DISSERTATION

SOURCE-TRACKING OF ANTIBIOTIC RESISTANCE GENES IN THE WATERSHED USING MOLECULAR PROFILING AND GEOSPATIAL ANALYSES

Submitted by

Heather Storteboom

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2009

COLORADO STATE UNIVERSITY

July 28, 2009

WE HEARBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY HEATHER STORTEBOOM ENTITLED, "SOURCE-TRACKING OF ANTIBIOTIC RESISTANCE GENES IN THE WATERSHED USING MOLECULAR PROFILING AND GEOSPATIAL ANALYSES," BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Jessica Davis

Amy Pruden

Mary Stromberger

Advisor: Mazdak Arabi

Department Head: Luis Garcia

ABSTRACT OF DISSERTATION

SOURCE-TRACKING OF ANTIBIOTIC RESISTANCE GENES IN THE WATERSHED USING MOLECULAR PROFILING AND GEOSPATIAL ANALYSES

Antibiotic resistance genes (ARG) have been found in many environmental matrices, including soils, groundwater, surface water, and sediments. Agricultural feeding operations and wastewater treatment plants are potential sources of ARG in rivers, or are sources of antibiotics that may select ARG from native river bacteria. The aim of this research is to identify ARG profiles that can characterize potential sources of ARG as well as native river environments and then use this knowledge to determine the sources and mechanisms involved in the spread of ARG to river environments. Initially, three wastewater treatment plants, six animal feeding operation lagoons, three sites along a pristine region of the Cache la Poudre River (PR), and a wildlife fish hatchery and rearing unit were compared with respect to the distribution, levels, and phylogenetic diversity of their ARG profiles. The tet genes tet(H), tet(Q), tet(S), and tet(T) were found to indicate agricultural influence, while high detection frequencies of tet(C), tet(E), and tet(O) were more typical of WWTP profiles. Sul(I) was detected in 100% of samples from source environments, but just once in the pristine river environment. The ARG profile of the pristine PR was dominated by tet(M) and tet(W), demonstrating their presence in an environment does not indicate anthropogenic disturbance. The tet(W) clone libraries from Pristine PR, WWTPs, and AFO lagoons, are each unique, as determined by both restriction fragment length polymorphism (RFLP) and phylogenetic analysis.

Secondly, samples from the PR and South Platte River (SPR) in Northern Colorado were characterized with respect to the distribution, levels, and diversity of their ARG profiles. On the basis of the ARG indicator variables derived in the study of source environments, most river samples were classified as WWTP influenced by discriminant analysis. The relationship between spatial explanatory variables and the ARG response variables was determined with classification and regression tree (CART) analysis. There was good agreement between the classification of river sites according to spatial variables and source indicator variables, demonstrating the effectiveness of these indicators in source-tracking ARG. According to multivariate linear regression, sul(I) was significantly correlated with the inverse-distance weighted (IDW) number of cattle upstream of each river point (R^2 = 0.83, p<0.0003), whereas tet(W) was not correlated with any explanatory variable tested. Tet(W) was isolated from two river environments: site PR4, located in Weld County downstream of Fort Collins; and site SPR3, located downstream of the confluence of the PR with the SPR. When compared to an existing clone library of tet(W) genes from animal feeding operations and wastewater treatment plants, PR4 was significantly different from the animal feeding operations (p<0.05); the SPR confluence (SPR3) was not significantly different from either environment. The PR4 environment was most similar to that of wastewater treatment plants, while SPR3 showed equal similarity with both source environments. A link between ARG indicator variables and spatial indicators was established. Furthermore, it was demonstrated that the ARG profiles of river samples were more similar to WWTPs than AFO lagoons or the pristine river. Based on this work, transport of ARG from sources may be a reasonable mechanism for ARG proliferation in riverine environments.

Heather Storteboom
Department of Civil and Environmental Engineering
Colorado State University
Fort Collins, CO 80523
Fall 2009

ACKNOWLEDGEMENTS

The time and energy devoted to this task has not only resulted in sacrifices for me but also for my family and friends. I want to thank all of you for putting up with me during the ups and downs of the last five years of graduate school. I would especially like to recognize the following:

To Dr. Amy Pruden: thank you for giving me a chance five years ago and encouraging me and helping me to grow along the way. I cannot thank you enough for all the opportunities you have given me. I continue to miss you and wish you success in Virginia.

To Dr. Mazdak Arabi: thank you for taking me in as a graduate student. I appreciate all the help and resources you provided to make my staying at CSU possible. Thanks for all your contributions to the analysis and writing of this dissertation. I am grateful for the opportunity to work with you and I have learned so much along the way.

To my parents: you could not be more supportive of me in all that I do. You have stood beside me in making so many of life's decisions, always lovingly offering your support and advice and allowing me to make the decision that I felt was best for me. I am so glad that you have always encouraged me and love me no matter what.

And most importantly, to my husband Andrew: You encourage me and gave me your support when I considered staying to do a PhD. And you would not allow me to give up this dream even when so many things came crashing down at once. You see something in me that I do not see in myself. You have sacrificed so much to make it possible for me to pursue this degree. Thank you for enduring all the adventures of the last five years with me with a willing spirit and smiling heart. I love you.

DEDICATION

In loving memory of
Dr. Lyle Eugene Hanson, my Poppie
October 2, 1920 – December 19, 2007

We shared a bond that is rare between granafather and granddaughter:

You were my grandfather and also my friend.
You were my mentor and also my colleague.

We shared a deep love for understanding the world around us
—something that never left you throughout your life

You were the academic and prcfessional I aspire to be:

You were intelligent and accomplished, but carried yourself with humility rather than pride.

You never looked down on others who had less education than you, but willingly engaged and learned from people of all education levels and backgrounds.

You understood that knowledge is not simply learned in books, but it is seen, heard, smelled, touched, and tasted in all the experiences God grants you in this life.

You were proud of your history and heritage and also had great interest in and respect for the background and culture of your friends and colleagues.

You had such a wonderful sense of humor through all life brought your way.

In many ways I feel that this dissertation is the culmination of the work that you began many decades ago. I wish with all my heart that you could have been here to see it to completion.

Successes will always carry with them the pain that you are not here to share them with me. And yet, I have a sneaking suspicion that you are watching from heaven, probably still elbowing the guy sitting next to you to tell him, "Did you see that?...she is *my* granddaughter."

I want all who read this work to know that I am the granddaughter cf Dr. Lyle Eugene Hanson—and cf that I could not be more proud.

Lyle Eugene Hanson, Ph.D., D.V.M., was Professor at the University of Illinois College of Veterinary Medicine from 1950-1985. He served as Department Head of Veterinary Pathology and Hygiene from 1967-1979 and Associate Dean for Veterinary Research and Graduate Training from 1979-1985.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Appendix 1

Appendix 2

CHAPTER 1: INTRODUCTION

1.1 The problem of antibiotic resistance

Antibiotic resistance is recognized by the WHO and the Center for Disease Control as one of the main public health challenges of our time (WHO 2000, CDC 2001). During the past century, the medical field has been revolutionized by antibiotics, which have reduced the mortality and morbidity associated with infectious disease. And yet, the rising number of reported antibiotic resistant infections (ARI) and the re-emergence of old scourges in developed countries, such as tuberculosis, that demonstrate enhanced virulence and drug resistance may signal an end to the antibiotic era (Levy 2002, IDSA 2004, Colin et al. 2007, Davies 2007). The incidence of methicillin-resistant *Staphylococcus aureus* (MRSA) infections in intensive care units has risen from 2% in 1974 to nearly 64% in 2004 (Klevens et al. 2007). ARI are no longer just the affliction of hospitals or immunocompromised patients as evidenced by the increasing incidence of community-acquired MRSA (CA-MRSA) infections in healthy children and adults (Bloomfield 2006, Colin et al. 2007). As ARI have been increasing, the numbers of new antibiotics brought to market has been steadily decreasing, from sixteen new antibiotics in the early 1980s to just four since 2003 (Fig 1.1) (Talbot et al. 2006). Since new drug discovery is no longer capable of keeping pace with emerging ARI, strategies for prolonging the efficacy of available antibiotics and reducing the spread of ARI must be developed (WHO 2000, CDC 2001, Talbot et al. 2006, Colin et al. 2007).

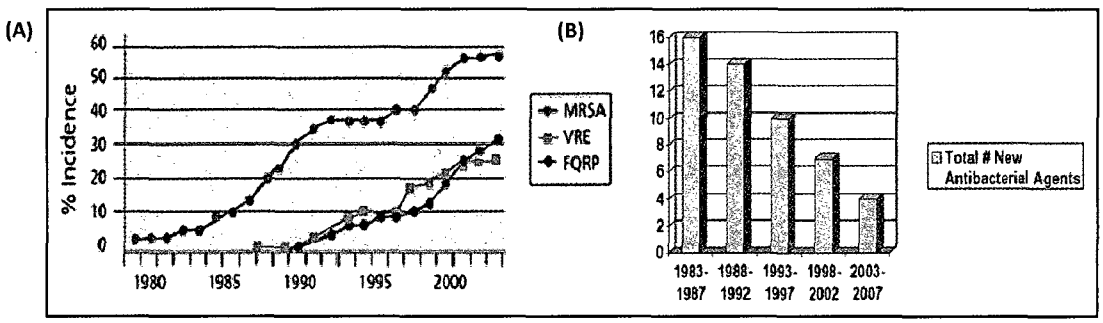

Figure 1.1 The Growing Problem of Antibiotic Resistance.
Rising incidence of antibiotic resistant pathogens in the hospital setting (A) and declining discovery of new antibiotics (B) is limiting their ability to fight infectious disease. MRSA: methicillin- resistant *Staphylococcus aureus*; VRE: vancomycin-resistant Enterococci; FQRP: Floroquinolone-resistant *Psuedomonas aeruginosa* (CDC 2001, Talbot et al. 2006).

1.2 Antibiotic functionality and use

1.2.1 Common modes of action

Originally, the term antibiotic was defined as a chemical derived from microorganisms capable of inhibiting other bacteria. With the introduction of synthetic and semi-synthetic chemicals possessing the same properties, the terminology "antibiotic" has been expanded to include all chemicals that inhibit bacteria, regardless of their origin (Kümmerer 2009). Another characteristic of antibiotics is their ability to act selectively against prokaryotes, either killing them (bacteriocidals) or preventing their growth (bacteriostatics). This allows antibiotics to be used effectively to eradicate bacterial infections without harming the host organism. Most antibiotics target three essential processes in bacteria: cell wall synthesis, protein synthesis, or DNA replication and repair (Walsh 2000). The β-lactam and cephalosporin antibiotics block the active sites of enzymes necessary for cell wall synthesis to render bacterial cells susceptible to lysis. To the same end, glycopeptides interfere with a key product of cell-wall synthesis. Tetracyclines, macrolides, aminoglycosides, and oxazolidinones target one of the many steps in the process of protein synthesis. The synthetic fluoroquinolones are the only antibiotics known to interfere with DNA replication and repair (Walsh 2000). The tetracyclines and sulfonamides

are most directly applicable for the work presented herein, and will be discussed in greater detail below.

1.2.2 Tetracyclines

Chlortetracycline and oxytetracycline were discovered in the early 1940s as natural products produced by *Streptomyces aurecfaciens* and *S. ramus,* respectively (Chopra and Roberts 2001). The low toxicity and broad-spectrum action of these early tetracyclines made them ideal candidates for clinical use (Duggar 1948, Broschard 1949, Nelson and Projan 2005). These properties also prompted further research into the structures of the compounds and eventually lead to the development of several semi-synthetic and synthetic analogs of tetracyclines for clinical use (Nelson and Projan 2005).

Tetracyclines are broad-spectrum antibiotics, which have proven to be effective against both gram-negative and gram-positive bacteria (Katiyar and Elend 1991, Speer et al. 1992, Chopra and Roberts 2001). By binding to the A-site of the 30S subunit of the ribosome, tetracycline blocks the binding of aminoacyl-tRNA to the ribosome; this impedes the elongation step of protein synthesis and in turn, bacterial growth (Goldman et al. 1983, Brodersen et al. 2000, Pioletti et al. 2001, Connell et al. 2002).

The first tetracyclines were introduced into clinical medicine in the 1950s and are still prescribed today, though their effectiveness has been limited by widespread tetracycline resistance (Chopra and Roberts 2001). In addition to their therapeutic uses in human medicine, tetracyclines are also used in veterinary medicine and livestock production for treatment of disease, prophylaxis, and growth promotion, where antibiotics in the feed or water are given to increase the weight gain to feed ratio of livestock animals (Nat. Acad. Sci. 1999, McEwen and Fedorka-Cray 2002). Chlortetracyline is used for growth promotion in cattle, swine, and poultry.

Oxytetracycline is used extensively in aquaculture and also as a growth promoter in cattle (McEwen and Fedorka-Cray 2002).

