

[c] After updating the library with fresh known source isolates

[d] Numbers in parentheses are what remained after removal of clonal isolates and old, poorly classified isolates, and were used to produce Final RCC and ARCC values

Additional steps were taken to ensure the validity of the KSL, including the application of a supplementary means of isolate classification, and the use of two challenge sets. Dickerson et al. (1) used two separate statistical algorithms, discriminant analysis and logistic regression, to determine the sources of fecal pollution at two Virginia beaches. Discriminant analysis was shown to offer better classification for a four-way split, while logistic regression was better for a two-way (human vs. non-human) split. However, the difference between the two, regardless of the size of the category split, was generally less than 5%. The two-way split was not considered for this project because a four-way split was sufficient, but logistic regression was utilized to compare the RCC produced by discriminant analysis for the four category classification. Table 8 summarizes the comparison between the two algorithms, showing that logistic regression presented a higher RCC for each source category (human, pets, livestock, wildlife), but the differences were all less than 5%, with the exception of the human classification. The ARCC for discriminant analysis was 3% lower than the value generated by logistic regression, but both were above 90% (Table 8).

Table 8. Discriminant Analysis vs. Logistic Regression for Four Category Classification[a]

Discriminant Analysis	Percentage (number) of isolates classified as				
	Human	**Pet**	**Livestock**	**Wildlife**	
Human (n = 178)	89 (158)[b]	<1 (1)	1 (2)	1 (7)	
Pet (n = 155)	5 (9)	94 (145)	7 (13)	4 (18)	ARCC = 93%
Livestock (n = 177)	2 (3)	1 (2)	89 (157)	<1 (2)	
Wildlife (n = 493)	4 (8)	5 (7)	3 (5)	95 (466)	
Logistic Regression	Percentage (number) of isolates classified as				
	Human	**Pet**	**Livestock**	**Wildlife**	
Human (n = 178)	96 (170)	0	2 (3)	1 (4)	
Pet (n = 155)	1 (2)	95 (147)	2 (4)	1 (7)	ARCC = 96%
Livestock (n = 177)	1 (2)	1 (2)	94 (167)	1 (4)	
Wildlife (n = 493)	2 (4)	4 (6)	2 (3)	97 (478)	

[a] Using the updated KSL

[b] Underlined values indicate RCC for each source category

The results of two separate challenge sets offered further information about the KSL. Table 9 summarizes the classification success for the repeatability and accuracy tests, and for the KSL. The ARCCs for the two challenge sets were lower than the ARCC for the KSL, but neither was substantially lower than other researchers have reported. Dickerson et al. (1), for example, reported ARCCs of 81.8% and 83.8% for the KSL and challenge set, respectively. Graves et al. (2) calculated an ARCC of 65% for a six-category challenge set. The MDP was calculated by averaging together the AFMs for the repeatability test, accuracy test, and KSL (Table 9). The MDP of 14% was slightly lower than the 17.3% for Dickerson et al. (1) and the 22.5% for Graves et al. (2).

Table 9. Challenge Set Classification Using Repeatability Test and Accuracy Test

Repeatibility Test[a]	Percentage (number) of isolates classified as				
	Human	**Pet**	**Livestock**	**Wildlife**	
Human (n = 20)	85 (17)[b]	0	0	0	ARCC = 86%
Pet (n = 20)	5 (1)	90 (18)	15 (3)	5 (1)	
Livestock (n = 20)	5 (1)	0	70 (14)	0	AFM = 14%
Wildlife (n = 20)	5 (1)	5 (7)	15 (3)	95 (19)	

Accuracy Test[a]	Percentage (number) of isolates classified as				
	Human	**Pet**	**Livestock**	**Wildlife**	
Human (n = 20)	65 (13)	0	0	0	ARCC = 78%
Pet (n = 20)	5 (1)	80 (16)	20 (4)	0	
Livestock (n = 20)	15 (3)	0	45 (9)	0	AFM = 22%
Wildlife (n = 20)	15 (3)	20 (4)	35 (7)	100 (20)	

Known Source Library[a]	Percentage (number) of isolates classified as				
	Human	**Pet**	**Livestock**	**Wildlife**	
Human (n = 178)	89 (158)	<1 (1)	1 (2)	1 (7)	ARCC = 93%
Pet (n = 155)	5 (9)	94 (145)	7 (13)	4 (18)	AFM = 7%
Livestock (n = 177)	2 (3)	1 (2)	89 (157)	<1 (2)	MDP = 14%[c]
Wildlife (n = 493)	4 (8)	5 (7)	3 (5)	95 (466)	

[a] Using discriminant analysis

[b] Underlined values indicate RCC for each source category

[c] Calculated by averaging AFM of each classification

Discriminant analysis (DA) was the main method of classification for the current project. DA is used to find the linear combination of features which best separates two or more classes of an event. Simply stated, the probability of an input *x* being in a class *y* is a function of the linear combination of the known observations. In this case, the classes are the known source categories (human, pet, livestock, wildlife) and the known observations are the resistance patterns. The probability calculation for each source (human, pets, livestock, wildlife) is made independently, and the each isolate is classified into only its most probable class.

DA classifies every isolate into a source category (human, pets, livestock, wildlife), even if the probability is low for the isolate to be in any category. This could be considered a drawback of DA, but there are several ways to mitigate the consequences. It has been suggested that the creation of an "unknown" category would enable isolates that do not fit any source category to be effectively removed (10). However, the number of isolates left for classification might shrink considerably, and would therefore require additional known source collection. The degree of additional manpower required for this solution was beyond the scope of this project. A second solution is known as cluster analysis. Cluster analysis is the relabeling of similar isolates of a source category into a subcategory (10). When applied before performing DA, clustering removes subcategories that expand the confidence interval of a source category. The tighter confidence intervals improve the classification. This method was applied by Touchton (10) on an library of enterococci isolates, with mixed results. Because the clustering mechanism had no percentage similarity, the designation of categories and subcategories was purely subjective. Furthermore, the results were based on enterococci isolates, and the source tracking results did not compare well with their *E. coli* library for the same watershed.

The best approach to interpreting the KSL was the calculation of the minimum detectable percentage (MDP) and verification with logistic regression. The MDP established a threshold for known source classification. Any known source category that was classified below the MDP was considered a negligible contributing source. Logistic regression validated the classification rates produced by DA. This was in accordance with the method used by Dickerson et al. (1).

After addition of the update isolates and removal of clonal isolate patterns, the KSL was comprised of 178 human, 155 pet, 177 livestock, and 493 wildlife isolates. The library may have been better served with a more even distribution of known source isolates, but the effect on the representativeness of the library would nevertheless be difficult to measure. More important than the distribution is the total number of isolates (9, 14). A library with more than 1000 isolates, regardless of whether the known sources are divided evenly, has been representative for multiple watersheds (2, 7, 8, 13, 14). Because this project is ongoing, the KSL for Prince William County will have to continue to updated on at least a yearly basis. No attempt was made to classify environmental water isolates (see Section III) using the older, non-updated (with clonal isolates still intact) KSL, as there would be no baseline for comparing the results. Because the updated KSL (with clonal isolates removed) was shown to have a higher RCC for each source category and a higher ARCC, it was used to make all subsequent classifications.

III. Environmental Water Isolates (Tables 10-34)

A total of 2854 environmental *E. coli* isolates were isolated for purposes of ARA. The pet category was classified as the major signature at 16 of the 21 sampling locations, a sizable majority (Table 10). The second most frequent major signature was livestock, classified four times. Livestock and pets were classified at identical percentages at Quarterly Site 4. This left only two sites, where wildlife was the major signature. Wildlife was the most common secondary signature, found at all but two sites. Livestock was a secondary signature at 13 locations, while pets were found at 5. The pet and wildlife categories were classified as either a minor or major signature at all 21 locations, while livestock fell below the minimum detectable percentage (MDP) at four locations. Perhaps the most interesting result was that a human source was not classified even a single time as either a major or minor signature. In fact, on a monthly basis, the human category was classified only four times above the MDP (Tables 14, 15, 18, 32). Table 11 shows the relative fraction of each classification for all source categories.

Table 10. Major and Minor Signatures at Each Sample Location After MDP Adjustment[a]

Site	Major[b]	Fraction	Minor[c]	Fraction	Fluorometry[d]
M1	Pet	0.510	Wildlife Livestock	0.240 0.180	63.0
M2	Pet	0.470	Wildlife Livestock	0.291 0.209	57.8
M3	Pet	0.529	Livestock Wildlife	0.235 0.221	31.8
Q1	Pet	0.616	Wildlife	0.286	39.3
Q2	Pet	0.483	Wildlife Livestock	0.268 0.168	81.8
Q3	Pet	0.450	Wildlife Livestock	0.362 0.188	80.0
Q4	Pet Livestock	0.375 0.375	Wildlife	0.250	39.7
Q5	Pet	0.580	Wildlife Livestock	0.203 0.200	28.4
Q6	Livestock	0.352	Wildlife Pet	0.340 0.296	46.6
Q7	Wildlife	0.358	Pet Livestock	0.332 0.308	41.2
Q8	Pet	0.453	Livestock Wildlife	0.318 0.218	29.6
Q9	Pet	0.410	Livestock Wildlife	0.300 0.280	56.8
Q10	Pet	0.562	Wildlife	0.346	58.2
Q11	Livestock	0.373	Pet Wildlife	0.315 0.298	73.0
Q12	Pet	0.600	Wildlife Livestock	0.223 0.148	37.8
Q13	Pet	0.605	Wildlife	0.263	49.0
Q14	Pet	0.378	Livestock Wildlife	0.345 0.253	75.6
Q15	Pet	0.497	Wildlife Livestock	0.283 0.207	47.4
Q16	Pet	0.404	Wildlife Livestock	0.362 0.188	52.0
Q17	Livestock	0.383	Pet Wildlife	0.370 0.230	54.0
Q18	Wildlife	0.473	Pet	0.427	66.5

[a] Using the MDP of 14%
[b] The source(s) with the highest average classification percentage(s) over 13 months
[c] The remaining source(s) with average classification percentage(s) higher than the MDP
[d] For the detection of optical brighteners

Table 11. Relative Fraction of Classified Isolates[a]

Site	Human	Pet	Livestock	Wildlife
M1	0.070	0.510	0.180	0.240
M2	0.031	0.470	0.209	0.291
M3	0.018	0.529	0.235	0.221
Q1	0.008	0.616	0.092	0.286
Q2	0.083	0.483	0.168	0.268
Q3	0.000	0.450	0.188	0.362
Q4	0.000	0.375	0.375	0.250
Q5	0.013	0.580	0.200	0.203
Q6	0.016	0.296	0.352	0.340
Q7	0.000	0.332	0.308	0.358
Q8	0.010	0.453	0.318	0.218
Q9	0.010	0.410	0.300	0.280
Q10	0.018	0.562	0.074	0.346
Q11	0.013	0.315	0.373	0.298
Q12	0.035	0.600	0.148	0.223
Q13	0.000	0.605	0.135	0.263
Q14	0.028	0.378	0.345	0.253
Q15	0.017	0.497	0.207	0.283
Q16	0.050	0.404	0.188	0.362
Q17	0.020	0.370	0.383	0.230
Q18	0.033	0.427	0.067	0.473

[a] Row totals may not equal 1.00 due to rounding

As mentioned in Chapter III, sediment samples were collected in June 2007 to compare with the environmental samples from the same month. The chi-squared test showed no statistically significant difference ($p > 0.05$) in the MST results for 11 of the 16 locations for which comparisons were able to be made (Table 12). Of the five locations where statistically significant differences were found, four showed different major signatures. The major signature at Quarterly Site 1 was livestock for the environmental isolates but switched to the pet category with the sediment isolates. For Quarterly Site 3, the major signature went from livestock for the environmental isolates to wildlife for the sediment isolates. Quarterly Site 8 went from being overwhelmingly classified as pet with the environmental isolates to a more even split between pet and livestock with the sediment isolates. Similarly, Quarterly Site 14 switched from pet being the major signature with the environmental isolates to livestock being classified in more than half of the sediment isolates. Quarterly Site 4, despite showing a statistically significant difference, maintained the pet category as the major signature, and it became even more dominant with the sediment samples.

Another important observation was the similarities between environmental and sediment isolates with regard to human classification. Among the locations for which sufficient isolates were collected for comparison, the human category was classified above the MDP a total of four times. Two of these instances occurred with the Quarterly Site 2 samples; both environmental and sediment samples were well above the MDP. The other two instances resulted in a difference between environmental and sediment classification. The environmental isolates for Monthly Site 2 had a human classification above the MDP, while the sediment isolates fell slightly below. Conversely, the sediment isolates for Monthly Site 3 were slightly above the MDP, while the environmental isolates were classified below.

A similar attempt involving environmental and sediment samples was attempted in Year 1 of the PWC project. Like the current project, the source tracking results for the Year 1 sediment samples essentially "mirrored" the results from the environmental samples, in that dominant sources in the environmental samples were also dominant in the sediments, and minor sources in the environmental samples were also minor in the sediments. It has been postulated in the scientific literature that fecal bacteria from some types of major sources such as wildlife, birds, (or livestock in Slate Run, for example) might die out quickly and not be found in sediments to any great extent, or fecal bacteria from minor sources such as dogs or humans might become established in sediments and persist even after the source was removed. The results from the previous 18 tables do not support either of these possibilities, and this should be good news to the scientific and regulatory communities.

The Year 1 sediment comparison led to many interesting observations. The same seasonality was observed for the sediment samples as had been recorded for the regular samples; birds, livestock, and pets were highest in the warm season and lower in the cool season, while the human signature was highest in the cool season and lowest in the warm season. This seasonality indicates that the fecal bacterial populations decline as pressure from the source declines. Cattle, for example, spend many hours per day standing in water during the summer if stream access is unrestricted, but rarely stand in the water during the winter. The reduction of the livestock signature in Cedar Run and Slate Run, in both the regular and the sediment samples in the winter months, shows this same trend and implies that if cattle were totally removed from the stream, the livestock signature would eventually disappear. It is hard to speculate on how long that might take, as there are no published studies that have examined this time element, but it will probably require at least a few months.

For the sites where a human signature was obtained, it was found almost always in the winter months when seasonal water tables are at their highest levels and septic-tank drainfields are much closer to elevated water tables (if not actually immersed). This proximity to a water table "short circuits" the aerobic treatment in the separation zone beneath drainfields, and expedites the transport of fecal bacteria into seasonal water tables where they can be carried into streams that drain the water tables. The absence of a human signature in the summer months indicates that human fecal bacteria are unable to become established and persist in the sediments (or the water column) without constant recharge from a contaminant source. This is good news, as it indicates that the fecal bacteria from a given source can be either eliminated or greatly reduced in a stream if specific sources can be located and closed off or removed.

Table 12. Comparison of Environmental and Sediment Samples from June 2007

			Number of *E. coli* Isolates Classified (%)					
Site	**CFU/100mL**	**Isolates**	**Human**	**Pet**	**Livestock**	**Wildlife**	**X^2 and df**	**p-value**
M1	72	24	12 (50)	4 (17)	8 (33)	0	---	---
M1 Sed[a]	35	---	---	---	---	---	---	
M2	210	24	4 (17)	2 (8)	15 (63)	3 (13)	$X^2 = 0.58$	0.902
M2 Sed	1100	24	3 (13)	3 (13)	16 (67)	2 (8)	df = 3	
M3	94	24	1 (4)	7 (29)	16 (67)	0	$X^2 = 3.19$	0.363
M3 Sed	180	24	4 (17)	6 (25)	13 (54)	1 (4)	df = 3	
Q1	650	24	0	8 (33)	11 (46)	5 (21)	$X^2 = 9.13$	0.010
Q1 Sed	420	24	0	17 (71)	7 (29)	0	df = 2	
Q2	1000	24	8 (33)	7 (29)	9 (38)	0	$X^2 = 5.45$	0.066
Q2 Sed	1190	24	10 (42)	1 (4)	13 (54)	0	df = 2	
Q3	140	24	0	1 (4)	21 (88)	2 (8)	$X^2 = 19.1$	0.000
Q3 Sed	450	24	0	8	6 (25)	10 (42)	df = 2	
Q4	800	24	0	10 (42)	9 (38)	5 (21)	$X^2 = 4.81$	0.019
Q4 Sed	1310	24	1 (4)	16 (67)	4 (17)	3 (13)	df = 3	
Q5	620	23	1 (4)	18 (78)	2 (9)	2 (9)	$X^2 = 3.45$	0.327
Q5 Sed	240	23	0	14 (61)	4 (17)	5 (22)	df = 3	
Q6	108	24	2 (8)	10 (42)	9 (38)	3 (13)	$X^2 = 4.85$	0.183
Q6 Sed	270	22	0	15 (68)	4 (18)	3 (14)	df = 3	
Q7	134	23	0	1 (4)	20 (87)	2 (9)	$X^2 = 1.36$	0.507
Q7 Sed	53	23	0	3 (13)	19 (83)	1 (4)	df = 2	
Q8	88	24	0	22 (92)	2 (8)	0	$X^2 = 15.8$	0.001
Q8 Sed	330	24	1 (4)	9 (38)	10 (42)	4 (17)	df = 3	
Q9	48	21	0	13 (62)	6 (29)	2 (10)	$X^2 = 6.30$	0.098
Q9 Sed	50	14	1 (7)	4 (29)	4 (29)	5 (36)	df = 3	
Q10	11	15	0	4 (27)	3 (20)	8 (53)	---	---
Q10 Sed[a]	6	---	---	---	---	---	---	
Q11	101	20	0	8 (40)	10 (50)	2 (10)	---	---
Q11 Sed[a]	38	---	---	---	---	---	---	
Q12	183	23	0	11 (48)	6 (26)	6 (26)	$X^2 = 4.85$	0.183
Q12 Sed	67	19	0	15 (79)	3 (16)	1 (5)	df = 3	
Q13	135	24	0	18 (75)	5 (21)	1 (4)	$X^2 = 3.90$	0.273
Q13 Sed	93	24	1 (4)	21 (88)	1 (4)	1 (4)	df = 3	
Q14	1840	23	0	15 (65)	6 (26)	2 (9)	$X^2 = 7.93$	0.019
Q14 Sed	2300	12	0	3 (25)	9 (75)	0	df = 2	
Q15[a]	15	---	---	---	---	---	---	---
Q15 Sed	170	22	3 (14)	15 (68)	2 (9)	2 (9)	---	
Q16	210	24	0	6 (25)	15 (63)	3 (13)	$X^2 = 5.09$	0.078
Q16 Sed	290	23	0	9 (39)	7 (30)	7 (30)	df = 2	
Q17	150	18	0	7 (39)	2 (11)	9 (50)	$X^2 = 0.210$	0.900
Q17 Sed	400	12	0	4 (33)	1 (8)	7 (58)	df = 2	
Q18	105	22	0	8 (36)	0	14 (64)	---	---
Q18 Sed[a]	7	---	---	---	---	---	---	

[a] Insufficient number of isolates for MST

Comparing MST results across different years is essential to understanding whether change has occurred, and whether improvements with regard to specific known sources have been made. Similar to the monitoring results, the purpose of this section is to compare the MST results of each site for the current project to the results obtained in earlier years. The analysis is organized by stream. Included with this historical analysis is a discussion of ways to clean up each site and/or stream, based on the major and minor signatures.