1.2.3 Sulfonamides

The first sulfonamide Prontosil (4-sulfonamide-2',4'-diaminoazobenzene) was produced synthetically in 1932 and shown to be effective *in vivo* against peritonitis caused by *Streptococcus pyogenes*, but *in vitro* it had no effect. Scientists discovered that Prontosil is metabolized *in vivo* to sulfanilamide—the active form of the drug (Skold 2000). Sulfonamides act by competitive inhibition of dihydropteroate synthase, a key enzyme in folic acid synthesis. The sulfa drugs are structurally similar to the intended substrate, p-aminobenzoic acid. All bacteria must synthesize their own folic, but eukaryotes can obtain folic acid from dietary sources. Thus, sulfonamides can act selectively against bacteria by attacking their folate synthesis pathway (Skold 2000).

Sulfonamides were introduced into clinical medicine in 1935. At the time, they were most often prescribed for the treatment of streptococcal, staphylococcal, pneumococcal, and urinary tract infections (Skold 2000). Sulfonamides were the only antibiotic available in the 1930s and early 1940s and were widely used during those years. However, many people showed allergic reactions to these sulfa drugs, and thus their use as the drug of choice decreased as other antibiotics that exhibited less allergic response became available. Sulfonamide resistance was seen early on after their introduction into clinical medicine, and as this resistance increased, the effectiveness and use of sulfonamides has diminished (Skold 2000). Despite their limited use in human medicine, sulfonamides are commonly used for both treatment and growth promotion in livestock production and for treatment in aquaculture (McEwen and Fedorka-Cray 2002).

1.3 Antibiotic Resistance

The complexity of the antibiotic resistance problem can be more completely understood within the context of bacterial evolution and genetics. Scientists initially believed that mutations occurring at random or as a response to stress were the cause of antibiotic resistance in pathogens (Davies 2007). This type of resistance is seen in nature and is termed intrinsic resistance (Davies 2007, Dzidic et al. 2008). In the 1950s, scientists in Japan discovered "resistant traits" that could be transferred between bacterial species (Davies 2007). Unlike humans who are only able to transfer DNA to their progeny, bacteria can also transfer pieces of DNA, termed mobile genetic elements (MGE) to even distantly related bacteria. This is called acquired resistance (Davison 1999, Frost et al. 2005). ARG are typically present on MGE, and therefore are able to transfer between populations in a process termed horizontal gene transfer (HGT). HGT occurs via three main processes: transformation, transduction and conjugation (Figure 1.2) (Davison 1999, Frost et al. 2005, Martinez et al. 2007, Dzidic et al. 2008).

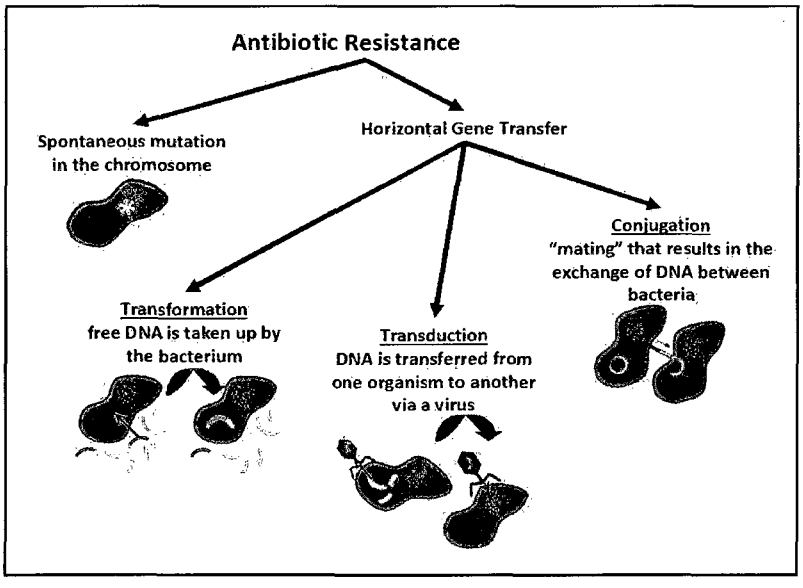

Figure 1.2 Inherent and acquired antibiotic resistance. Antibiotic resistance can be inherent from a spontaneous mutation in the chromosome. It is more frequently acquired through the processes of horizontal gene transfer: transformation, transduction, and conjugation.

1.3.1 Processes of horizontal gene transfer (HGT)

Transformation is the process in which a bacterium takes up free DNA from its environment, integrates the DNA into its cell and maintains it (Dröge et al. 1999). The number of species known to possess competence for natural transformation is limited (~ 50 species), but many other species may also possess the ability for natural transformation but have not been studied in depth (Lorenz and Wackernagel 1994, Dröge et al. 1999). Because transformation is the uptake of extracellular DNA, it does not require physical proximity of the donor and the recipient. The donor may excrete DNA or the DNA may be expelled during cell death and lysis. This means that it is only necessary for the recipient to be in a physiologically active state (Lorenz and Wackernagel 1994). For transformation to take place, competence must be induced in the recipient by appropriate environmental conditions, which can vary depending on the species.

Transduction is the transfer of DNA by means of viruses. There are two main forms of transduction: generalized transduction and specialized transduction. In generalized transduction, bacterial DNA is broken up and packaged into defective virus particles (in other words, bacterial DNA is packaged as a phage, even though the virus no longer contains any viral DNA). These transduceable elements can then infect another bacterium, when the DNA is subsequently released inside the cell and can be incorporated into the recipient's DNA (Matigan et al. 2003). HGT via transduction does not require physical proximity of the host to the donor since transfer of the genetic material is accomplished via a phage. Phages have limited host ranges so both the donor and the host must be susceptible to infection by the same phage for transduction to occur (Lorenz and Wackernagel 1994, Dröge et al. 1999).

In conjugation, the donor must have a conjugative element that can be transferred, be in close proximity to the recipient, establish a physical connection to the recipient, such as a sex pilus, and both donor and recipient must be metabolically active (Lorenz and Wackernagel 1994). Either plasmid or chromosomal DNA can be transferred through conjugation (Matigan et al. 2003).

Plasmids are extra-chromosomal elements of DNA that are not necessary for cell function but can replicate independent of a host cell's replication process. They can occur in one copy or in several hundred copies and can vary greatly in size. Transposable elements are MGE that function by moving DNA within a bacterium. DNA can be moved between locations on a chromosome or between a plasmid and the chromosomal DNA. MGE capable of initiating conjugation are termed conjugative plasmids and transposons (Matigan et al. 2003, Snyder and Champness 2003).

1.3.2 Factors influencing the spread of antibiotic resistance

Antibiotics create selective pressure in an environment; bacteria that have intrinsic resistance or are carrying ARG survive and multiply, while antibiotic-sensitive bacteria die off. It has also been demonstrated that low doses of antibiotics can induce or increase the HGT of ARG (Martinez et al. 2007, Schluter et al. 2007, Dzidic et al. 2008). ARG are often linked to other ARG or to genes conferring resistance to metals, disinfectants, or toxic chemicals, such as chlorinated aromatics. As a result, the presence of any antibiotic, heavy metal, disinfectant, or toxic chemical can create an environment suitable for the selection of ARG (Singer et al. 2006, Schluter et al. 2007). ARG can result in a fitness cost to the host organism. Therefore, it was believed in absence of selective pressure, there would be no advantage to retaining the ARG and the genes would be lost through successive generations (Andersson 2006, Martinez et al.

2007, Schluter et al. 2007). However, ARG have been found to persist in environments after selective pressure has been removed (Davies 2007, Martinez et al. 2007, Schluter et al. 2007, Storteboom et al. 2007). In addition to co-selection, which suggests the presence of additional unknown selective forces in the environment, there are other mechanisms that bacteria have developed to maintain ARG (Davison 1999, Schluter et al. 2007). Evolutionary changes occur rapidly in bacteria, making it possible for bacteria to adapt quickly to changing environments (Davison 1999, Martinez et al. 2007, Schluter et al. 2007). As a result, bacteria can adapt to reduce or eliminate the fitness cost associated with ARG (Davies 2007, Martinez et al. 2007). ARG may confer other advantages to a host, such as improved biofilm formation or surface attachment (mediated by enzymes used in pilus formation during HGT) that off-set the cost of resistance (Couce and Blasquez 2009). Another way ARG can persist in a population is by the incorporation of a gene into chromosomal DNA, as is done when a transposon transfers genetic information from a plasmid to the host's chromosome. Once the gene is part of the chromosome of the bacterium, it is grouped along with genetic information that is necessary for the survival of the bacterium and will be maintained (Matigan et al. 2003).

Since ARG are typically spread through HGT, factors that have been shown to affect rates of HGT will in turn affect the rate of ARG proliferation. Dröge et al. (1999) has reviewed the different conditions that facilitate HGT in the natural environment. In bacteria that are more closely related, conjugation can occur more readily than in bacteria that are very distantly related. Frequency of transfer of genetic information varies widely in nature (10^{-2} - 10^{-7} transconjugates per recipient) (Dröge et al. 1999). For transformation, favorable environmental conditions, such as high bacterial density and abundance of nutrients, may be more suitable for HGT (Lorenz and Wackernagel 1994, Dröge et al. 1999). Some studies have shown that bacteria more readily express and promote the transfer of mobile genetic elements. For example, in a

study of SOS gene (stress gene) expression, it was noted that when SOS genes were expressed, gene excision and expression were induced, sometimes up to 100X (Walsh 2006).

When HGT occurs, the genes may not always be sustained in a population. In the case of plasmids, a plasmid may be lost immediately due to incompatibility with a new plasmid. Certain plasmids, often those that are more closely related, are not able to exist inside the same cell. Therefore, one of the plasmids must be lost. Plasmids that are not able to "co-exist" are grouped into incompatibility groups or *Inc* plasmid groups. When a cell possesses a plasmid of one *Inc* group, that plasmid will not be compatible with another plasmid within the same *Inc* plasmid group (Datta 1979).

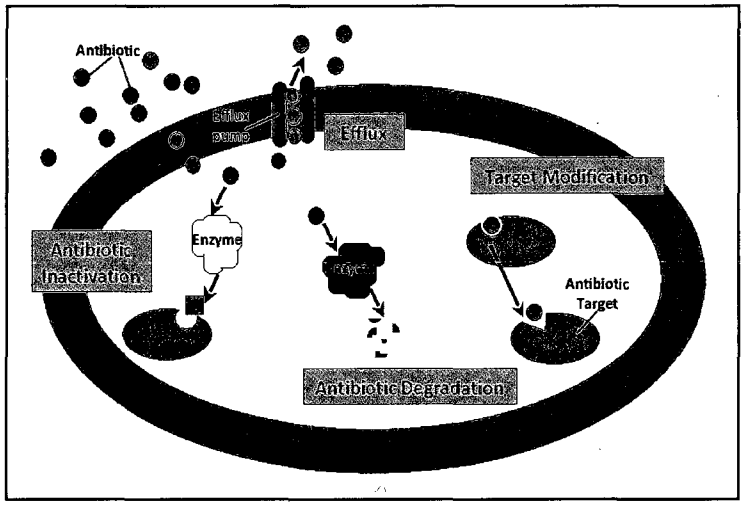

Figure 1.3 Mechanisms of Antibiotic Resistance
(1) Efflux
(2) Antibiotic Inactivation
(3) Antibiotic degradation
(4) Target modification

1.3.3 Mechanisms of antibiotic resistance

Bacteria have evolved four major strategies of antibiotic resistance: (1) efflux of the antibiotic, (2) antibiotic inactivation, (3) antibiotic degradation, and (4) target modification (Figure 1.3). Bacteria can survive by pumping the antibiotic out of its cell, termed efflux, so that the intracellular antibiotic concentration never reaches the lethal or inhibitory dose. ARG conferring this type of resistance code for efflux pumps which are capable of pumping out many types of drugs and chemicals, or are specific for an antibiotic or class of antibiotics (Walsh 2000).

The second and third mechanisms of resistance are related to altering the antibiotic, either by some chemical change that makes it no longer function as intended (inactivation) or by forming an enzyme capable of metabolizing the antibiotic (degradation). When resistance is conferred through target modification, the resistant bacterium is able to alter the structure of the binding site of the antibiotic, such as protection of the ribosome in the case of tetracycline resistance (Walsh 2000, Chopra and Roberts 2001). Alternatively, the bacterium can bypass the target, by creating an alternative pathway that is not affected by the antibiotic, as occurs in vancomycin resistance bacteria (Walsh 2000).

Some research has shown that ARG in pathogens may have originated from antibiotic producers since bacteria producing antibiotics must have some protective mechanism that allows them to produce a toxic substance. While this may be true in some cases, there is a growing body of evidence to suggest that the enzymes coded for in ARG are more often modifications of proteins that fulfill other necessary roles for the bacteria. This working hypothesis is more plausible given that it does not require pathogens to have come in contact with antibiotic producers. In addition, the resistance in antibiotic producers is more often inherent and thus, not carried on MGEs. Aminov and Mackie (2007) wrote an excellent review that discusses the likely ancestors of ARG in environmental and pathogenic bacteria.