The first year of the project used a five category library, the same four as the current project plus a bird category, for all locations. In Year 2 the bird category was combined with the wildlife isolates for a four category classification. Year 3 went back to the five category classification and the current project elected to use four categories. Table 13 shows the sample locations and the years sampled.

In the current project, contamination from pets was the most frequent major signature, occurring at 16 of the 21 sites. Livestock was the second highest, followed by wildlife. Minor signatures consisted mostly of wildlife and livestock. At no site was human contamination either a major or minor signature.

Table 13. Sample Locations and Years Sampled

Site #	Location	Year(s) Sampled[a]
Q1	Neabsco Creek, Lindendale Road	1,3,4
Q2	Neabsco Creek, Benita Fitzgerald Road	1,2,3,4
Q3	Neabsco Creek, Neabsco Mills Rd. & Route 1	1,3,4
Q4	Cow Branch, Montgomery Ave.	4
Q5	Cow Branch, Rippon Landing Park	1,3,4
Q6	Powell's Creek, Fox Mills Apt. & Route 1	1,3,4
Q7	Quantico Creek, South Fork, Joplin Road	1,3,4
Q8	Quantico Creek, Main Stem, Mine Road & I-95 Overpass	1,3,4
Q9	Cedar Run, Carraige Ford Road	1,2,3,4
Q10	Cedar Run, Bristow Road	1,3,4
Q11	Slate Run, Old Church Road	1,3,4
Q12	Bull Run, Route 28	2,3,4
Q13	Catharpin Run, Robin Drive	2,3,4
Q14	Flat Branch, Lomond Drive	2,3,4
Q15	South Run, Buckland Mill Road	2,3,4
Q16	Broad Run, Route 28	2,3,4
Q17	Kettle Run, Valley View Road	2,3,4
Q18	North Fork of Lake Manassas	4
M1	Little Bull Run, Pageland Lane	4
M2	Powell's Creek, Northgate Dr.	4
M3	Broad Run, Route 55	4

[a] Year 1 = July 2003 - June 2004
Year 2 = July 2004 - June 2005
Year 3 = July 2005 - June 2006
Year 4 = July 2006 - June 2007

Quarterly Sites 1-3: Neabsco Creek

Wildlife and birds were the primary contributors of fecal contamination according to the classification in Year 1. The pet signature was minor but indicates that dogs were an issue. The human and livestock signatures were below the Year 1 MDP (8%) and could therefore be disregarded. The large allocations for birds and wildlife were most likely associated with the unimproved green areas directly adjacent to each sampling location (6). Each sample location was situated in a commercial or highly suburbanized area of the county, where any wildlife and birds that may reside would be forced into tight green spaces. This would lead to a higher concentration in fecal matter at specific locations.

Quarterly Site 2 was the only Neabsco Creek location monitored in Year 2. There was virtually no change in the classification averages. Again, no major changes occurred when all three locations were monitored again in Year 3, no major changes. Wildlife was still the major signature at all locations, followed by birds and pets. The human and livestock signatures remained below the MDP. It was not until the current project that a shift began to take notice. Pets were classified as the major signature, while wildlife (including birds) remained a persistent but minor signature. The biggest surprise in Year 4 was the emergence of a livestock allocation at Quarterly Sites 2 and 3. In each case it was the most minor signature and could be attributed to misclassification that managed to occur marginally above the MDP. Importantly, the human signature remained below the MDP and was considered insignificant. Although a higher average was observed for pet than for wildlife, this was not altogether surprising. This stream is neither at nor near any rural areas of the county, so it was expected that pets and wildlife would be primary contributors. The fact that one was higher than the other is not as significant. These

results are consistent with expectation and future projects will most likely demonstrate a continued predominance of these two sources.

Dealing with high levels of wildlife contamination is a formidable task. In general, there are three methods for reducing wildlife contamination: reduce wildlife populations, reduce wildlife access, and/or reduce wildlife habitat (10). Each of these methods has major limitations. Populations of wildlife, such as deer, can be reduced by increased hunting, but this is highly impractical at this suburban location. Reducing wildlife access to necessities such as water is generally only useful with waterfowl. The third approach, reducing habitat, is a natural consequence of increased urbanization, but removing green zones is an unattractive alternative for many other reasons. The best approach may be to concede the wildlife signature, at least in highly developed areas, but continue to focus on the other sources. Posting signs encouraging the proper disposal of pet waste and having stricter fines may decrease the pet influence. Strict enforcement of leash laws may also prove beneficial.

Quarterly Sites 4-5: Cow Branch

Quarterly Site 5 was the only Cow Branch location for which more than one year of data has been collected. It was monitored in Years 1 and 3, during which wildlife was by far the largest source allocation. Similar to Neabsco Creek, the only secondary signature of significance was birds. This location was marked by multiple construction jobs during summer collection periods, and is rooted in a commercial district of PWC. The absence of a livestock signature was consistent with expectation, and low to absent human allocations indicate no major septic or sewer issues. The current project deviated from previous results with respect to the pet signature. Current results suggest that dogs are becoming an increasing issue for Cow Branch.,

while wildlife persists at lower levels than detected previously. It is difficult to explain the increased pet signature, and the emergence of a relatively minor livestock allocation, except to suggest that fecal bacteria may be concentrated upstream. These results should be confirmed by future study.

Monitoring at Quarterly Site 4 began with the current project. Different from Quarterly Site 5, this Cow Branch location is nestled in a quiet residential neighborhood. Although farms are no longer in the immediate vicinity, the surroundings could still be described as rural for PWC. Interestingly, this was the only location where the major signature was shared evenly by two sources, pets and livestock. These results suggest that in this residential area, unleashed dogs and improper disposal of pet waste has become an issue. Moreover, livestock waste may be washing down from other locations. Wildlife would not play as much of a role in this area, and the lower wildlife signature is therefore consistent with expectation. It is critical to collect additional samples at this site in order to assess any changes that may occur. It is difficult to assess exactly where the livestock signature is coming from, but future monitoring could determine whether it persists.

Quarterly Site 6 and Monthly Site 2: Powell's Creek

The results for Quarterly Site 6 in Year 1 were very similar to those obtained for Neabsco Creek. Wildlife and birds were the dominant signatures, while a significant pet signature indicated that dogs were a concern. Very little change occurred in Year 3, when quarterly monitoring commenced. Wildlife and birds still dominated, and the minor pet signature persisted. Neither year produced any allocation of human or livestock. The current project presented a surprising change in major source allocations. Livestock, which had never

previously surfaced, was classified at a slightly higher percentage than wildlife. Wildlife and pets were significant signatures, consistent with earlier years, but the livestock signature is difficult to account for. This site is in a quiet residential area which does not make a livestock signature as surprising as a highly commercial or industrial area might. Nevertheless, future source tracking at this location should be done in order to ascertain if this result continues. Consistent with earlier years, the human signature was below the MDP.

The new sampling location on Powell's Creek, Monthly Site 2, is situated in a quiet residential neighborhood surrounded by dense forest. The MST results suggest that dogs and wildlife are the biggest issues. A minor livestock signature was found but was not consistent with the surroundings. No human signature was detected. Efforts to reduce contamination should focus on the pet allocation.

Quarterly Sites 7-8: Quantico Creek

The two sampling locations on Quantico Creek were within the Quantico Marine Base, making them perhaps the most isolated sites. A large deer population makes it easy to expect a heavy wildlife signature, and Year 1 results did not disappoint. Wildlife and birds were the major sources of impairment at both sites, and only a minor pet signature was detected. Year 3 results were nearly identical. No human or livestock signature was detected above the MDP in either year.

Relatively little change occurred with the current project results for Quarterly Site 7. Wildlife remained the major source of impairment, while pets were secondary. The biggest difference was the appearance of a minor livestock signature. Due to the rural surroundings, it is

plausible that nearby farms could be a culprit. As expected, no human signature was detected. It would be interesting to see how future results might compare.

Quarterly Sites 9-10: Cedar Run

Quarterly Site 9 has been sampled consecutively for the last four years, providing a wealth of source tracking data. The results of Year 1 showed a wide array of contamination sources. Wildlife, livestock, bird, and human signatures were all found at levels above the MDP. Only a pet signature went unclassified. The predominant sources were wildlife and livestock, a logical result for this rural location. The mixed distribution is typical of larger streams and indicated the many different land uses taking place in the watershed surrounding Cedar Run. Year 2 brought about little change, the most notable being the emergence of a minor, but still insignificant, pet signature. Wildlife and livestock still dominated, but the human signature was still significant. The bird allocation increased slightly in Year 3, but was still lower than wildlife and livestock. The human signature dropped to the MDP threshold of 10%, and the pet signature remained insignificant.

The current results for Quarterly Site 9 suggest that unleashed dogs have become a bigger concern. Pets were the primary contributor, followed by livestock and wildlife. Similar to Year 3, the human signature was below the MDP. These results are not overly surprising, and the continued decrease in human allocation is encouraging. Contamination from livestock is expected in rural parts of the watershed where farming operations and homes are typically served by on-site septic systems. Additionally, on almost every sampling occasion multiple unleashed dogs were observed running in the surrounding fields.

The early results of Quarterly Site 10, monitored in Years 1 and 3, mirrored those of Quarterly Site 9. A discrepancy emerged, however, with the current project results. Only wildlife and pets were classified above the MDP, with pets being primary. The lack of a livestock signature was explained by the fact that this site has become more residential in recent years. The increased development can also explain the increased pet presence. Subsequent sampling should confirm that pets remain an increasing issue, while efforts to reduce livestock influence should be implemented at other points in the stream.

Quarterly Site 11: Slate Run

Similar to Cedar Run, the Year 1 allocation averages indicated several significant sources of contamination. This part of Slate Run had the highest average human allocation of any Year 1 site, but the wildlife and livestock were slightly higher. Birds classified above the MDP, and a negligible pet signature was found. The high livestock and wildlife signatures were consistent with the rural makeup. When monitored quarterly in Year 3, the biggest change was the wildlife signature. It went from being primary in Year 1 to the fourth highest signature in Year 3. Livestock was now the primary allocation, and the bird and human signatures both increased slightly. The pet signature remained below the MDP. The current project results suggest a continuing livestock trend, with an increasing pet issue. The wildlife signature has not disappeared and should not be disregarded. However, cleanup efforts in this rural area should focus on livestock and pets. Proper fencing of streams would go a long way towards decreasing contamination.

Quarterly Site 12: Bull Run

Sampled consecutively for the last three years, this location has seen little change in major source allocation. In Year 2, wildlife was classified in more than half of the isolates, and pets was the only significant secondary source. Wildlife and pets remained the two biggest signatures in Year 3. The only difference was the emergence of a small livestock signature. Albeit minor, it was considered significant because it was classified above the MDP. The only change in the current project results was the switching of the major and minor signatures; pets became primary while wildlife became secondary. The primary pet signature reflects the nature of the surroundings. Bull Run Regional Park is directly adjacent and maintains several walking paths for people to exercise, walk their dogs, etc. Additionally, this particular section has become a popular swimming hole, for people and their pets. The smaller livestock signature has endured and should be used as confirmation of the Year 2 results. No significant human signature was found in any year.

Quarterly Site 13: Catharpin Run

Similar to Bull Run, the sources of contamination have undergone little change over three years. Wildlife was the primary source of impairment in Year 2, followed by livestock and a very minor pet signature. Livestock took over as the primary contaminant in Year 3, but most likely because birds and wildlife were split into different categories. Wildlife and birds were both highly classified so it would be reasonable to expect that combining the two categories would have produced a higher classification than livestock. The pet signature remained low but significant.

Interestingly, the livestock signature was not classified by the current project library. A much higher pet signature was obtained with a slightly lower wildlife allocation. The livestock signature of years past was consistent with the sampling location being upstream from a large agricultural area. However, the decreased livestock allocation could be attributed to better fencing of streams, or a change in the type or concentration of antibiotics fed to livestock. Follow up studies should be done to indicate whether this signature has indeed faded. No human signature was classified above the MDP in any year.

Quarterly Site 14: Flat Branch

One of the highest human allocations was obtained at this Flat Branch location in Year 2. It was posited that this was due to a malfunctioning sewer line. Wildlife was the primary signature but did not cause as much concern as the human allocation. A secondary pet signature was also recorded. No major changes occurred in Year 3. A relatively high human allocation remained, and birds and wildlife comprised the major sources of contamination. No livestock signature was detected in the first two years.

The only surprise in the current project results was the classification of a minor livestock signature. Encouragingly, the human signature decreased below detectable limits. Pets was the highest allocation, and wildlife was minor. With the exception of the livestock signature, the results at Flat Branch are consistent with expectation. Several deer and unleashed dogs were observed on multiple occasions. The questionable livestock signature could have been a consequence of slight library misrepresentation. In the future, livestock isolates should be obtained from multiple locations around PWC. Collecting from different locations in the country should increase the potential for better representation.

Quarterly Site 15: South Run

This sampling location is in one of the more rural areas of the county, where large horse farms are a recurring theme. The wooded buffer zone supports a large deer and waterfowl population, and was consistent with the large wildlife signature detected in Year 2. Not surprisingly, livestock was secondary. The only notable change in Year 3 was the detection of a minor pet signature. The current project results suggest dogs are becoming a bigger issue, while wildlife and livestock persist at lower levels.

Quarterly Site 16 and Monthly Site 3: Broad Run

As with most commercial areas, fewer different sources were detected. Wildlife was detected in more than half of the Year 1 isolates, and pets was the only secondary signature above the MDP. In Year 2, the only change was a minor livestock signature disregarded as negligible. The pet signature increased but remained secondary to wildlife and birds. The current project results brought no major surprises. The pet and wildlife signatures remained primary. A questionable livestock signature was detected at a level just above the MDP, but should probably be disregarded. The ongoing construction and increased number of housing developments over the past couple years has forced wildlife into smaller green spaces. Additionally, the increasing human population suggests that dog waste is entering the waters of PWC at an increasing rate.

Monthly Site 3 is yet another example of an area where pet waste has become an issue. Compared to Quarterly Site 16, the minor livestock signature better fits the more rural surroundings of this location. The minor wildlife signature is consistent with the heavy wooded buffer zone. No human contamination was detected at either site, in any year.

Quarterly Site 17: Kettle Run

Three sources of impairment, wildlife, livestock, and pets, have characterized this location of Kettle Run for the past three years. The only difference from one year to the next lay in which source was primary. Wildlife was primary in Years 2 and 3 (when combined with birds), and livestock edged out the other two in the current project. The persistent livestock signature indicates that the area hobby farms are still active and are numerous enough to impact water quality. The pet signature can be explained by the presence of a community park, where walking and exercising of dogs is a frequent occurrence. The wildlife signature is typical of a rural setting. Future sampling will more than likely continue to demonstrate a three-source pattern, unless efforts are made to affect change.

Quarterly Site 18: North Fork of Lake Manassas

A beaver dam highlighted this site during the majority of samplings. No major surprises occurred. Although in a commercial part of PWC, a wildlife signature would still be expected due to the wooded surroundings of the stream. The wildlife signature was primary and was consistent with observations of deer, waterfowl, and deer scat. An almost equal percentage of isolates were classified as pets, a typical result for a commercial area. No human signature was detected. The absence of a livestock allocation is supported by the lack of nearby farms or agricultural areas.

Monthly Site 1: Little Bull Run

Human contamination was the only signature not classified at this location. Pet and wildlife signatures, both secondary, can be explained by the heavily forested, residential

framework. The only questionable allocation was livestock. Additional sampling should be done in order to assess whether livestock contamination is a legitimate concern.

IV. Fluorometric Analysis (Tables 14-34)

Fluorometric analysis (FA) was used to confirm ARA results that suggested the presence or absence of a human signature. As mentioned in Chapter III, a concentration value of 100 was used as the minimum threshold for an environmental sample to be recorded as positive for the presence of optical brighteners. The current project produced a total of 124 FA readings. Of the 124, eight were recorded above 100 (Tables 14-34). One of the nine values, the FA value of 120 in June 2007 at Quarterly Site 2, correlated with a human signature that was higher than the MDP. This value was deemed a true positive. The other eight values were deemed false positives, or numbers that suggested a human signature but was not confirmed by ARA. Conversely, the term false negative was used to define FA values that fell below 100 mg/L, but did not agree with ARA results that suggested a human signature above the MDP. Of the 116 values below 100 mg/L, three were recorded as false negatives. These were seen at Monthly Site 1 in June 2007, Monthly Site 2 in June 2007, and Quarterly Site 16 in September 2006. In total, the percentage of false positives was 5.6% while the percentage of false negatives was 2.4%. The overall rate of false positives was 6.7% while the overall rate of false negatives was 75%.

FA has been incorporated as part of the toolbox approach with this project since the spring of 2004. Although it saw limited use in Year 1 because it was not purchased until later, the fluorometer corroborated the human signatures in Cedar Run and Slate Run that occurred when testing was available. FA was applied to all samples in Year 2. It again corroborated the above-MDP human signature at Cedar Run during the same months that ARA identified one.

The only other site that year with a human signature, according to ARA, was Flat Branch. Over the course of Year 2, six of 13 months showed positive fluorescence readings. All six months correlated with ARA results that indicated at least a minor human signature. The months for which ARA did not show a human contribution were supported by negative fluorometry readings. In Year 3, two sampling locations showed monthly hotspots of human contamination. The yearly averages for human contamination for Quarterly Site 2 (Neabsco Creek) and a Slate Run location (not part of current project) were not above the MDP according to ARA, but they both had positive fluorometry readings in September 2005 and June 2006. The human signature at Slate Run was supported by ARA during the same months that the positive FA readings were obtained. The FA reading at Quarterly Site 2 in September 2005 was supported by ARA, but unfortunately, contamination resulted in the loss of ARA data for the June 2006 samples. Raw data from Year 3 was not available for calculation of the false positive and false negative rates for FA.

In summary, many if not all of the performance criteria mentioned in Chapter 1 were applied to this project. The **experimental design** included the well-researched selection of *E. coli* as the indicator bacterium. *E. coli* from known fecal sources (human, pets, livestock, wildlife) in the project watershed was used to create a known source library (KSL). The known source library was comparable in size to other successful MST projects, and was updated with fresh isolates. Additionally, the toolbox approach was employed through the use of fluorometric analysis as a confirmatory measure of human contamination. The **minimum detectable**

percentage, or sensitivity, was calculated based on methods published previously and was applied to analysis of the KSL. The KSL was **quantified** through the calculation of average rates of classification (ARCC), a method well established in the primary literature. Calculating the ARCC also allowed for the determination of the project **accuracy**. **Specificity** was measured by the calculation of the false positive and false negative rates with regard to the comparison between environmental samples and the FA results. The false positive rate was low but the false negative rate should be improved in future studies. **Range of applicability** is a term that has to be looked at on a broader scale. ARA has been applied in sub-tropical waters (8), marine waters (1), temperate climates (12-14) to name just a few. To date, there has been nothing to suggest that ARA can not be applied in any watershed under any standard conditions. The **practicality** of ARA is perhaps its biggest weapon. Capital costs are minimal, training can be accomplished in a week, and the per-sample cost is much lower than genotypic-based MST techniques.