1.3.4 Tetracycline resistance genes

Resistance to tetracyclines is conferred by three mechanisms: efflux of tetracyclines, synthesis of ribosomal protection proteins, and enzymatic inactivation of tetracycline (Chopra and Roberts 2001). The tetracycline efflux genes code for membrane-associated proteins that pump tetracyclines out of the cell using a reverse proton gradient (Yamaguchi et al. 1990). The ribosomal protection proteins (RPP) produced by tetracycline resistant bacteria are cytoplasmic proteins that interact with the ribosome, which is allosterically altered at the binding site for

tetracycline (Connell et al. 2003). Most tetracycline resistance (tet) genes confer resistance by these two mechanisms. In addition, three genes are known to encode for proteins that can enzymatically alter tetracycline so that it cannot bind to the ribosome.

To date, there are twenty-six known tet efflux genes, eleven ribosomal protection genes and three genes coding for enzymatic inactivation of tetracycline (Table 1.1). Standard nomenclature dictates that tet genes that share less than 80% similarity represent separate genes (Levy et al. 1999). Tet efflux genes are typically carried on MGE, and this has contributed to their presence in diverse types of bacteria, including gram-negative, gram-positive, aerobic and anaerobic bacteria (Roberts 2005, Supanaric et al. 2005). Tetracycline resistance in gram-negative organisms is usually conferred by efflux genes. The efflux genes with the widest host range are tet(B), 25 genera; tet(A), 15 genera; tet(L), 15 genera; tet(D), 14 genera; tet(K), 13 genera; and tet(C), 12 genera (Roberts 2005).

Table 1.1 Host range of tetracycline resistance genes. Host range is defined as the number of genera in which a gene has been detected.T able was adapted and updated from Roberts (2005). Genes studied in this work are in bold-face type.

Efflux (n = 26)				Ribosomal Protection Protein (n = 11)		Enzymatic inactivation (n=3)		Unknown (n = 1)	
tet (A)	15	*tetA* (P)	1	***tet* (M)**	**42**	*tet* (X)	1	*tet* (U)	2
***tet* (B)**	**25**	*otr* (B)	2	***tet* (O)**	**11**	*tet* (34)*	4		
***tet* (C)**	**12**	*otr* (C)*	1	***tet* (Q)**	**15**	*tet* (37)*	1		
tet (D)	14	*tcr3*	1	***tet* (S)**	**6**				
***tet* (E)**	**7**	*tet* (30)	1	***tet* (T)**	**1**				
tet (G)	7	*tet* (31)	1	***tet* (W)**	**19**				
***tet* (H)**	**5**	*tet* (33)	1	***tetB* (P)**	**1**				
tet (J)	1	*tet* (35)*	2	*otr* (A)	1				
tet (K)	13	*tet* (38)*	1	*tet*	1				
tet (L)	15	*tet* (39)*	1	*tet* (32)*	1				
tet (V)	1	*tet* (40)*	1	*tet* (36)*	3				
tet (Y)	1	*tet* (41)*	1						
tet (Z)	1	*tet* (42)*	7						

* Indicates discovery within the last decade.

RPP genes are commonly found in anaerobic gram-negative commensal bacteria and have also been found in gram-positive and aerobic bacteria (Roberts 2005b). These genes are commonly carried on plasmids or on conjugative transposons that generally have wider host ranges than plasmids. The tet(M), tet(Q) and tet(W) genes are known to be associated with conjugative transposons and have the widest host range of all tet RPP genes (42, 15, and 19 genera, respectively). Of the remaining tet RPP genes, only tet(O), tet(S), and tet(36) have been identified in more than one genera. The tet(O) and tet(S) genes were previously thought to be carried exclusively on plasmids, but recent work has shown they are also found on conjugative transposons (Giovanetti et al. 2003, Brenciani et al. 2004, Lancaster et al. 2004).

There are just three genes that confer resistance through enzymatic alteration: tet(X), tet(34), and tet(37). The tet(X) gene was discovered in 1988 and until recently has been thought to be non-functional in natural environments; it was isolated from a strict anaerobe and yet the NADP-requiring oxidoreductase enzyme that it codes for is only active in the presence of oxygen (Speer and Salyers 1988, Roberts 1994). Recent work by Ghosh and colleagues (2009) have identified tet(X) in an aerobic, tetracycline-resistant isolate of *Sphingobacterium* sp. PM2-P1-29. The tet(34) and tet(37) genes were discovered in the past decade (Nonaka and Suzuki 2002, Diaz-Torres et al. 2003). Due to these recent discoveries, some suggest that enzymatic alteration tet genes may be more prevalent than previously thought (Roberts 2005). Another gene, tet(U), has been found in a bacterium exhibiting low-level resistance to tetracycline. However, the gene does not share significant homology with any of the other tet genes. The protein encoded by tet(U) is most similar to the RPP but does not have the GTP-binding region present on all other RPP (Ridenhour et al. 1996).

1.3.5 Sulfonamide resistance genes

Mutations within the chromosome can confer resistance to sulfonamides by slightly altering the the bacterium's dihydropteroate synthase. These modifications do not allow the binding of sulfonamides and thus free the enzyme's active site for p-aminobenzoic acid to bind and folic acid synthesis to continue (Sköld 2000). The same type of resistance can be carried on plasmids. Sulfonamide resistance (sul) genes on the plasmid encode for variations of dihydropteroate synthase that allow p-aminobenzoic acid to competitively bind to it in place of the sulfa drugs. These sul genes are the only means of resistance known for gram-negative enteric bacteria, which carry either the sul(I) or sul(II) gene with equal frequency. Given their need to encode a variant enzyme that can function in the same manner as the native dihydropteroate synthase, these genes are quite homologous (Sköld 2000). The sul(I) gene is most often found on an integron (Tn21), and the gene is frequently linked with other ARG. The sul(II) gene is carried on plasmids, typically those belonging to the incompatability group *IncQ* (Sköld 2000).

1.4 Pathways and occurrence of antibiotics in the environment

A fraction of the antibiotic dosage consumed by humans and animals will not be metabolized by the body, but will pass unaltered into the feces (Feinman and Matheson 1978). These antibiotics can then enter the natural environment via discharge of municipal wastewater effluent or run-off from agricultural lagoons and manure-amended soils (Kolpin et al. 2002, Kümmerer and Henninger 2003, Yang and Carlson 2003, Davis et al. 2006, Singer et al. 2007). Improved analytical capabilities in the field of environmental chemistry have led to the detection of low concentrations of antibiotics in many natural and engineered environments, such as surface waters and sediments, wastewater treatment facilities, animal waste lagoon systems, groundwater underlying these lagoons, and fields fertilized with animal manures

(Kolpin et al. 2002, Kümmerer and Henninger 2003, Yang and Carlson 2003, Pei et al. 2006, Cha et al. 2006, Kim and Carlson 2006, Kim and Carlson 2007a).

1.4.1 Production of antibiotics

Many antibiotics are produced naturally but at low concentrations by soil-dwelling bacteria, though bacteria in other ecosystems may also produce antibiotics. In studies of undisturbed soils, the concentrations of antibiotics in the soils has been below the limit of detection of the methods applied (Kümmerer 2009, Segura et al. 2009). Antibiotics are produced industrially, but there has been little evidence of pollution from pharmaceutical plants in the United States or Europe (Kümmerer 2009). However, in countries with less stringent environmental law, antibiotics have been quantified in concentrations up to 1 ppm in factory effluent (Kümmerer 2009). Joakim Larsson and colleagues recently analyzed streams in India downstream from industrial pharmaceutical factories and found levels of ciprofloxacin at levels high enough to treat infections. Many other drugs were also detected in the stream, raising concerns for the potential negative health effects of receiving mixtures of many drugs at high doses (Associated Press 2009).

1.4.2 Human medicine

Antibiotics used for human medicine may be improperly disposed of by being discarded in refuse (without taking appropriate precautions such as addition of an adsorbant) or by flushing of antibiotics into the sewage system. The former will make its way to a landfill where the antibiotics could enter the environment via landfill leachate (Segura et al. 2009). In addition to direct disposal of antibiotics to the sewage system, antibiotics and metabolites excreted by humans will make their way to the sewer. From there, some antibiotics may reach the natural environment through cracks or leaks in sewer lines. Most of the inputs from human use of antibiotics, however, are thought to be from WWTPs. Treatment processes are not currently

designed for effective removal of the dissolved antibiotic compounds, so antibiotics can reach natural environments through effluent discharge or application of biosolids (Segura et al. 2009, Kummerer 2009a).

Wastewater has been characterized in terms of antibiotics in raw sewage, activated sludge, and treated effluent. Median antibiotic concentrations in wastewater vary but are typically in the range of 10-100 µg/L in hospital sewage and effluent, and 1-10 µg/L in municipal sewage and effluent (Kim and Aga 2007, Kümmerer 2009, Segura et al. 2009). Wastewater treatment has been shown to reduce the levels of parent antibiotic compounds to some degree with maximum effluent antibiotic concentrations generally lower, in the range of 0.1-1 µg/L. (Kim and Aga 2007, Segura et al. 2009). Factors such as increased residence time, and method of disinfection may be crucial in the removal of antibiotics though the effects of treatment are still being investigated (Jones et al. 2004, Kim and Aga 2007, Baquero et al. 2008, Segura et al. 2009). However, the mechanisms that facilitate this removal are unclear due to the difficulties in identifying and quantifying antibiotic metabolites. If antibiotics remain in the effluent, they may retain their activity as metabolites or may be transformed back to the active parent compound once discharged into the environment (Jones et al. 2004).

1.4.3 Veterinary medicine and livestock production

Antibiotic use in livestock production is seen as the major source of antibiotics from animal use. Livestock are treated with antibiotics to control infection (therapeutic) or for prevention of disease (prophylaxis). When a disease outbreak in a livestock population occurs, or at crucial points in production where incidence of disease is higher than normal, it is common for the entire herd to be treated with therapeutic doses (McEwen and Fedorka-Cray 2002). The use of antibiotics for growth promotion in livestock is a more controversial issue. When used for

therapy or prophylaxis, antibiotics are delivered at clinical doses that, when administered properly are less likely to select for resistance. However, when antibiotics are used as growth promoters, they are administered at levels below the effective (sub-therapeutic) dose. Environments with sub-therapeutic levels of antibiotics, such as the animal gut, are ideal habitats for selection and transfer of genetic material, such as ARG (Dröge et al. 1999). As in humans, the active antibiotics and their metabolites are excreted in urine and feces. Due to the high bacterial density and availability of nutrients, animal feces and manure are other ideal habitats for ARG selection and transfer (Chee-Sanford et al. 2001).

The treatment of animal manure varies depending on the size of the operation, the type of animal production (*e.g.* dairy or beef), and the local environmental regulations. Much of the work concerning antibiotic pathways from animal sources to natural environments has focused on operations with large numbers of animals that are typically concentrated in small areas. These operations are called concentrated animal feeding operations (CAFOs). Dairies and swine facilities typically use water to wash out pens. This water is then collected in lined or unlined lagoons and can then be applied in liquid form as fertilizer to nearby fields. These waste systems may involve multiple lagoons or a single lagoon and may or may not be aerated. Lagoons are dredged occasionally to remove the build-up of sludge that can then be applied as a solid fertilizer. On beef and sheep feedlots, stockpiling or composting of solid manure are more typical management practices. In areas where feedlots and other CAFOs are subject to environmental regulations, lined or unlined lagoons are usually present to collect the run-off from the pens and stockpiled manure.

Concentrations of antibiotics in agricultural wastewaters vary over a wide range, with median quantities between 0.0001 – 10,000 µg/L (Segura et al. 2009). As leaching or run-off

from agricultural fields is thought to be the predominant pathway of agricultural antibiotics in the environment, concentrations of antibiotics in agricultural fields have also been well documented. Field application of manures can lead to the build-up of antibiotics in high quantities since most antibiotics appear to be recalcitrant in this environment (Khan et al. 2008). More comprehensive information is available for the contamination of soils with antibiotics due to field application of solid or liquid manure (Jjemba et al. 2002, Kemper et al. 2008). Sulfonamides and tetracyclines have been quantified in agricultural soils with ranges between 1 – 11 μg/kg for sulfonamides and 39 – 1000 μg/kg for tetracyclines (Kemper 2008).