V. Real Value of Analysis

The primary benefit of this project is that it provides locations and targets for reduction of *E. coli*. Because *E. coli* is an indicator organism, this in turn should lead to general reduction of waterborne disease. As pets and wildlife were the most commonly identified sources, the most resources and planning will be given to economically feasible methods of reducing these sources. Directed action will be taken by officials of Prince William County (PWC) and will allow consideration to be given by the state to enforcement of water quality standards in the impacted areas. The impact of the research on the TMDL and TMDL reduction plan are determined by levels of *E. coli* reported and by the sources of fecal pollution. Pet sources were found in a wide part of the region. As a primary source at many locations, this project provides the basis for

increased bagging of dog feces and increased enforcement of this practice. As housing increases in the Occoquan Basin, the reduction of potential pet fecal influence will become more important for meeting bacterial water quality standards.

The more often secondary signature of livestock will provide PWC with minor initiative to reduce its fecal influence in the affected areas. Livestock farm targets near the sampling regions will be required to implement best management practices (BMPs) for runoff reduction. Possible approaches include fencing of creeks and creation of separate watering holes. Horse trails in the region could either be diverted away from waterways, or additional storm water management ponds could be added to the area. In areas where such changes are not possible, diapering may be appropriate for horses.

VI. Conclusions

One goal of this project was to monitor and evaluate the identification of 21 stream locations as *E. coli* impaired. This goal was successfully completed. This objective also included determining the source(s) of bacterial impairment. This objective was satisfactorily met, but more work is needed to either confirm or deny the results at several locations. Wildlife, pets, and livestock were indicated as the major and minor source of impairment by ARA. Yearly site averages indicated no human signatures at any site.

A second objective was to determine the best design for an ARA library. This objective was completed in part, because the extent which older isolates and those from outside the region can be used was not determined quantitatively. The best ARA library design was the one that used only unique isolates and removed conflicting isolates. Continuing examination of the

representation of library data as binary is necessary to determine whether the statistical assumptions in DA prevent meaningful results.

Local libraries must remain dominant but regional information is useful in filling in gaps. The multi-year library created for this study may have contained regional or dated data, which inflated the variation and importance of the one or more known source classes. This study cannot confirm the use of library data from both an entire region and multiple years. The third objective for the study was to evaluate the fluorometer for measurement of optical brighteners in fresh water. This was partially successful due to a low rate of false positives but a questionably higher rate of false negatives. The fluorometer continues to have potential as a metric of waste in freshwater. More work must be done to show its utility.

VII. Study Revision Recommendations

Several measures could be taken by future researchers to improve the results of this project. It might be beneficial to revert back to a five category classification for the KSL, as was used in some of the earlier years. This would allow a more streamlined comparison and may enable better specificity with regard to the difference between birds and other wildlife, such as deer. The KSL may also be enhanced by the development and implementation of a more concrete and uniform method of plate reading. Although a strict guideline was followed for this project, it remains to be seen whether the next researcher will follow the same protocol. Having a document that details, in clear and distinct language, what makes a particular isolate be recorded as resistant or not resistant may go a long way in helping to pass along the technique to new researchers. Another approach is to incorporate the use of an automated plate reader. The benefit of such a device would certainly outweigh the increased capital expense. Similarly, it

must be kept in mind that the type and/or concentration of antibiotics might have to be adjusted. The biggest problem with changing antibiotics is that it would make all prior libraries for this project useless. But starting over with fresh known source isolates might be the best approach.

With regard to the collection of the environmental samples, a couple approaches might enhance results. Although stream flow was observed qualitatively, a more quantitative measure of stream flow rate might improve knowledge about the limitations of ARA. Collecting stream flow data was not an objective at the beginning of this project, and the lack of materials and resources prevented it from happening at a later date. Finally, the performance criterion known as repeatability was not properly emphasized in this project. In the future, efforts should be made to collect duplicate environmental samples in order to improve comparison of ARA results.

VIII. References

1. **Dickerson, J. W. J., C. Hagedorn, and A. Hassall.** 2007. Detection and remediation of human-origin pollution at two public beaches in Virginia using multiple source tracking methods. Water Res. **(accepted)**.
2. **Graves, A. K., C. Hagedorn, A. Brooks, R. L. Hagedorn, and E. Martin.** Microbial source tracking in a ruralwatershed dominated by cattle,Water Research (2007), doi:10.1016/j.watres.2007.04.020.
3. **Hagedorn, C.** 2005. Identifying sources of fecal pollution in impaired waters in Prince William County, Virginia, p. 1-45. Virginia Polytechnic Institute and State University.
4. **Hagedorn, C.** 2005. Identifying sources of fecal pollution in impaired waters in Prince William County, Virginia: sites monitored quarterly in 2004-2005, p. 1-4. Virginia Polytechnic Institute and State University.
5. **Hagedorn, C.** 2006. Identifying sources of fecal pollution in impaired waters in Prince William County, Virginia: sites monitored quarterly in 2005-2006, p. 1-8. Virginia Polytechnic Institute and State University.
6. **Hagedorn, C.** 2004. Updated final project report to Department of Public Works, Prince William County, VA, p. 1-29. Virginia Polytechnic Institute and State University.
7. **Hagedorn, C., S. L. Robinson, J. R. Filtz, S. B. Grubbs, T. A. Angier, and J. R. B. Reneau.** 1999. Determining sources of fecal pollution in a rural Virginia watershed with antibiotic resistance patterns in fecal streptococci. Appl. Environ. Microbiol. **65:**5522-5531.
8. **Harwood, V. J., J. Whitlock, and V. Withington.** 2000. Classification of antibiotic resistance patterns of indicator bacteria by discriminant analysis: use in predicting the

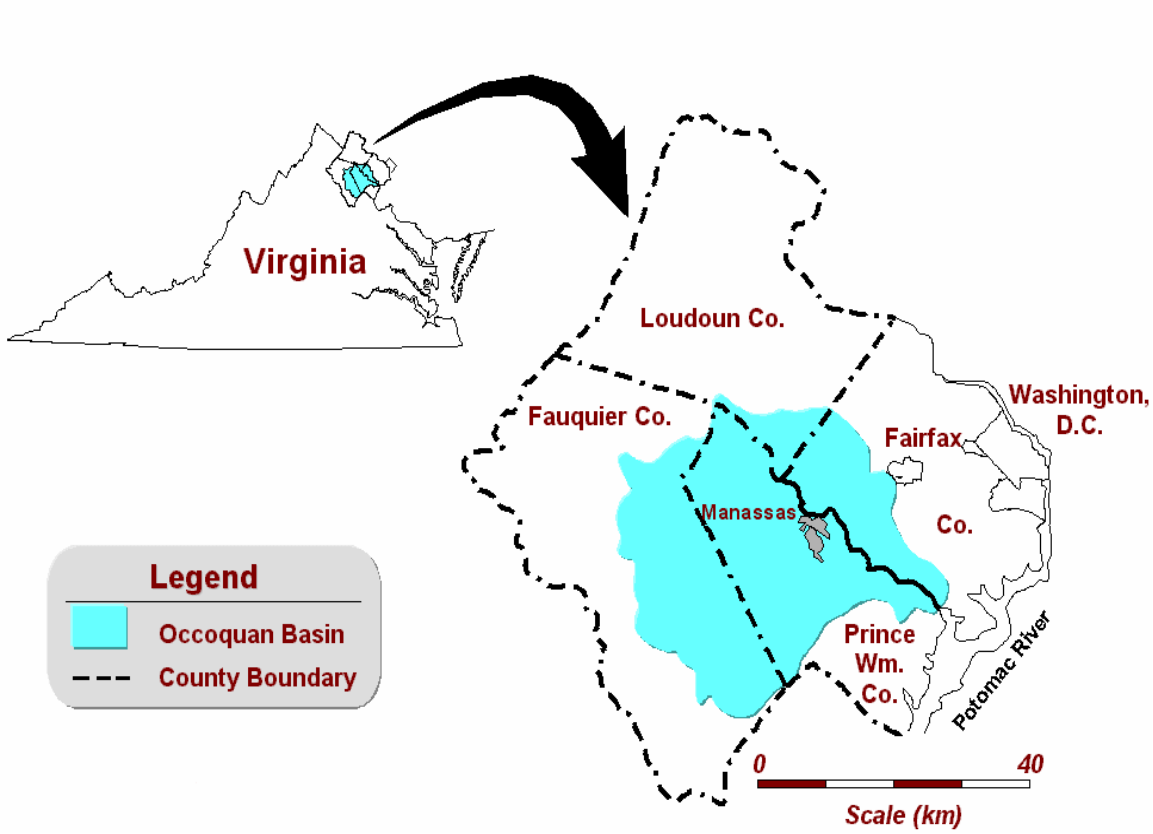
source of fecal contamination in subtropical waters. Appl. Environ. Microbiol. **66:**3698-3704.

9. **Stoeckel, D. M., and V. J. Harwood.** 2007. Performance, design, and analysis in microbial source tracking studies. Appl. Environ. Microbiol. **73:**2405-2415.
10. **Touchton, G. D.** 2005. Microbial source tracking in a mixed use watershed in Northern Virginia. M.S. Thesis. Virginia Polytechnic Institute and State University, Blacksburg, VA.
11. **USEPA.** 1986. Ambient water quality criteria for bacteria - 1986, p. 1-24.
12. **Wiggins, B. A.** 1996. Discriminant analysis of antibiotic resistance patterns in fecal streptococci, a method to differentiate human and animal sources of fecal pollution in natural waters. Appl. Environ. Microbiol. **62:**3997-4002.
13. **Wiggins, B. A., R. W. Andrews, R. A. Conway, C. L. Corr, E. J. Dobratz, D. P. Dougherty, J. R. Eppard, S. R. Knupp, M. C. Limjoco, J. M. Mettenberg, J. M. Rinehardt, J. Sonsino, R. L. Torrijos, and M. E. Zimmerman.** 1999. Use of antibiotic resistance analysis to identify nonpoint sources of fecal pollution. Appl. Environ. Microbiol. **65:**3483-3486.
14. **Wiggins, B. A., P. W. Cash, W. S. Creamer, S. E. Dart, P. P. Garcia, T. M. Gerecke, J. Han, B. L. Henry, K. B. Hoover, E. L. Johnson, K. C. Jones, J. G. McCarthy, J. A. McDonough, S. A. Mercer, M. J. Noto, H. Park, M. S. Phillips, S. M. Purner, B. Smith, E. N. Stevens, and A. K. Varner.** 2003. Use of antibiotic resistance analysis for representativeness testing of multiwatershed libraries. Appl. Environ. Microbiol. **69:**3399-3405.

Appendix

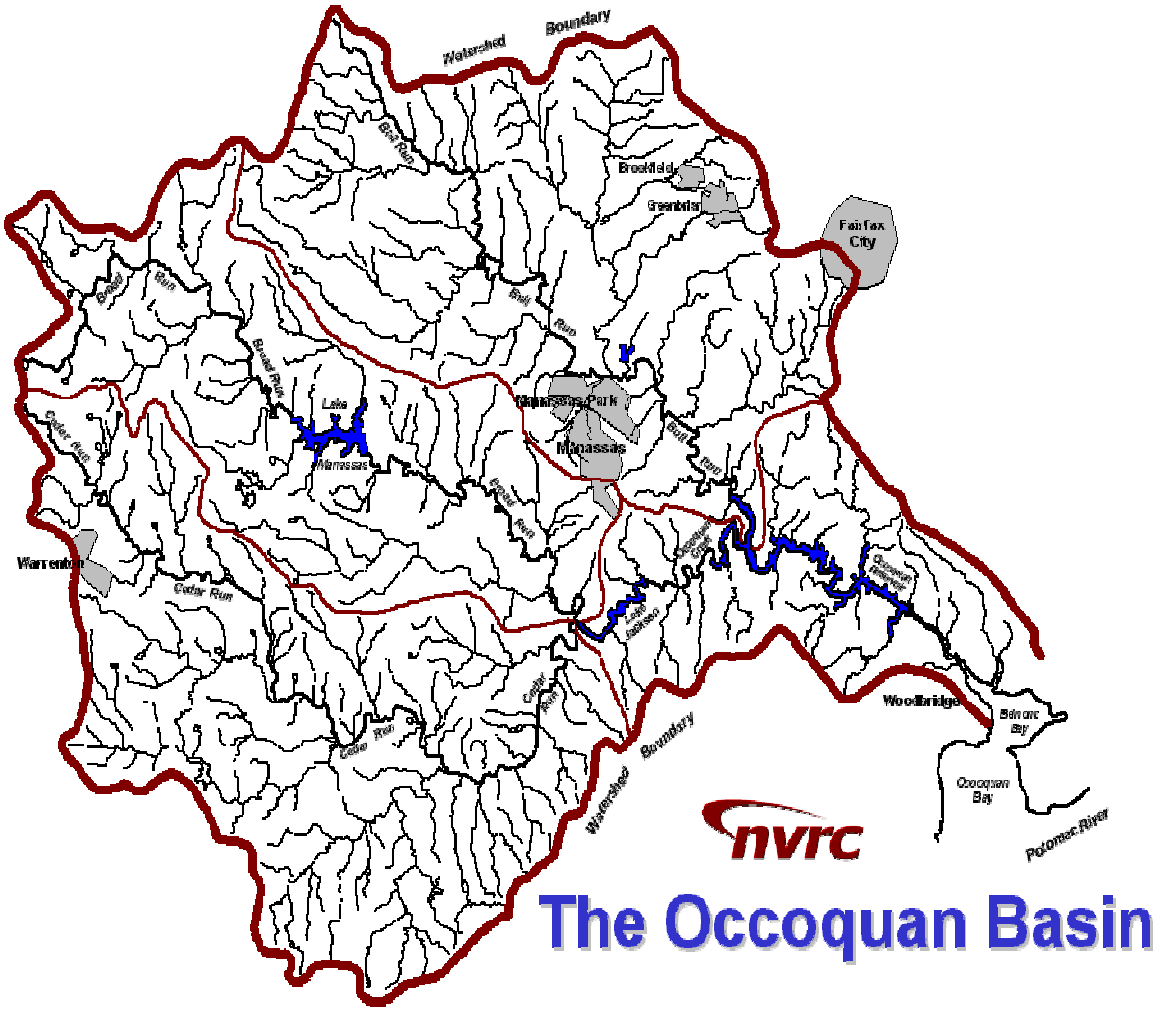
A. Occoquan Basin Maps

Figure 5. The Occoquan Basin of Virginia



Courtesy of the Northern Virginia Regional Commission – www.novaregion.org

Figure 6. Stream Locations within the Occoquan Basin



Courtesy of the Northern Virginia Regional Commission – www.novaregion.org

B. Source Tracking Site Data

Table 14. Monitoring and MST Results for Site M1: Little Bull Run, Old Carolina Rd.

Site	N38°49.2758' W077°37.8274'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL**[a]	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06[b]	---	---	---	---	---	---	---
July	315	23	73	0	10 (44)	7 (30)	6 (26)
Aug	3300	21	145	0	8 (38)	11 (52)	2 (10)
Sept	173	24	59	1 (4)	17 (71)	4 (17)	2 (8)
Oct	395	24	103	0	12 (50)	0	12 (50)
Nov	295	24	79	2 (8)	6 (25)	7 (29)	9 (38)
Dec	85	24	47	3 (13)	17 (71)	2 (8)	2 (8)
Jan 07	9	22	39	0	21 (95)	0	1 (5)
Feb	86	24	31	0	20 (83)	0	4 (17)
Mar[c]	5	---	20	---	---	---	---
Apr	25	12	47	0	12 (50)	5 (21)	7 (29)
May	44	24	47	0	5 (21)	2 (8)	17 (71)
June	72	24	67	12 (50)	4 (17)	8 (33)	0
Total	---	246	---	18 (7)	132 (54)	46 (19)	62 (25)
Average	400	22	63	1.63 (7)	12 (51)	4 (18)	5.6 (24)
Std. Dev.	923	---	3.59	3.59 (15)	6 (25.9)	3.73 (27.5)	5.3 (21.9)

[a] Colony Forming Units per 100mL

[b] Sampling of monthly sites began in July 06

[c] Insufficient number of isolates for MST

Table 15. Monitoring and MST Results for Site M2: Powell's Creek, Northgate Dr.

Site	N38°36.4787' W077°19.9019'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL**[a]	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06[b]	---	---	---	---	---	---	---
July	78	24	59	2 (8)	15 (63)	2 (8)	5 (21)
Aug	160	23	49	0	14 (61)	2 (9)	7 (30)
Sept	143	23	42	1 (4)	9 (39)	6 (26)	7 (30)
Oct	232	24	83	0	15 (63)	5 (21)	4 (17)
Nov	90	24	69	1 (4)	5 (21)	11 (46)	7 (29)
Dec	475	23	71	0	14 (61)	5 (22)	4 (17)
Jan 07	19	24	68	0	24 (100)	0	0
Feb	154	24	58	0	24 (100)	0	0
Mar	23	15	36	0	1 (7)	4 (27)	10 (67)
Apr	14	19	61	0	8 (42)	3 (16)	8 (42)
May	64	24	51	1 (4)	0	3 (13)	20 (83)
June	210	24	47	4 (17)	2 (8)	15 (63)	3 (13)
Total	---	271	---	9 (3.32)	131 (48)	56 (20.66)	75 (27.7)
Average	163.5	22.58	57.83	0.75 (3.08)	10.91 (47)	4.67 (20.9)	6.3 (29.1)
Std. Dev.	167.3	---	13.54	1.22 (5.11)	8.2 (33.9)	4.4 (18.43)	5.3 (24.9)

[a] Colony Forming Units per 100mL

[b] Sampling of monthly sites began in July 06

Table 16. Monitoring and MST Results for Site M3: Broad Run, Rt. 55

Site	55 N38°49.3822' W077°42.3242'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL[a]**	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06[b]	---	---	---	---	---	---	---
July[c]	105	---	39	---	---	---	---
Aug	38	17	42	0	13 (76)	2 (12)	2 (12)
Sept	55	24	33	0	8 (33)	6 (25)	10 (42)
Oct	245	24	49	0	7 (29)	7 (29)	10 (42)
Nov	1100	24	61	1 (4)	6 (25)	9 (38)	8 (33)
Dec	53	24	20	0	17 (71)	3 (13)	4 (17)
Jan 07	10	17	19	2 (12)	15 (88)	0	0
Feb	27	24	13	0	9 (38)	5 (21)	10 (42)
Mar	157	20	8	0	14 (70)	1 (5)	5 (25)
Apr	123	23	23	0	12 (52)	8 (35)	3 (13)
May	123	24	26	0	17 (71)	3 (13)	4 (17)
June	94	24	48	1 (4)	7 (29)	16 (67)	0
Total	---	245	---	4 (1.6)	125 (51)	60 (24.5)	56 (22.9)
Average	177.5	20.4	31.8	0.36 (1.8)	11.4 (52.9)	5.4 (23.5)	5.1 (22.1)
Std. Dev.	297.7	---	16.2	0.67 (3.7)	4.1 (22.9)	4.5 (18.8)	3.9 (15.9)

[a] Colony Forming Units per 100mL

[b] Sampling of monthly sites began in July 06

[c] Insufficient number of isolates for MST

Table 17. Monitoring and MST Results for Site Q1: Neabsco Creek, Lindendale Rd.