1.4.4 Plant agriculture

A few antibiotics, mainly streptomycin and oxtetracycline, are used to control infectious disease in tree-fruit production in the US and in Latin America (Vidaver et al. 2002, Kummerer et al. 2009a). In the US this use constitutes only a minor (<0.5%) fraction of total antibiotic use; however since antibiotics used on plants must be capable of withstand to harsh environmental conditions, such as intense UV rays and precipitation, they are likely to be more persistent in the environment (Kümmerer 2009). Antibiotics can enter the environment via run-off from orchards, but detailed information regarding the presence of antibiotics in orchard soils or surrounding waters is not available.

1.4.5 Aquaculture

In practice, antibiotics are frequently used in aquaculture for therapeutic purposes. They may be administered in the feed, by injection, or by dosing of ponds. In the US, antibiotics are not allowed for growth promotion in aquaculture. But in Asia, where the aquaculture industry is growing rapidly, there is less regulation (Kümmerer 2009). Antibiotics may be directly discharged from ponds, overflow as run-off during rainfall events, enter surface waters through

Table 1.2 Occurrence in nanograms per liter (ppt) of antibiotics in urban and agricultural wastewater, surface water, and drinking water. Table is adapted from Segura et al. 2009. Sulfonamides and tetracyclines are highlighted as the most important to this dissertation.

Class of Antibiotic	No. > LOQ[a]	No. papers[b]	Mean	Median	75%	95%	Max	LOQ_{low}[c]	LOQ_{high}[d]
Wastewaters									
parent compounds									
Azoles	17	6	5,987	26	50	61,920	90,200	5	112
β-Lactams	42	15	4,633	300	1,200	5,360	153,000	1	100,000
Quinolaxine-dioxide	0	5	5	100					
Lincosamides	33	10	18,715	60	2,125	190,500	240,000	0.3	100
Macrolides	217	43	352	110	271	1,000	27,000	0.2	1,155
Poliether ionophores	13	2	29	11	26	167	190	1	3
Quinolones	420	51	152,247	205	570	41,922	31,000,000	1	20,600
Sulfonamides	289	57	11,972	330	800	31,000	1,158,680	1	300,000
Tetracyclines	161	32	11,642,200	530	7,250	6,095,000	920,000,000	1	700,000
Trimethoprime	210	52	1,351	270	795	5,000	55,200	1	150,000
Other	17	11	750	39	1,115	4,101	5,000	3	667
metabolites/degradation products									
β-Lactams	10	1	55,794,600	7,630,000	44,500,000	389,000,000	389,000,000	410	1,360
Macrolides	52	16	2,848	450	1,414	5,650	83,000	1	500
Sulfonamides	13	5	779	570	1,200	2,230	2,235	20	212
Tetracyclines	6	1	5,092,167	1,445,000	9,420,000	18,100,000	18,100,000	600	1,300
Natural waters									
parent compounds									
Azoles	39	11	18	18	23	44	58	1	370
β-Lactams	6	16	73	11	48	350	350	2	24,000
Quinolaxine-dioxide	0	9	35	1,400					
Lincosamides	46	21	147	18	100	1,020	1,400	0.04	198
Macrolides	128	38	58	11	46	197	1,022	0.02	1,155
Poliether ionophores	4	4	312	35	606	1,172	1,172	0.1	380
Quinolones	78	31	199	27	108	640	5,600	0.3	7,000
Sulfonamides	234	60	66,531	120	700	472,000	1,600,000	0.2	33,000
Tetracyclines	47	33	97,369	192	658	623,550	712,000	0.07	1,650
Trimethoprim	90	24	94	18	42	510	3,000	0.2	4,000
Other	5	21	136	127	207	266	266	2	2,000
metabolites/degradation products									
β-Lactams	4	1	4,719,500	4,037,500	8,840,000	10,540,000	10,540,000	410	1,360
Macrolides	51	24	184	40	146	1,186	1,700	0.3	250
Sulfonamides	6	6	86	14	239	240	240	2	50
Tetracyclines	27	25	12,367	11,100	14,750	32,840	34,200	10	1,300
Drinking waters									
parent compounds									
Azoles	0	1	370	370					
β-Lactams	0	0	NA	NA					
Quinolaxine-dioxide	0	1	100	100					
Lincosamides	0	2	0.07	5					
Macrolides	4	2	3	3	5	5	5	0.07	220
Poliether ionophores	0	0	NA	NA					
Quinolones	5	2	2	2	3	4	4	0.3	10
Sulfonamides	2	5	0.4	0.4	0.5	0.5	0.5	0.2	1,155
Tetracyclines	0	1	3	12					
Trimethoprim	0	4	0.5	250					
Other	0	1	5	5					
metabolites/degradation products									
β-Lactams	0	0	NA	NA					
Macrolides	0	1	10	10					
Sulfonamides	0	0	NA	NA					
Tetracyclines	0	0	NA	NA					

NA, not available.; [a]Number of values reported as being > LOQ for each class; [b]Number of papers reporting analysis in each matrix for each class; [c]Lowest LOQ reported; [d]Highest LOQ reported.

river bank filtration, or seep from ponds into groundwater. Antibiotics in aquaculture ponds have been detected at concentrations ranging from 0.0001 – 13,000 μg/L, and it has been suggested that 70-80% of the drugs used in the aquaculture industry will reach the natural environment (Segura et al. 2009).

1.5 Occurrence of antibiotic resistance in the environment

Many have proposed that the natural environment has become a reservoir of ARG, where novel resistance strategies can evolve, persist and be passed on to pathogenic bacteria (Rysz and Alvarez 2004, Yu et al. 2005, D'Costa et al. 2006, Pei et al. 2006, Pruden et al. 2006, Davies 2007, Schluter et al. 2007). ARG can easily be transferred between diverse bacteria, selected (amplified) within a population, co-selected by other chemicals, and persist even in the absence of selective pressure. In this way, ARG can be viewed as unique environmental contaminants (Pruden et al. 2006).

1.6 Overview of past work on the Cache La Poudre River

The Poudre River is a highly-suitable model system for characterizing the sources and processes involved in the transport of ARG. It has a pristine source arising from snow melt in the Rocky Mountains and few tributaries. It is also well-zonated between pristine, agriculture and urban influences (Yang and Carlson 2003, Pruden et al. 2006). Levels of many classes of antibiotics including tetracyclines, β-lactams, macrolides, sulfonamides, and ionophores, have been characterized in the Poudre River system (Yang and Carlson 2003, Yang and Carlson 2004a, Yang and Carlson 2004b, Yang et al. 2004, Cha et al. 2005, Yang et al. 2005, Kim and Carlson 2005, Cha et al. 2006, Kim and Carlson 2006, Yang et al. 2006, Kim and Carlson 2007a, Kim and Carlson 2007b). Concentrations of antibiotics increased downstream, and agricultural antibiotics were found at higher levels in agricultural zones, as were human antibiotics in urban

zones (Yang and Carlson 2003, Kim and Carlson 2006, Pei et al. 2006, Pruden et al. 2006, Kim and Carlson 2007). To date, no antibiotics have been detected in the pristine zone (Kim and Carlson 2007). Additionally, Pei et al. (2006) and Pruden et al. (2006) have demonstrated increased levels of ARG at more heavily impacted sites along the river. These results validate the Poudre River as a model system for the study of ARG transport.

Two processes are implicated as driving factors for the high levels of ARGs observed in surface waters: (1) The presence of low levels of antibiotics, metals, and disinfectants causes native bacteria carrying ARG to be selected and/or causes an increase in the HGT of ARG (Singer et al. 2006, Schluter et al. 2007); or (2) Non-native bacteria possessing ARG are transported to surface waters via point-source pollution, such as WWTP effluents such as runoff from feedlots, or non-point source pollution, such as runoff from agricultural fields (Singer 2007, Baquero 2009). To be able to apply management practices to control the spread of ARG, the dominant sources and mechanisms involved in the spread of ARG through natural environmental matrices must be determined.

Recent studies concerning ARG in the environment (Kummerer and Henninger 2003, Schwartz et al. 2003, Rysz and Alvarez 2004, D'Costa et al. 2006, Pei et al. 2006, Pruden et al. 2006, Auerbach et al. 2007, Koike et al. 2007, Hu et al. 2008, Meinersmann et al. 2008, Tuckfield and McArthur 2008) often neglected to analyze their collected data in a way that explores how potential variables may be involved in the presence and transport of ARG in the natural environment (Singer et al. 2006). One exception is the work of Koike et al. (2007). By evaluating the phylogenetic diversity of ARG in an unlined swine lagoon and the underlying groundwater and building on past work (Chee-Sanford et al. 2003), Koike et al. (2007) was able to establish evidence supporting the transport of ARG from the lagoon to the groundwater.

1.7 Research Objectives

The goal of this research is to identify ARG profiles that effectively classify potential sources and reservoirs of ARG in the environment and then to apply this approach to determine the sources and mechanisms involved in the transport of ARG to the Poudre River environment. To guide this study, six objectives are outlined:

(1) Characterize the background ARG profile of undisturbed river water and sediments.

(2) Characterize potential source environments with respect to the distribution, frequency of detection, and phylogenetic diversity of their ARG profiles.

(3) Determine if certain ARG or genetic signatures of ARG can serve as indicators of either urban or agricultural influence.

(4) Characterize and classify river samples according to the distribution, frequency of detection and phylogenetic diversity of their ARG profiles.

(5) Create a geospatial database of the study area to assist in developing variables that reflect the measure of urban and agricultural inputs in the watershed.

(6) Determine which spatial indicators are effective in explaining the variations in the ARG profiles of river samples.

The research presented herein is a novel approach that combines molecular methods for determining the presence, concentration, and diversity of ARG in environmental samples with geospatial analysis of the study watershed to achieve these objectives. Chapter 2 will focus on the characterization of source environments. Then in Chapter 3, the results and approach will be applied to the Poudre River environment. Finally, Chapter 4 will discuss the potential next steps in this field of research.

REFERENCES

Aminov RI and Mackie RI (2007) Evolution and ecology of antibiotic resistance genes. *FEMS Microbiology Letters*, 271: 147-161.

Aminov, RI, Garrigues-Jeanjean, N, and Mackie, RI (2001) Molecular ecology of tetracycline resistance: Development and validation of primers for detection of tetracycline resistance genes encoding ribosomal protection proteins. *Applied and Environmental Microbiology,* 67: 22-32.

Aminov, RI, Chee-Sanford, JC, Garrigues, N, Teferedegne, B, Krapac, IJ, White, BA, Mackie, RI (2002) Development, validation, and application of PCR primers for detection of tetracycline efflux genes of gram-negative bacteria. *Applied and Environmental Microbiology,* 68: 1786-93.

Andersson, DI (2006) The biological cost of mutational antibiotic resistance: any practical conclusions? *Current Opinion in Microbiology,* 9: 461-465.

Associated Press (2009) World's highest drug levels entering India stream. *The Washington Post.* Originally published January 26, 2009.

Auerbach, EA, Seyfried, EE, McMahon, KD (2007) Tetracycline resistance genes in activated sludge wastewater treatment plants. *Water Research,* 41: 1143-1151.

Baquero F, Martinez JL, Canton R (2008) Antibiotics and antibiotic resistance in water environments. *Current Opinion in Biotechnology,* 19: 260-265.

Bloomfield, SF (2006) Community-acquired MRSA. *Lancet*, 368 (9544): 1328-1328.

Brenciani, A, Ojo, KK, Monachetti, A, Menzo, S, Roberts, MC, Varaldo, PE, and Giovanetti, E (2004) Distribution and molecular analysis of mef(A)-containing elements in tetracycline-susceptible and -resistant Streptococcus pyogenes clinical isolates with efflux-mediated erythromycin resistance. *The Journal of Antimicrobial Chemotherapy,* 54: 991-998.

Brodersen, DE, Clemons, JWM, Carter, AP, Morgan-Warren, RJ, Wimberly, BT, and Ramakrishnan, V (2000) The Structural Basis for the Action of the Antibiotics Tetracycline, Pactamycin, and Hygromycin B on the 30S Ribosomal Subunit. *Cell*, 103: 1143-1154.

Broschard, RW, Dornbush, AC, Gordon, S, Hutchings, BL, Kohler, AR, Krupka, G et al (1949) Aureomycin, a New Antibiotic. *Science*, 109: 199-200.

CDC, Centers for Disease Control (2001) National Nosocomial Infections Surveillance (NNIS) System Report, Data Summary from January 1992-June 2001, Issued August (2001). *American Journal of Infection Control*, 29: 404-421.

Cha, JM, Yang, S, Carlson, KH (2005) Rapid analysis of trace levels of antibiotic polyether ionophores in surface water by solid-phase extraction and liquid chromatography with ion trap tandem mass spectrometric detection. *Journal of Chromatography A*, 1065 (2): 187-198.

Cha, JM, Yang, S, Carlson, KH (2006) Trace determination of beta-lactam antibiotics in surface water and urban wastewater using liquid chromatography combined with electrospray tandem mass spectrometry. *Journal of Chromatography A*, 1115 (1-2): 46-57.