Site	N38°38.7273' W077°21.9542'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL[a]**	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	1600	21	109	1 (4)	18 (86)	0	2 (10)
Sept	353	24	32	0	13 (54)	0	11 (46)
Dec	24	18	25	0	13 (72)	0	5 (28)
Mar 07	42	24	9	0	15 (63)	0	9 (38)
June	650	24	61	0	8 (33)	11 (46)	5 (21)
Total	---	111	---	1 (0.9)	67 (60.3)	11 (9.9)	32 (28.9)
Average	533.8	22.2	47.2	0.2 (0.8)	13.4 (61.6)	2.2 (9.2)	6.4 (28.6)
Std. Dev.	649	---	39.3	0.45 (1.8)	3.65 (19.9)	4.9 (20.6)	3.6 (14.1)

[a] Colony Forming Units per 100mL

Table 18. Monitoring and MST Results for Site Q2: Neabsco Creek, Benita Fitzgerald Rd.

Site	N38°37.5141' W077°18.8082'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL[a]**	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06[b]	2960	---	110	---	---	---	---
Sept	275	24	77	0	5 (21)	7 (29)	12 (50)
Dec	103	24	75	0	17 (71)	0	7 (29)
Mar 07	36	18	26	0	13 (72)	0	5 (28)
June	1000	24	120	8 (33)	7 (29)	9 (38)	0
Total	---	90	---	8 (8.9)	42 (46.7)	16 (17.8)	24 (26.7)
Average	874.8	22.5	81.6	2 (8.3)	10.5 (48.3)	4 (16.8)	6 (26.8)
Std. Dev.	1227.1	---	36.9	4 (16.5)	5.5 (23.4)	4.7 (19.7)	5 (20.5)

[a] Colony Forming Units per 100mL

[b] MST results lost due to contamination

Table 19. Monitoring and MST Results for Site Q3: Neabsco Creek, Neabsco Mills Rd. & Rt. 1

Site	N38°36.6421' W077°17.4307'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL**[a]	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	1140	16	94	0	13 (81)	1 (6)	2 (13)
Sept	60	22	80	0	7 (32)	0	15 (68)
Dec	33	24	74	0	17 (71)	0	7 (29)
Mar 07	38	19	29	0	7 (37)	0	12 (63)
June	140	24	123	0	1 (4)	21 (88)	2 (8)
Total	---	105	---	0	45 (42.9)	22 (21)	38 (36.2)
Average	282.2	21	80	0	9 (45)	4.4 (18.8)	7.6 (36.2)
Std. Dev.	481.4	---	34.2	0	6.2 (31.2)	9.3 (38.8)	5.9 (27.9)

[a] Colony Forming Units per 100mL

Table 20. Monitoring and MST Results for Site Q4: Cow Branch, Montgomery Ave.

Site	N38°38.1860' W077°16.6908'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL**[a]	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06[b]	---	---	---	---	---	---	---
Sept	135	24	18	0	8 (33)	9 (38)	7 (29)
Dec[c]	<10	---	20	---	---	---	---
Mar 07[c]	0	---	6	---	---	---	---
June	800	24	93	0	10 (42)	9 (38)	5 (21)
Total	---	48	---	0	18 (37.5)	18 (37.5)	12 (25)
Average	236	24	34.3	0	9 (37.5)	9 (37.5)	6 (25)
Std. Dev.	381	---	39.7	0	1.4 (6.4)	1.4 (6.4)	1.4 (5.7)

[a] Colony Forming Units per 100mL

[b] Sampling began in Sept. 06

[c] Insufficient number of isolates for MST

Table 21. Monitoring and MST Results for Site Q5: Cow Branch, Rippon Landing Park

Site	N38°37.0715' W077°16.4505'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL**[a]	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	320	13	53	0	5 (38)	5 (38)	3 (23)
Sept	45	24	35	0	14 (58)	3 (13)	7 (29)
Dec[b]	<10	---	17	---	---	---	---
Mar 07[b]	7	---	5	---	---	---	---
June	620	23	32	1 (4)	18 (78)	2 (9)	2 (9)
Total	---	60	---	1 (1.7)	37 (61.7)	10 (16.7)	12 (20)
Average	200.2	20	28.4	0.3 (1.3)	12.3 (58)	3.3 (20)	4 (20.3)
Std. Dev.	268.6	---	18.3	0.58 (2.3)	6.7 (20)	1.5 (15.7)	2.6 (10.3)

[a] Colony Forming Units per 100mL

[b] Insufficient number of isolates for MST

Table 22. Monitoring and MST Results for Site Q6: Powell's Creek, Fox Mills Apt. & Rt. 1

Site	N38°35.7780' W077°18.1155'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL**[a]	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	220	13	47	0	4 (31)	7 (54)	2 (15)
Sept	318	26	42	0	0	12 (46)	14 (54)
Dec	78	24	60	0	16 (67)	5 (21)	3 (13)
Mar 07	18	12	30	0	1 (8)	2 (17)	9 (75)
June	108	24	54	2 (8)	10 (42)	9 (38)	3 (13)
Total	---	99	---	2 (2)	31 (31.3)	35 (35.4)	31 (31.3)
Average	148.4	19.8	46.6	0.4 (1.6)	6.2 (29.6)	7 (35.2)	6.2 (34)
Std. Dev.	119.9	---	11.5	0.89 (3.6)	6.7 (26.9)	3.8 (15.9)	5.2 (28.8)

[a] Colony Forming Units per 100mL

Table 23. Monitoring and MST Results for Site Q7: Quantico Creek, South Fork at Joplin Rd.

Site	N38°35.2507' W077°25.7305'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL**[a]	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	40	22	55	0	9 (41)	5 (22)	8 (36)
Sept	113	22	47	0	6 (27)	7 (32)	9 (41)
Dec	<10	15	26	0	10 (67)	2 (13)	3 (20)
Mar 07	19	22	12	0	6 (27)	0	16 (73)
June	134	23	66	0	1 (4)	20 (87)	2 (9)
Total	---	104	---	0	32 (30.8)	34 (32.7)	38 (36.5)
Average	63	20.8	41.2	0	6.4 (33.2)	6.8 (30.8)	7.6 (35.8)
Std. Dev.	56.8	---	21.9	0	3.5 (23.1)	7.9 (33.6)	5.6 (24.4)

[a] Colony Forming Units per 100mL

Table 24. Monitoring and MST Results for Site Q8: Quantico Creek, Main Stem

Site	N38°34.1153' W077°20.1667'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL**[a]	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	160	21	38	1 (4)	11 (52)	4 (19)	5 (24)
Sept	128	24	41	0	2 (8)	16 (67)	6 (25)
Dec[b]	<10	---	20	---	---	---	---
Mar 07	13	20	9	0	6 (29)	7 (33)	8 (38)
June	88	24	40	0	22 (92)	2 (8)	0
Total	---	89	---	1 (1.1)	41 (46.1)	29 (32.6)	19 (21.3)
Average	79.6	22.3	29.6	0.25 (1)	10.3 (45.3)	7.3 (31.8)	4.75 (21.8)
Std. Dev.	67.6	---	14.4	0.5 (2)	8.7 (36)	6.2 (25.6)	3.4 (15.8)

[a] Colony Forming Units per 100mL
[b] Insufficient number of isolates for MST

Table 25. Monitoring and MST Results for Site Q9: Cedar Run, Carriage Ford Rd.

Site	N38°38.5709' W077°35.1393'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL[a]**	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	40	21	56	1 (5)	11 (52)	4 (19)	5 (24)
Sept	63	23	88	0	1 (4)	15 (65)	7 (30)
Dec	430	24	41	0	14 (58)	1 (4)	9 (38)
Mar 07	30	21	17	0	6 (29)	7 (33)	8 (38)
June	48	21	82	0	13 (62)	6 (29)	2 (10)
Total	---	110	---	1 (0.9)	45 (40.9)	33 (30)	31 (28.2)
Average	122.2	22	56.8	0.2 (1)	9 (41)	6.6 (30)	6.2 (28)
Std. Dev.	172.5	---	29.3	0.5 (2.2)	5.4 (24.3)	5.2 (22.5)	2.8 (11.7)

[a] Colony Forming Units per 100mL

Table 26. Monitoring and MST Results for Site Q10: Cedar Run, Bristow Rd.