Chee-Sanford, JC, Aminov, RI, Krapac, IJ, Garrigues-Jeanjean, N, and Mackie, RI (2001) Occurrence and diversity of tetracycline resistance genes in lagoons and groundwater

underlying two swine production facilities. *Applied and Environmental Microbiology*, 67: 1494-1502.

Chopra, I, and Roberts, M (2001) Tetracycline antibiotics: Mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiology and Molecular Biology Reviews*, 65: 232-260.

Colin, TJ, Koch, K, Jost, K Davis, OB, Dela Rosa, D (eds) (2007) Fighting Superbugs: Are disease-resistant bacteria becoming unstoppable? *CQ Researcher*, 17 (29): 673-696.

Connell, SR, Trieber, CA, Stelzl, U, Einfeldt, E, Taylor, DE, and Nierhaus, KH (2002) The tetracycline resistance protein Tet(O) perturbs the conformation of the ribosomal decoding centre. *Molecular Microbiology,* 45: 1463-1472.

Connell, SR, Tracz, DM, Nierhaus, KH, and Taylor, DE (2003) Ribosomal Protection Proteins and Their Mechanism of Tetracycline Resistance. *Antimicrobial Agents and Chemotherap,* 47: 3675-3681.

Datta, N. Plasmid classification: Incompatibility grouping. (1979), In: K.N. Timmis and A. Pühler, Editors, *Plasmids of Medical, Environmental, and Commercial importance. Developments in Genetics* vol. 1, Elsevier/North-Holland Biomedical press, Amsterdam: 3–12.

Davies, J (2007) Microbes have the last word. *EMBO Reports*, 8 (7): 616-621.

Davis, JG, Truman, CC, Kim, SC, Ascough, JC, Carlson, K (2006) Antibiotic transport via runoff and soil loss. *Journal of Environmental Quality,* 35 (6): 2250-2260.

Davison, J (1999) Genetic exchange between bacteria in the environment. Plasmid 42 (2):73-91

D'Costa, VM, McGrann, KM, Hughes, DW, Wright, GD (2006) Sampling the antibiotic resistome. *Science*, 311 (5759): 374-377.

Diaz-Torres, ML, McNab, R, Spratt, DA, Villedieu, A, Hunt, N, Wilson, M, and Mullany, P (2003) Novel Tetracycline Resistance Determinant from the Oral Metagenome. *Antimicrobial Agents and Chemotherapy*, 47: 1430-1432.

Droge M, Puhler A, Selbitschka W (1999) Horizontal gene transfer among bacteria in terrestrial and aquatic habitats as assessed by microcosm and field studies. *Biology and Fertility of Soils*, 29: 221-245.

Duggar, BM (1948) Aureomycin: a product of the continuing search for new antibiotics. *Annals of the New York Academy of Sciences USA*, 51: 171-181.

Dzidic, S, Suskovic, J, Kos, B (2008) Antibiotic resistance mechanisms in bacteria: Biochemical and genetic aspects. *Food Technology and Biotechnology,* 46 (1):11-21.

Feinman, SE, and Matheson, JC (1978) Draft environmental impact statement: subtherapeutic antibacterial agents in animal feeds. *Food and Drug Administration Department of Health, Education and Welfare Report*, p372. Food and Drug Administration, Washington, DC.

Frost, LS, Leplae, R, Summers, AO, Toussaint, A (2005) Mobile genetic elements: The agents of open source evolution. *Nature Reviews Microbiology*, 3 (9):722-732.

Giovanetti, E, Brenciani, A, Lupidi, R, Roberts, MC, and Varaldo, PE (2003) Presence of the tet(O) Gene in Erythromycin- and Tetracycline-Resistant Strains of Streptococcus pyogenes and Linkage with either the mef(A) or the erm(A) Gene. *Antimicrobial Agents and Chemotherapy*, 47: 2844-2849.

Goldman, RF, Hasan, T, Hall, CC, Strycharz, WA, and Cooperman, BS (1983) Photoincorporation of Tetracycline into Escherichia coli Ribosomes Identification of the Major Proteins Photolabeled by Native Tetracycline and Tetracycline Photoproducts and Implications for the Inhibitory Action of Tetracycline on Protein Synthesis. *Biochemistry,* 22: 359-368.

Ghosh S, Sadowsky MJ, Roberts MC, Gralnick JA, and LaPara TM (2009) *Sphingobacterium* sp strain PM2-P1-29 harbours a functional tet(X) gene encoding for the degradation of tetracycline. *Journal of Applied Microbiology*, 106 (4): 1336-1342.

Hu, JY, Shi, JC, Chang, H, Li, D, Yang, M, Kamagata, YC (2008) Phenotyping and genotyping of antihiotic-resistant Escherichia coli isolated from a natural river basin. *Environmental Science & Technology*, 42 (9):3415-3420.

Infectious Disease Society of America (2004) Bad Bugs, No Drugs: As Antibiotic Discovery Stagnates...A Public Health Crisis Brews (http://wwwidsocietyorg/badbugsnodrugshtml) accessed online June 30, 2008.

Jjemba PK (2002) The potential impact of veterinary and human therapeutic agents in manure and biosolids on plants grown on arable land: a review. *Agriculture Ecosystems & Environment*, 93: 267-278.

Jones OAH, Voulvoulis N, Lester JN (2004) Potential ecological and human health risks associated with the presence of pharmaceutically active compounds in the aquatic environment. *Critical Reviews in Toxicology*, 34: 335-350.

Katiyar, SK, and Elend, TD (1991) Enhanced Antiparasitic Activity of Lipophilic Tetracyclines: Role of Uptake. *Antimicrobial Agents and Chemotherapy*, 35: 2198-2202.

Kemper N (2008) Veterinary antibiotics in the aquatic and terrestrial environment. *Ecological Indicators*, 8: 1-13.

Khan, SJ, Roser, DJ, Davies, CM, Peters, GM, Stuetz, RM, Tucker, R, Ashbolt, NJ (2008) Chemical contaminants in feedlot wastes: Concentrations, effects and attenuation. *Environment International*, 34: 839-859.

Kim, S, Aga, DS (2007) Potential ecological and human health impacts of antibiotics and antibiotic-resistant bacteria from wastewater treatment plants. *Journal of Toxicology and Environmental Health-Part B-Critical Reviews,* 10: 559-573.

Kim, SC, Carlson, K (2005) LC-MS2 for quantifying trace amounts of pharmaceutical compounds in soil and sediment matrices. *TrAC Trends in Analytical Chemistry*, 24 (7):635-644.

Kim, SC, Carlson, K (2006) Occurrence of ionophore antibiotics in water and sediments of a mixed-landscape watershed. *Water Research*, 40 (13):2549-2560.

Kim, SC, Carlson, K (2007) Quantification of human and veterinary antibiotics in water and sediment using SPE/LC/MS/MS. *Analytical and Bioanalytical Chemistry*, 387 (4):1301-1315.

Kim, SC, Carlson, K (2007) Temporal and spatial trends in the occurrence of human and veterinary antibiotics in aqueous and river sediment matrices. *Environmental Science & Technology*, 41: 50-57.

Klevens, RM, Morrison, MA, Fridkin, SK, Reingold, A, Petit, S, Gershman, K, Ray, S, Harrison, LH, Lynfield, R, Dumyati, G, Townes, JM, Craig, AS, Fosheim, G, McDougal, LK, Tenover, FC, Active Bacterial Core Surveillance (2006) Community-associated methicillin-resistant

Staphylococcus aureus and healthcare risk factors. *Emerging Infectious Diseases*, 12 (12):1991-1993.

Koike, S, Krapac, IG, Oliver, HD, Yannarell, AC, Chee-Sanford, JC, Aminov, RI, Mackie, RI (2007) Monitoring and source tracking of tetracycline resistance genes in lagoons and groundwater adjacent to swine production facilities over a 3-year period. *Applied and Environmental Microbiology*, 73 (15):4813-4823.

Kolpin, DW, Furlong, ET, Meyer, MT, Thurman, EM, Zaugg, SD, Barber, LB, and Buxton, HT (2002) Pharmaceuticals, hormones, and other organic wastewater contaminants in US streams, 1999-2000: A national reconnaissance. *Environmental Science & Technology*, 36: 1202-1211.

Kümmerer, K, and Henninger, A (2003) Promoting resistance by the emission of antibiotics from hospitals and households into effluent. *Clinical Microbiology and Infection*, 9: 1203-1214.

Kümmerer (2009a) Antibiotics in the aquatic environment - A review - Part I. *Chemosphere*, 75: 417-434.

Lancaster, H, Roberts, AP, Bedi, R, Wilson, M, and Mullany, P (2004) Characterization of Tn916S, a Tn916-Like Element Containing the Tetracycline Resistance Determinant tet(S). *Journal of Bacteriology*, 186: 4395-4398.

Levy, SB (2002) *The Antibiotic Paradox: How the Misuse of Antibiotics Destroys Their Curative Powers.* Cambridge: Perseus Publishing.

Levy, SB, McMurry, LM, Barbosa, TM, Burdett, V, Courvalin, P, Hillen, W et al (1999) Nomenclature for New Tetracycline Resistance Determinants. *Antimicrobial Agents and Chemotherapy*, 43: 1523-1524.

Lorenz MG, Wackernagel, W (1994) Bacterial Gene-Transfer by Natural genetic transformation in the Environment. *Microbiology Reviews*, 58: 563-602.

Martinez, JL, Baquero, F, Andersson, DI (2007) Predicting antibiotic resistance. *Nature Reviews Microbiology*, 5: 958-965.

Matigan, MT, Martinko, JM, and Parker, J (eds) (2003) *BROCK Biology of Microorganisms*, Upper Saddle River, NJ: Pearson Education.

McEwen, S, and Fedorka-Cray, PJ (2002) Antimicrobial Use and Resistance in Animals. *Clinical Infectious Diseases*, 34: S93-S106.

Meinersmann, RJ, Berrang, ME, Jackson, CR, Fedorka-Cray, P, Ladely, S, Little, E, Frye, JG, Mattsson, B (2008) Salmonella, Campylobacter and Enterococcus spp: Their antimicrobial resistance profiles and their spatial relationships in a synoptic study of the Upper Oconee River basin. *Microbial Ecology*, 55 (3):444-452.

National Academy of Sciences Committee on Drug Use in Food Animals (1999) The use of drugs in food animals: benefits and risks. Washington, DC: National Academy Press.

Nelson, ML, and Projan, SJ (2005) Discovery and Industrialization of Therapeutically Important Tetracyclines, In *Frontiers in Antimicrobial Resistance: A Tribute to Stuart B Levy*. White, DG, Alekshun, MN, and McDermott, PF (eds) Washington, DC: ASM Press, pp 29-38.

Nonaka, L, and Suzuki, S (2002) New Mg2+-Dependent Oxytetracycline Resistance Determinant Tet 34 in Vibrio Isolates from Marine Fish Intestinal Contents. *Antimicrobial Agents and Chemotherapy*, 46: 1550-1552.

Pei, R, Kim, S-C, Carlson, KH, and Pruden, A (2006) Effect of River Landscape on the sediment concentrations of antibiotics and corresponding antibiotic resistance genes (ARG). *Water Research*, 40: 2427-2435.

Pei, R, Cha, J, Carlson, KH, Pruden, A (2007) Response of antibiotic resistance genes (ARG) to biological treatment in dairy lagoon water. *Environmental Science & Technology*, 41 (14): 5108-5113.

Pioletti, M, Schlünzen, F, Harms, J, Zarivach, R, Glühmann, M, Avila, H et al (2001) Crystal structures of complexes of the small ribosomal subunit with tetracycline, edeine and IF3. *The EMBO Journal*, 20: 1829-1839.

Pruden, A, Pei, RT, Storteboom, H, Carlson, KH (2006) Antibiotic resistance genes as emerging contaminants: Studies in northern Colorado. *Environmental Science & Technology*, 40 (23): 7445-7450.

Ridenhour, MB, Fletcher, HM, Mortensen, JE, and Daneo-Moore, L (1996) A Novel Tetracycline-Resistant Determinant, tet(U), Is Encoded on the Plasmid pKQ10 inEnterococcus faecium. *Plasmid*, 35: 71-80.

Roberts, MC (1994) Epidemiology of tetracycline-resistance determinants. *Trends in Microbiology*, 2: 353-357.

Roberts, MC (2005) Tetracycline Resistance Due to Ribosomal Protection Proteins, In *Frontiers in Antimicrobial Resistance: A Tribute to Stuart B Levy*. White, DG, Alekshun, MN, and McDermott, PF (eds) Washington DC: ASM Press, pp 19-28.

Roberts, MC (2005) Update on acquired tetracycline resistance genes. *FEMS Microbiology Letters*, 245: 195-203.

Rysz, M, and Alvarez, PJJ (2004) Amplification and attenuation of tetracycline resistance in soil bacteria: aquifer column experiments. *Water Research*, 38: 3705-3712.

Schluter, A, Szczepanowski, R, Puhler, A, Top, EM (2007) Genomics of IncP-1 antibiotic resistance plasmids isolated from wastewater treatment plants provides evidence for a widely accessible drug resistance gene pool. *FEMS Microbiology Reviews*, 31 (4):449-477.

Schwartz, T, Kohnen, W, Jansen, B, Obst, U (2003) Detection of antibiotic-resistant bacteria and their resistance genes in wastewater, surface water, and drinking water biofilms. *FEMS Microbiology Ecology*, 43(3): 325-335.

Segura PA, Francois M, Gagnon C, Sauve S (2009) Review of the Occurrence of Anti-infectives in Contaminated Wastewaters and Natural and Drinking Waters. *Environmental Health Perspectives*, 117: 675-684.

Singer, RS, Ward, MP, Maldonado, G (2007) Can landscape ecology untangle the complexity of antibiotic resistance? *Nature Reviews Microbiology*, 4: 943 – 952.

Sköld O (2000) Sulfonamide resistance: mechanisms and trends. *Drug Resistance Updates*, 3: 155-160.

Snyder, L, and Champness, W (eds) (2003) Transposition and Site-Specific Recombination In Molecular Genetics of Bacteria. Washington, DC: ASM Press, pp 303-340.

Speer, BS, and Salyers, AA (1988) Characterization of a Novel Tetracycline Resistance That Functions Only in Aerobically Grown Escherichia-Coli. *Journal of Bacteriology*, 170: 1423-1429.

Speer, BS, Shoemaker, NB, and Salyers, AA (1992) Bacterial-Resistance to Tetracycline - Mechanisms, Transfer, and Clinical-Significance. *Clinical Microbiology Reviews*, 5: 387-399.

Storteboom, HN, Kim, SC, Doesken, KC, Carlson, KH, Davis, JG, Pruden, A (2007) Response of antibiotics and resistance genes to high-intensity and low-intensity manure management. *Journal of Environmental Quality*, 36: 1695-1703

Sapunaric, FM, Aldema-Ramos, M, and McMurry, LM (2005) Tetracycline Resistance: Efflux, Mutation, and Other Mechanisms, In *Frontiers in Antimicrobial Resistance: A Tribute to Stuart B Levy*. White, DG, Alekshun, MN, and McDermott, PF (eds) Washington, DC: ASM Press, pp 3-18.

Talbot, GH, J Bradley, et al (2006) Bad bugs need drugs: an update on the development pipeline from the Antimicrobial Availability Task Force of the Infectious Diseases Society of America. *Clinical Infectious Diseases*, 42 (5): 657-668.

Tuckfield, RC, McArthur, JV (2008) Spatial analysis of antibiotic resistance along metal contaminated streams. *Microbial Ecology*, 55 (4): 595-607.

Vidaver, AK (2002) Uses of antimicrobials in plant agriculture. *Clinical Infectious Diseases*, 34: S107-S110.

Walsh C (2000) Molecular mechanisms that confer antibacterial drug resistance. *Nature*, 406: 775-781.

Walsh, TR (2006) Combinatorial genetic evolution of multiresistance. *Current Opinion in Microbiology*, 9: 476-482.

Wegener, HC (2003) Antibiotics in animal feed and their role in resistance development. *Current Opinion in Microbiology*, 6: 439-445.

WHO, World Health Organization (2000) WHO annual report on infectious disease: Overcoming antimicrobial resistance http://wwwwhoint/infectious-disease-report/2000/.

Yamaguchi, A, Ono, N, Akasaka, T, Noumi, T, and Sawai, T (1990) Metal-tetracycline/H+ antiporter of Escherichia coli encoded by a transposon, Tn10 The role of the conserved dipeptide, Ser65-Asp66, in tetracycline transport. *The Journal of Biological Chemistry*, 265: 15525-15530.

Yang, S, and Carlson, K (2003) Evolution of antibiotic occurrence in a river through pristine, urban and agricultural landscapes. *Water Research*, 37: 4645-4656.

Yang, S, Carlson, K (2004) Routine monitoring of antibiotics in water and wastewater with a radioimmunoassay technique. *Water Research*, 38 (14-15): 3155-3166.

Yang, S, Carlson, KH (2004) Solid-phase extraction-high-performance liquid chromatography-ion trap mass spectrometry for analysis of trace concentrations of macrolide antibiotics in natural and waste water matrices. *Journal of Chromatography A*, 1038 (1-2): 141-155.

Yang, SW, Cha, J, Carlson, K (2004) Quantitative determination of trace concentrations of tetracycline and sulfonamide antibiotics in surface water using solid-phase extraction and liquid chromatography/ion trap tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, 18 (18): 2131-2145.

Yang, WR, Moore, IF, Koteva, KP, Bareich, DC, Hughes, DW, and Wright, GD (2004) TetX is a flavin-dependent monooxygenase conferring resistance to tetracycline antibiotics. *Journal of Biological Chemistry*, 279: 52346-52352.

Yang, SW, Cha, JM, Carlson, K (2005) Simultaneous extraction and analysis of 11 tetracycline and sulfonamide antibiotics in influent and effluent domestic wastewater by solid-phase extraction and liquid chromatography-electro spray ionization tandem mass spectrometry. *Journal of Chromatography A*, 1097 (1-2): 40-53.

Yang, SW, Cha, JM, Carlson, K (2006) Trace analysis and occurrence of anhydroerythromycin and tylosin in influent and effluent wastewater by liquid chromatography combined with electrospray tandem mass spectrometry. *Analytical and Bioanalytical Chemistry*, 385 (3): 623-636.

Yu, ZT, Michel, FC, Hansen, G, Wittum, T, and Morrison, M (2005) Development and application of real-time PCR assays for quantification of genes encoding tetracycline resistance. *Applied and Environmental Microbiology*, 71: 6926-6933.

Zhang XX, Zhang T, Fang H (2009) Antibiotic resistance genes in water environment. *Applied Microbiology and Biotechnology*, 82: 397-414.

CHAPTER 2: CHARACTERIZING ANTIBIOTIC RESISTANCE GENE (ARG) PROFILES OF PRISTINE RIVER AND PUTATIVE URBAN AND AGRICULTURAL SOURCE ENVIRONMENTS[1]

Abstract

Antibiotic resistance genes (ARG) have been found in many environmental matrices, including soils, groundwater, surface water, and sediments. Animal feeding operations (AFO) and wastewater treatment plants (WWTP) are potential sources of ARG in rivers and/or are sources of antibiotics that may select for ARG from native river bacteria. In this study, we aim to identify ARG profiles that can characterize potential sources of ARG including the native river. To this end, three WWTPs, six AFO lagoons, three sites along a pristine region of the Cache la Poudre River (PR), and a wildlife fish rearing unit (FRU) were compared with respect to the distribution, levels, and phylogenetic diversity of their ARG profiles. From correspondence analysis of the distributions of ARG in the source environments and pristine PR, tet(H), tet(Q), tet(S), and tet(T) were found to indicate AFO influence, while high detection frequencies of tet(C), tet(E), and tet(O) were more typical of WWTP profiles. Sul(I) was detected in 100% of samples from source environments, but just once in the pristine river environment. The ARG profile of the native PR was dominated by tet(M) and tet(W), demonstrating their presence in an environment does not indicate anthropogenic disturbance. The tet(W) gene libraries from Pristine PR, WWTPs, and AFO lagoons, were each unique, as determined by both restriction fragment length polymorphism (RFLP) and phylogenetic analysis. The libraries of different

[1]This manuscript will be submitted to *Environmental Science & Technology* with the following co-authors: M. Arabi, B. Crimi, J.G. Davis, and A. Pruden.

WWTPs were fairly similar, but the AFO libraries were quite diverse. These methods and findings will be highly valuable for the development of source-tracking tools and to characterize environmental transport of ARG.

2.1 Introduction

Antibiotic resistance is recognized by the World Health Organization (WHO) and the Center for Disease Control (CDC) as one of the main public health challenges of our time (WHO 2000, CDC 2001). As antibiotic resistance infections (ARI) have been increasing, the numbers of new antibiotics brought to market has been steadily decreasing, from sixteen new antibiotics in the early 1980s to just four since 2003 (Talbot et al. 2006). Since new drug discovery is no longer capable of keeping pace with emerging ARI, strategies for prolonging the efficacy of available antibiotics and reducing the spread of antibiotic resistance must be developed (WHO 2000, CDC 2001, Talbot et al. 2006, Colin et al. 2007). It has been shown that environments with heavier anthropogenic influence tend to have increased levels and more unique types of antibiotic resistance genes (ARG) (Pei et al. 2006, Pruden et al. 2006, Auerbach et al. 2007, Koike et al 2007, Hu et al. 2008). Still, the primary mechanism(s) involved in the proliferation and spread of ARG via environmental pathways has not been determined.

When antibiotics are consumed, up to 75% of the antibiotic passes through the body unmetabolized into the feces (Chee-Sanford et al. 2001). These antibiotics can then enter the natural environment through a variety of pathways, including point discharges from wastewater treatment plants (WWTP), animal feeding operations (AFO), and fish hatcheries, or via non-point sources such as overland flow from fields where manure or biosolids have been applied. Antibiotics have been detected at part per trillion (ppt) concentrations in many natural and engineered environments, such as surface waters and sediments, municipal WWTP, AFO lagoon

systems, groundwater underlying these lagoons, and fields fertilized with animal manures or biosolids (Chee-Sanford et al. 2001, Kolpin et al. 2002, Kümmerer and Henninger 2003, Thorsten et al. 2003; Yang and Carlson 2003, Pei et al. 2006, Batt et al. 2006a, Batt et al. 2006b, Kim and Carlson 2007b).

The presence of antibiotics, even at sub-inhibitory concentrations, may drive natural selection of ARG from native river bacteria. Alternatively, we propose that non-native bacteria possessing ARG, or ARG as free DNA, may themselves be transported to surface waters via similar pathways as antibiotics. Once present in rivers, ARG are capable of being transferred among bacteria, including pathogens, through horizontal gene transfer. The fate and transport of ARG is expected to be unique relative to other contaminants considering that they may be amplified in the presence of selecting agents such as antibiotics, transferred between diverse types of bacteria, and exist as either intracellular or extracellular entities.

Culture-independent analyses of potential source environments provide an effective approach for determining the structure of the ARG profiles in diverse environments. Previous studies of ARG profiles in source environments are frequently limited to pure cultures representing either human or animal fecal sources. To date, there has been no in-depth analysis of ARG profiles in pristine river environments. An understanding of the background ARG occurrence is crucial for determining if ARG present at impacted sites are selected from native river sediments or are transported from source environments. Additionally, differentiation of human and animal sources of ARG can shed light on areas where intervention may be most effective in helping reduce the spread of ARG contaminants through natural environmental matrices.

The goal of this study was to identify ARG profiles that are effective for characterizing potential sources of ARG as well as native river environments. To this end, three specific objectives were defined: first, to characterize the background level of resistance in pristine river water and sediments; second, to characterize putative source environments with respect to the distribution, frequency of detection, and phylogenetic diversity of their ARG profiles; finally, to determine if certain ARG or genetic signatures of ARG are suitable indicators of either urban or agricultural influence. Ultimately, the approach is expected to be of broad value for identifying sources of antibiotic resistance input to target environments. This may then provide the basis for effective mitigation strategies.

2.2 Experimental Section

The Cache La Poudre River (PR) in Northern Colorado offers a unique opportunity to analyze a fairly undisturbed river ecosystem. It has a pristine source arising from snow melt in the Rocky Mountains with few tributaries or anthropogenic disturbances. Riverine samples from sites along this undisturbed system were analyzed and compared with samples from WWTP facilities, AFO lagoons and a fish hatchery for presence/absence of eleven ARGs. The sampling effort entailed multiple points in time in order to capture seasonal variations and potential impact(s) of low/high flow conditions. Correspondence analysis was employed to identify ARG profiles most suitable for characterizing a site as pristine, AFO, WWTP, or FRU.

2.2.1 Pristine environment: Poudre River water and sediment samples

The upstream PR analyzed in this study lies within the Roosevelt National Forest. A map of the land use in the upstream PR region along with the sampling sites is available in supplementary information (Figure A1.1). The PR watershed has been analyzed with respect to both antibiotics and ARG in previous works (Yang and Carlson 2003, Yang and Carlson 2004, Pei

et al. 2006, Pruden et al. 2006, Kim and Carlson 2006, 2007a, Kim and Carlson 2007b). To date, no antibiotics have been detected at PR site 1, though ARG were found to be present (Yang & Carlson 2003, 2004; Kim & Carlson 2006, 2007a, 2007b). Thus, PR1 and two sites further upstream located at Elephant Rock landmark, designated PR0a, and Profile Rock landmark, designated PR0b, in Roosevelt National Forest were chosen for analysis in this study. River water and sediment (sed) samples were collected three times during low flow: October 2006, February 2007, and October 2007 and once during high flow conditions: May 2007.