Site	N38°41.2152' W077°29.4485'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL[a]**	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	60	12	56	0	11 (92)	1 (8)	0
Sept	85	21	99	0	14 (67)	0	7 (33)
Dec	36	23	40	2 (9)	18 (78)	2 (9)	1 (4)
Mar 07	7	12	18	0	2 (17)	0	10 (83)
June	11	15	78	0	4 (27)	3 (20)	8 (53)
Total	---	83	---	2 (2.4)	49 (59)	6 (7.2)	26 (31.3)
Average	39.8	16.6	58.2	0.4 (1.8)	9.8 (56.2)	1.2 (7.4)	5.2 (34.6)
Std. Dev.	33.1	---	31.7	0.9 (4)	6.7 (32.6)	1.3 (8.2)	4.4 (34.7)

[a] Colony Forming Units per 100mL

Table 27. Monitoring and MST Results for Site Q11: Slate Run, Old Church Rd.

Site	N38°40.5688' W077°30.5324'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL**[a]	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	200	24	93	0	12 (50)	8 (33)	4 (17)
Sept[b]	68	---	104	---	---	---	---
Dec	245	24	55	0	6 (25)	11 (45)	7 (29)
Mar 07	18	19	24	1 (5.26)	2 (11)	4 (21)	12 (63)
June	101	20	89	0	8 (40)	10 (50)	2 (10)
Total	---	87	---	1 (1.1)	28 (32.2)	33 (37.9)	25 (28.7)
Average	126.4	21.8	73	0.3 (1.3)	7 (31.5)	8.3 (37.3)	6.3 (29.8)
Std. Dev.	93.9	---	32.9	0.5 (2.5)	4.2 (17)	3.1 (13)	4.3 (23.5)

[a] Colony Forming Units per 100mL
[b] Insufficient number of isolates for MST

Table 28. Monitoring and MST Results for Site Q12: Bull Run, Rt. 28

Site	N38°48.1749' W077°26.9728'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL**[a]	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	140	13	37	1 (8)	10 (77)	1 (8)	1 (8)
Sept	38	22	52	0	13 (59)	3 (14)	6 (27)
Dec	32	18	34	1 (6)	10 (56)	2 (11)	5 (28)
Mar 07[b]	1	---	15	---	---	---	---
June	183	23	51	0	11 (48)	6 (26)	6 (26)
Total	---	76	---	2 (2.6)	44 (57.9)	12 (15.8)	18 (23.7)
Average	78.8	19	37.8	0.5 (3.5)	11 (60)	3 (14.8)	4.5 (22.3)
Std. Dev.	78.3	---	15.1	0.6 (4.1)	1.4 (12.2)	2.2 (7.9)	2.4 (9.5)

[a] Colony Forming Units per 100mL
[b] Insufficient number of isolates for MST

Table 29. Monitoring and MST Results for Site Q13: Catharpin Run, Robin Dr.

Site	N38°50.6622' W077°32.8825'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL**[a]	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	85	17	65	0	14 (82)	2 (12)	1 (6)
Sept	225	24	62	0	9 (38)	5 (21)	10 (42)
Dec	100	19	33	0	9 (47)	0	10 (53)
Mar 07[b]	0	---	13	---	---	---	---
June	135	24	72	0	18 (75)	5 (21)	1 (4)
Total	---	84	---	0	50 (59.5)	12 (14.2)	22 (26.2)
Average	109	21	49	0	12.5 (60.5)	3 (13.5)	5.5 (26.3)
Std. Dev.	81.7	---	25	0	4.4 (21.3)	2.4 (9.9)	5.2 (25)

[a] Colony Forming Units per 100mL
[b] Insufficient number of isolates for MST

Table 30. Monitoring and MST Results for Site Q14: Flat Branch, Lomond Dr.

Site	N38°46.9103' W077°29.2204'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL**[a]	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	15200	21	151	1 (5)	6 (29)	8 (38)	6 (29)
Sept	5	23	49	0	3 (13)	17 (74)	3 (13)
Dec	23	16	33	1 (6)	7 (44)	0	8 (50)
Mar 07[b]	23	---	12	---	---	---	---
June	1840	23	133	0	15 (65)	6 (26)	8 (9)
Total	---	83	---	2 (2.4)	31 (37.3)	31 (37.3)	25 (30.1)
Average	3430	20.8	75.6	0.5 (2.8)	7.8 (37.8)	7.8 (34.5)	6.3 (25.3)
Std. Dev.	6626	---	62.3	0.6 (3.2)	5.1 (22.1)	7 (30.7)	2.4 (18.6)

[a] Colony Forming Units per 100mL
[b] Insufficient number of isolates for MST

Table 31. Monitoring and MST Results for Site Q15: South Run, Buckland Mill Rd.

Site	N38°46.1814' W077°39.9624'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL[a]**	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	45	21	47	1 (5)	9 (43)	11 (52)	0
Sept	43	22	38	0	16 (73)	1 (5)	5 (23)
Dec	38	21	52	0	7 (33)	1 (5)	13 (62)
Mar 07[b]	7	---	36	---	---	---	---
June[b]	15	---	64	---	---	---	---
Total	---	64	---	1 (1.6)	32 (50)	13 (20.3)	18 (28.1)
Average	29.6	21.3	47.4	0.3 (1.7)	10.7 (49.7)	4.3 (20.7)	6 (28.3)
Std. Dev.	17.4	---	11.3	0.6 (2.9)	4.7 (20.8)	5.7 (27.1)	6.6 (31.3)

[a] Colony Forming Units per 100mL
[b] Insufficient number of isolates for MST

Table 32. Monitoring and MST Results for Site Q16: Broad Run, Rt. 28

Site	N38°44.1842' W077°32.0215'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL[a]**	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	85	24	42	0	15 (62)	4 (17)	5 (21)
Sept	45	24	76	6 (25)	8 (33)	1 (4)	9 (38)
Dec	43	21	53	0	13 (62)	2 (10)	6 (29)
Mar 07	7	15	24	0	3 (20)	0	12 (80)
June	210	24	65	0	6 (25)	15 (63)	3 (13)
Total	---	108	---	6 (5.6)	45 (41.7)	22 (20.4)	35 (32.4)
Average	78	21.6	52	1.2 (5)	9 (40.4)	4.4 (18.8)	7 (36.2)
Std. Dev.	78.8	---	20.2	2.7 (11.2)	4.9 (20.3)	6.1 (25.5)	3.5 (26.2)

[a] Colony Forming Units per 100mL

Table 33. Monitoring and MST Results for Site Q17: Kettle Run, Valley View Rd.

Site	N38°42.1771' W077°32.0032'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL**[a]	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06	260	24	69	1 (4)	6 (25)	16 (67)	1 (4)
Sept	230	24	69	1 (4)	5 (21)	14 (58)	4 (17)
Dec	133	24	74	0	15 (63)	4 (17)	5 (21)
Mar 07[b]	6	---	41	---	---	---	---
June	150	18	17	0	7 (39)	2 (11)	9 (50)
Total	---	90	---	2 (2.2)	33 (30.6)	36 (33.3)	19 (17.6)
Average	155.8	22.5	54	0.5 (2)	8.3 (37)	9 (38.3)	4.8 (23)
Std. Dev.	99.2	---	24.4	0.6 (2.3)	4.6 (19)	7 (28.3)	3.3 (19.4)

[a] Colony Forming Units per 100mL

[b] Insufficient number of isolates for MST

Table 34. Monitoring and MST Results for Site Q18: North Fork of Lake Manassas, Rt. 29 S

Site	N38°47.5768' W077°37.5020'			Number of *E. coli* Isolates Classified (%)			
Month	**CFU/100mL**[a]	**Isolates**	**Fluorometry**	**Human**	**Pet**	**Livestock**	**Wildlife**
June 06[b]	---	---	---	---	---	---	---
Sept	100	20	74	2 (10)	5 (25)	4 (20)	9 (45)
Dec	73	21	59	0	14 (67)	0	7 (33)
Mar 07[b]	23	---	25	---	---	---	---
June	105	22	81	0	8 (36)	0	14 (64)
Total	---	63	---	2 (3.2)	27 (42.9)	4 (6.3)	30 (47.6)
Average	75.3	21	66.5	0.7 (3.3)	9 (42.7)	1.3 (6.7)	10 (47.3)
Std. Dev.	37.6	---	13.3	1.2 (5.8)	4.6 (21.8)	2.3 (11.5)	3.6 (15.6)

[a] Colony Forming Units per 100mL

[b] Sampling began in Sept. 06

C. Vita

Tim Wade was born in Harrisonburg, VA on May 4, 1982. As a junior and senior at Fort Defiance High School, he took classes at the Shenandoah Valley Regional Governor's School, a program designed for gifted, talented and highly motivated students. He graduated with an Associate of Arts and Sciences degree from Blue Ridge Community College (BRCC) in Weyers Cave, Va. While at BRCC, he was actively involved with virtually every student organization, including acting as a Student Ambassador, and participating in the Student Government Association and the Phi Theta Kappa National Honor Society. He transferred to James Madison University (JMU) where he completed a Bachelor of Science degree in Biology. Similar to his time at BRCC, Tim was very active with the student body of JMU. His extracurricular involvement included the Alpha Epsilon Delta Pre-Professional Honor Society, Pre-Dental Society (charter member), Pre-Physical Therapy Club, JMU Club Tennis, and acting as a volunteer pharmacy technician for the Harrisonburg Free Clinic. As a graduate student at Virginia Tech he acted as a graduate research assistant and graduate teaching assistant. He became a member of the American Society for Microbiology, the Phi Sigma National Biological Sciences Honor Society (Alpha Psi Chapter), Virginia Tech Club Tennis, and participated as a department representative for the Graduate Student Assembly and a judicial representative for the Graduate Honor System. Of all the activities he has been involved with and accomplishments he has garnered, Tim is most proud of earning the rank of Eagle Scout while a member of his local Boy Scouts of America troop.