Water was collected in sterile 1 L Nalgene bottles from the center of the river.. A metal spade was used to collect ~0.5 kg of sediment from the top 5 cm of the river floor. The sample was mixed and a portion was collected in sterile 50 mL centrifuge tubes. Samples were preserved on ice for transport to the laboratory (max 12 hrs).

2.2.2 Source environments: Animal Feeding Operation (AFO) lagoons, municipal Wastewater Treatment Plants (WWTP), and Fish Rearing Unit (FRU)

Water column and settled solids from AFO lagoons were collected during sampling events coordinated with river sampling events: October-November 2006, February-March 2007, May 2007, and October 2007. The AFO lagoons (small organic dairy (SOD), large organic dairy (LOD), small conventional dairy (SCD), large conventional dairy (LCD), and two beef feedlots (BF1, BF2)) analyzed in this work have been described previously (McKinney et al. 2009).

Three municipal WWTPs were sampled in April 2008 and April 2009, when 250 mL of activated sludge (as) from each facility and 50 mL biosolids from WWTP1 were collected. In April 2009, 250 mL of activated sludge and 500 mL effluent was collected at each WWTP. 50-mL of biosolids were again collected from WWTP1. All samples were stored on ice for transport. Two of the wastewater treatment plants, WWTP1 and WWTP3, treat nearly 100% municipal

wastewater; whereas WWTP2 also receives 15% commercial and 20% light industrial wastewater. A summary of the sites analyzed in this study is presented in Table 2.1.

A wildlife fish rearing unit (FRU) was the only known potential point source of ARG in the pristine region of the PR sampled in this study. To determine if it had any influence on the native ARG profile, discharge from the FRU was sampled in October 2008. Water from the discharge stream and the underlying sediment were collected from two separate raceway discharges of the FRU.

In the text, the term "site" refers to all samples (water column/sediment, suspended sediment/sediment, or activated sludge/biosolids/effluent) from a specific geographical location at all time points. The term "environment" refers to all sites of a particular classification. Classification of environments is as follows: upstream PR (3 sites): PR0a, PR0b, PR1; AFO lagoons (6 sites): LBF1, LBF2, LCD, LOD, SCD, SOD; WWTPs (3 sites): WWTP1, WWTP2, WWTP3; and the FRU (1 site).

2.2.3 Sample preparation and DNA extraction

Suspended sediments from water were collected via filtration through 0.45μm filter membranes (Millipore), placed in sterile 10 mL Petri dishes, and frozen prior to DNA extraction. Filters were cut into small pieces and added directly to extraction tubes. 1 g of sediment samples was extracted in duplicate using the UltraClean Soil DNA Kit (MoBio Laboratories, Inc.) according to manufacturer's instructions. Lagoon and activated sludge samples were centrifuged at 6000 rpm for 10 min. 0.25 g of the sludge pellet was extracted using the PowerSoil DNA Kit (MoBio Laboratories, Inc.) according to manufacturer's instructions. DNA quality was verified by agarose gel electrophoresis. All DNA extracts were diluted 1:10 with sterile water to be used as template for PCR assays.

2.2.4 Polymerase Chain Reaction (PCR) assays

To measure the distribution and frequency of ARG in source and pristine environments, the presence of eleven tetracycline resistance genes and two sulfonamide resistance genes was determined using primers and PCR assays as described previously (Aminov et al. 2001, Aminov et al. 2002, Pei et al. 2006, Koike et al. 2007). A typical reaction mixture (25 µL total volume) for PCR of tet genes contained: 1X ExTaq Buffer (Takara Bio Inc., Pittsburgh, PA), 0.5 µM each primer, 0.2 µM each dNTP, 1.25 units ExTaq polymerase and 1- µL template. The presence of two sul genes, sul(I) and sul(II), was determined using PCR assays described in Pei et al. (2006). The reaction mixture (25-µL total volume) contained 5 µL 5X buffer, 2.5 µL 10X buffer, 0.2 µM dNTPs, 0.2 µM each primer, 1.75 units of Taq polymerase and 1- µL template. The thermoprofile for all the genes (except tet(H)) was as follows: initial denaturing at 94°C, followed by 40 cycles of 15s at 94°C, 30 s at the annealing temperature (see Table S1), 30s at 72°C, and a final extension step for 7 min at 72°C. The thermoprofile for tet(H) has been described previously (Koike et al. 2007): initial denaturing at 94°C, followed by 40 cycles of 94°C for 30s, 66°C for 30s , 72°C for 30s, followed by 10 cycles of 94°C for 30s, 63°C for 30s, 72°C for 30s, and a final extension at 72°C for 10 minutes. 10 µL aliquots of PCR products were mixed with 1X SybrGreen loading dye and analyzed by electrophoresis on 2% agarose gels (w/v). Negative and positive controls were analyzed with each PCR assay. In addition, samples in which the ARG of interest was not detected were randomly subject to a spiked matrix test to determine if PCR inhibitions may have contributed to false negatives. Spiked controls were 2-3 orders of magnitude greater than the limit of detection (1000 genes). Spiked controls were not inhibited in any of the matrices examined (all sample matrices were tested for inhibition of at least one gene).

2.2.5 Genetic profiling of tet(W) with shot-gun cloning and Restriction Fragment Length Polymorphism (RFLP) analysis

To assess phylogenetic diversity of ARG in source environments, an 1152-bp region of the *tet*(W) gene was amplified via PCR as described by Koike et al. (2007) with primer (2pM) and Mg^{2+} (4mM) concentration optimized for the DNA matrices of all samples. PCR products were then cloned into the pCR1™ vector using the Topo-TA™ cloning kit. A restriction digest assay for the amplicons was designed with NEBcutter V2.0 (New England Biolabs, Inc.) based on sequences with the following accession numbers: DQ309636, DQ309637, DQ309647, DQ309651, DQ309659, DQ309659, DQ309667, DQ309687, DQ309688, DQ309691, and five sequences cloned from organic dairy sludge. Clones were screened for diversity using a BstUI digest with a typical reaction mixture (20µL): 2 µL m13 PCR product, 18 µL New England Biolabs Buffer 2, and 1 unit BstUI restriction enzyme. Digests were carried out at 60°C for 1 hr and visualized on a 2% agarose gel.

Selected clones were sequenced in both directions using the T3 and T7 primers by Colorado State University's Proteomics and Metabolomics facility. Large amplicons were assembled using ChromasPro 1.41 software (Technelysium Pty Ltd.). Phylogenetic and molecular evolutionary analyses, including sequence alignments and tree construction, were carried out using *MEGA* version 4 (Tamura et al. 2007).

2.2.6 Statistical analyses

The frequency of detection (FOD) of each gene at a geographical location for all time-points was calculated as the total number of detections (g^+) of that ARG divided by the total possible number of detections of that ARG (FOD = $ARG^+/(ARG^+ + ARG^-)$. Frequency of detection of 11 ARG at each of the geographical locations (six AFO lagoons, three WWTPs, and three pristine

River sites) was analyzed by correspondence analysis, a type of ordination method used to determine associations between variables (Storfer et al. 2007) using SAS 9.1 (SAS Institute Inc., Cary, NC) to determine correlations between ARG and their relationship to environments.

For the purpose of assessing phylogenetic similarity between the tet(W) clone libraries of source environments, alignments of sequences generated from the clone libraries of AFO lagoons WWTP activated sludge, and pristine river sediment were used to generate neighbor-joining trees using MEGA 4 (Tamura et al. 2007). Trees were rooted to the elongation factor gene, fusA, of *Aquifex aeolicus* and used for performing similarity analysis between environments using the online environmental analysis software UniFrac (available at <http://bmf2.colorado.edu /unifrac/index.psp>) (Lozupone et al. 2006). Principal component analysis, UniFrac significance tests, and jackknife cluster analysis were performed using abundance weights, determined as the number of identical sequences for each environment (Lozupone and Knight 2005, Lozupone et al. 2007).

2.3 Results and Discussion

2.3.1 Presence and distribution of ARG determined by PCR

Forty-seven AFO lagoon, eleven WWTP, twenty-four upstream PR, and four hatchery samples were screened by PCR for the presence of four tetracycline efflux genes: tet(B), tet(C), tet(E), tet(H); seven ribosomal protein protection (RPP) genes: tet(M), tet(O), tetB(P), tet(Q), tet(S), tet(T), tet(W); and two sulfonamide resistance genes: sul(I), sul(II) (Tables A1.2a and A1.2b, Appendix 1).

The AFO lagoon environment was observed to be the most diverse with each ARG being detected in at least two samples. The tet(B) and tetB(P) genes were detected in only one lagoon. The tet(B) gene was detected twice at the BF2 site and tetB(P) twice in the SCD site. Since these

genes were found at only one site, they are unlikely to be good indicator genes and were excluded from downstream analysis. For the eleven remaining ARG, the FOD was highest in AFO lagoons (0.82), followed by WWTPs (0.66), the FRU (0.30), and the native PR (0.09) (Table 2.1). The sul(I) gene was detected in 100% of AFO lagoon, WWTP, and FRU samples, but only once in the upstream river at site PR0b, immediately downstream of the FRU. Thus, presence of sul(I) may be an indicator of influence from a non-native environment.

While the above observations can provide general comparisons, a more robust method is necessary to determine the distribution of the eleven ARG in the various source environments. Moreover, understanding the relationships between the various genes will be useful in guiding further analysis and ultimately classification of source environments. To this end, correspondence analysis was performed by comparing the pooled number of detections of each ARG at each site in the pristine PR, AFO lagoon, and WWTP environments based on the frequency of detection of eleven ARG (Figure 2.1).

The first and second dimensions combined explain 72.7% of the total variation in the dataset. Pristine PR, WWTP and AFO lagoon environments were not dispersed but grouped together within their respective environments, indicating the uniqueness of the ARG profile to an environment. Vectors from the origin indicate the influence, or loading, of the eleven ARG on the variation in the data. ARG further from the origin (longer vectors) have the greatest contribution to the variation in the data, and thus could indicate genes with the potential for classification. Genes that are positively correlated will lie along the same axis. WWTPs and AFO lagoons were separated along dimension 1. The loading of tet(S) was clearly associated with the AFO lagoons. In fact, tet(S) was detected at least once in each AFO lagoon, but was not detected in any other environment. While this difference would be apparent from simple examination of the raw data, other patterns would be less likely to emerge in absence of some

level of dimension reduction. For example, tet(E) was the other major contributor to the variation in the data and it was loaded in the direction of all three WWTPs. In the AFO environment, tet(E) was detected in each dairy lagoon, but not in the beef feedlot lagoons. The overall frequency of detecting tet(E) was just 0.28 in the AFO environment. With the exception of tet(S) (FOD=0.49), all other genes were detected in at least 75% of the AFO samples. Tet(E) was detected in each of the three WWTPs, with a higher FOD (0.46) than AFO lagoons. Similarly, tet(C) was loaded in the same direction as tet(E), and its FOD was 0.91 in WWTP compared to 0.77 in AFOs. The RPP genes, tet(Q) and tet(T) appear to be positively correlated and predictors of AFO influence, being found in 87% (tet(Q)) and 96% (tet(T)) of AFO samples, but detected in just 45% of WWTP samples. Of the genes loaded in the direction of the upstream PR sites, tet(W), tet(M), and sul(II) genes account for 33%, 25%, 16% of the genes detected in the upstream PR, suggesting that these genes are native to the pristine river environment.

These results demonstrated the power of correspondence analysis for exploring the differences in the distribution and frequency of ARG profiles of different environments, especially with large datasets. Most importantly, correspondence analysis can be used to determine which genes are the strongest indicators of those differences. The tet genes tet(H), tet(S), tet(Q), and tet(T) are strongly correlated and loaded with AFOs. Similarly, the tet genes tet(C), tet(E), and tet(O) were associated and more indicative of WWTP sources. The potential to use these groups of ARG as indicators of WWTP and AFO sources is demonstrated when the average normalized frequency of detection of the tet(H), tet(Q), tet(S), and tet(T) (tet group HQST) genes was plotted against the normalized frequency of detecting tet(C), tet(E), and tet(O) for pristine PR, AFO, and WWTP environments (Figure 2.1B). Removing ARG that did not explain a large fraction of variability in the data resulted in enhanced separation of AFO, WWTP and

pristine PR environments. We hypothesize that these genes will be valuable for future classification of impacted environments according to agricultural and urban influence.

2.3.2 Comparison of upstream PR environment to the FRU environment

Due to the low level of anthropogenic influence in the upstream region of the PR, this environment was expected to posess a less diverse ARG profile. In total, eight of the eleven ARG were detected (tet(E), tet(Q), and tet(S) were not detected). Because of the location of the FRU within the upstream PR watershed, its possible impact on the river was also explored. Commercial aquaculture facilities are well known to have an intense impact on the surrounding environment, such as the release of ammonia and other pollutants (Cabello 2006, Cao et al. 2007). However, the FRU is a wildlife rearing unit, and likely to have stricter regulations concerning antibiotic use. Still, the FRU environment was distinct from the upstream PR in both overall detection and distribution of ARG (Figure 2.2). Whereas the upstream PR was dominated by tet(M) and tet(W) genes, tet(M) was detected just once in the FRU and no other tet RPP genes were detected (FOD of tet RPP genes = 0.04). In contrast, the sul(I), tet(C), and tet(H) genes were the most commonly detected ARG in the FRU. It is noteworthy that although these genes were also detected in the pristine PR, they were only found at site PR0b, located immediately downstream of the FRU, suggesting influence from the FRU at this site. Yet, the ARG profile of PR1, the site located ~50 km downstream from site PR0b, exhibits greater similarity to that of PR0a, than of PR0b. Thus, it is likely that any impact of the FRU on the ARG profiles in the river attenuates as the river flows downstream through pristine wilderness to the mouth of the canyon. These observations are in agreement with published literature regarding the predominance of tetracycline efflux genes in similar aquatic environments (Henriques et al. 2008). In addition, tet(C), which was predominant in the FRU, is commonly found in *Aeromonas*, a genus including fish pathogens (e.g. *Aeromonas hydrophila*) (Roberts 2005).

The tet(M) and tet(W) genes were frequently detected in the pristine PR, AFO, and WWTP environments. The literature cites that the tet(M) gene has been found in 42 genera, and tet(W) in 20 genera (Roberts 2005). These genes are both located on conjugative transposons—mobile genetic elements that code for their own transfer, can integrate themselves into a host's chromosomal DNA, and can quickly pass from organism to organism (Roberts 1994, Roberts 2005). These factors may contribute to their widespread distribution across environments as described here and elsewhere (Roberts 2005). It is evident that the tet(M) and tet(W) genes are ubiquitous in environments both with and without anthropogenic disturbances, and therefore their presence in an environment offers little information about the gene's source. However, their ubiquity across diverse environments can make them ideal ARG for elucidating the phylogenetic in the ARG among diverse environments. The tet(W) gene has been analyzed for diversity in a previous work (Koike et al. 2007). To examine these relationships among the source environments, the phylogeny of the tet(W) was further analyzed using restriction fragment length polymorphism (RFLP) and nucleotide sequencing.

2.3.3 Restriction Fragment Length Polymorphism analysis of tet(W)

A 1152-bp region of the tet(W) gene was amplified in six AFO samples and three WWTP samples. Clones from AFO lagoon samples (n = 311) and clones from WWTP activated sludge (n = 162) were generated, and RFLP analysis was performed. There were four predominant restriction patterns conserved between the AFO lagoons and WWTP samples (pattern 1, pattern 2, pattern 3, pattern 4) that were easily identifiable between libraries and accounted for 95% of all restriction patterns. All other unique patterns were screened with each other for further classification. Correspondence analysis was performed with the number of the conserved restriction patterns (normalized to total clones from each library) (Figure 2.3a). There is a clear distinction between the AFO and the WWTP tet(W) clone libraries. The WWTP libraries appear

highly similar, as they are found in only the lower right quadrant, whereas the AFO lagoon clone libraries are scattered across the other three quadrants, suggesting greater diversity within the AFO environment. In terms of loadings of the restriction patterns, pattern 2 appears to be an indicator of AFO lagoons and accounts for 17% of AFO clones but only 1% of WWTP clones. Pattern 1 is loaded in the direction of the WWTP libraries and represents 9% of all WWTP clones versus 3% of AFO clones. RFLP patterns 3 and 4 exhibit a split loading between the two environments and are negatively correlated (loaded in opposite directions). These patterns are not strong indicators of either environment and are less useful for classification of AFOs and WWTPs. Plotting the percent distribution of RFLP patterns 1 and 2 is equally informative for discrimination between source environments (Figure 2.3b). Environments similar to WWTP are predicted to have a higher proportion of clones possessing pattern 1 than 2, and those similar to AFOs are predicted to correspond to a higher proportion of clones with pattern 2 than 1.

From the initial sequencing of the libraries, clones with RFLP pattern 3 were fairly similar, while those of pattern 1, 2, and 4 exhibited more diversity. To determine the full measure of diversity in the samples, 100% of clones exhibiting patterns 1, 2, and 4, and 50% of clones with pattern 3 were sequenced.

2.3.4 Phylogenetic analysis of the tet(W) gene

Though presence of tet(W) in pristine environments was demonstrated in this study and others (Pei et al. 2006), amplifying the 1152-bp region of the gene in upstream PR samples was hindered, likely due to several factors such as matrix effects, low copy number of the gene, or damaged template DNA. From the alignment of 223 tet(W) sequences from the six AFO libraries and 107 sequences from three WWTP libraries, a tree was constructed using the neighbor-joining method in MEGA version 4 (Tamura et al. 2007). The sequence of the *Aquifex aeolicus*

fusA gene, coding for translation elongation factor EF-G, the proposed ancestor of RPP tetracycline resistant genes, was used as the out-group for rooting the tree (Aminov and Mackie 2007). The tree was further analyzed using UniFrac (Lozupone et al. 2006) to compare the libraries from the nine environments. UniFrac is designed to compare profiles based on the combined shared branch length of the phylogenies and is useful for understanding trees generated with large numbers of sequences generated from two or more different environments (Lozupone and Knight 2005, Lozupone et al. 2007).

When all six AFO lagoons were considered as one environment, the AFO lagoon environment was significantly different ($p < 0.01$) from the WWTP environment (activated sludge from all three WWTPs) according to the UniFrac Significance test. While the WWTP libraries appear highly similar, AFO lagoon libraries displayed notable divergence from each other, as indicated by the jackknife cluster analysis of the clone libraries (Figure 2.4) and the RFLP patterns (Figure A1.3) produced similar trees. In both trees, WWTPs form a cluster with WWTP1 and WWTP3 having the greatest similarity. Also, the small and large conventional dairies were shown to be similar using both methods. Since the phylogenetic analysis supports the results of the RFLP analysis, the latter may have good potential for application in absence of sequence information to discriminate between sources and provide a means of classifying impacted environments.

The striking divergence of AFO-derived ARG from other sequences in the clone library and from reference sequences has been noted previously (Koike et al. 2007, Yu et al. 2005). A clone derived from suspended solids in BF1 lagoon and a clone from the water column of the SCD lagoon were particularly divergent from other clones. These clones shared 99% and 96% identity, respectively, to the tet(W) gene of a gram-negative animal pathogen, *Lawsonia intracellularis* that is typically transmitted through environmental contamination with feces

from infected animals (Collins et al. 2000). The most divergent clones shared closest identity with other uncultured bacterial clones isolated from swine feedlots characterized by Koike et al. (2007) and AFOs characterized by Yu et al. (2007).

There were two groups of consensus sequences (clones that when sequenced were 100% identical). An identical group of 15 clones (3 WWTP, 12 AFO) shared 100% homology with the tet(W) reference gene of *Bifidobacterium* sp. ISO3519, a high-GC gram-positive bacteria. Another identical group of 14 clones (8 WWTP, 6 AFO) shared complete identity with the tet(W) gene found in the gram-positive rumen anaerobe *Butyrivibrio fibrisolvens*, the gram-negative commensal human-colonic anaerobe *Mitsuokella multacida,* and the gram-negative rumen anaerobe *Megasphaera elsdenii.*

The greatest diversity (divergence) was seen in the libraries derived from the small and large conventional dairy water column samples and beef feedlot 1 settled solids sample.. The clone libraries of the organic dairies exhibited less divergence from the other environments. This could be explained by the existence of greater selection pressures on conventional dairies and feedlots compared to organic farms due to more prevalent use of antibiotics (for therapeutic reasons in dairies and for growth promotion in feedlots). In fact, this hypothesis is in line with the trends seen in tetracycline (TC) concentration in the lagoons (McKinney et al. 2009). TCs were detected in organic dairy lagoons ([total TC] = 5 – 116 μg/kg dry weight) but at lower concentrations than in conventional dairy lagoons ([total TC] = 213 - 2300 μg/kg dry weight). Total TC concentration in the beef feedlot 1 lagoon was in between that of the organic and conventional dairies ([total TC] = 97 – 599 μg/kg dry weight) (McKinney et al. 2009).

To determine how native tet(W) genes compared to the genes found in WWTPs and AFO lagoons, a 167bp region of the tet(W) gene amplified from PR0b-sed in the PCR screening study

was cloned and sequenced for comparison to the existing clone library of the larger tet(W) amplicons. To reduce bias that could have been introduced by the use of a different reverse primer, long sequences with mismatches within the priming region (148-167bp) were eliminated from analysis. The remaining sequences (129) were aligned and trees created as described above. In total ten clones from the PR0b-sed sample were sequenced and represented two unique sequences. Jackknife clustering was performed using UniFrac (Figure 2.5). Despite the small sample size of the PR0b tet(W) clone library, the nodes of the tree were recovered more than 99% of the time, indicating there was sufficient sample size to determine the relationships among environments. The WWTP and AFO environments were more similar to each other than to the Pristine, PR0b clone library. When the PR0b sequences were compared to the entire clone library built in this study and to the nucleotide (nr/nt) database using the Basic Alignment Search Tool (BLAST) (available online: <http://blast.ncbi.nlm.nih.gov/Blast.cgi>), no sequences sharing 100% identity were found.

Though site PR0b appears to be influenced by the FRU, the PCR distribution study and of the FRU have shown that its ARG profile was dominated by tetracycline efflux and sulfonamide resistance genes, rather than RPP tet genes, such as tet(W). In fact, only one tet RPP gene tet(M) was ever detected at the FRU site. And in Q-PCR analysis, which has a lower limit of detection than traditional PCR, tet(W) was detected only once in the FRU at very low concentrations just above the limit of quantification (data not shown). Therefore, it is unlikely that the diversity of the PR0b sediment sequences could be explained solely by influence or selection by the FRU.

2.4 Summary

In this study, we were the first to examine urban (WWTPs), agricultural (AFOs), and pristine river environments within the same watershed using methods of pattern detection and

environmental analysis with the purpose of developing methods for the source-tracking of ARG in impacted river watersheds. This approach can be applied to impacted environments to help determine sources of ARG and to identify dominant processes of ARG propagation. Classification of environments using these methods may be built upon using geospatial analysis of the surrounding watershed to determine which variables are effective for explaining the distribution, quantities, and phylogenetic diversity of ARG profiles in the river environments.

REFERENCES

Aminov RI and Mackie RI (2007) Evolution and ecology of antibiotic resistance genes. *FEMS Microbiology Letters,* 271: 147-161.

Aminov, RI, Chee-Sanford, JC, Garrigues, N, Teferedegne, B, Krapac, IJ, White, BA, Mackie, RI (2002) Development, validation, and application of PCR primers for detection of tetracycline efflux genes of gram-negative bacteria. *Applied and Environmental Microbiology,* 68: 1786–93.

Aminov, RI, Garrigues-Jeanjean, N, and Mackie, RI (2001) Molecular ecology of tetracycline resistance: Development and validation of primers for detection of tetracycline resistance genes encoding ribosomal protection proteins. *Applied and Environmental Microbiology*, 67: 22-32.

Auerbach, EA, Seyfried, EE, McMahon, KD (2007) Tetracycline resistance genes in activated sludge wastewater treatment plants. *Water Research,* 41: 1143-1151.

Batt, AL, Bruce, IB, Aga, DS (2006a) Evaluating the vulnerability of surface waters to antibiotic contamination from varying wastewater treatment plant discharges. *Environmental Pollution,* 142: 295-302.

Batt, AL, Snow, DD, Aga, DS (2006b) Occurrence of sulfonamide antimicrobials in private water wells in Washington County, Idaho, USA. *Chemosphere,* 64: 1963-1971.

Cao, L, Wang, W, Yang, Y, Yang, C, Yuan, Z, Xiong, S, Diana, J. (2007) Environmental Impact of Aquaculture and Countermeasures to Aquaculture Pollution in China. *Environmental Science Pollution Research*, 14 (7): 452–462.

CDC, Centers for Disease Control. (2001) National Nosocomial Infections Surveillance (NNIS) System Report, Data Summary from January 1992-June 2001, Issued August (2001). *American Journal of Infection Control* 29: 404-421.

Chee-Sanford, JC, Aminov, RI, Krapac, IJ, Garrigues-Jeanjean, N, and Mackie, RI (2001) Occurrence and diversity of tetracycline resistance genes in lagoons and groundwater underlying two swine production facilities. *Applied and Environmental Microbiology,* 67: 1494-1502.

Colin, TJ, Koch, K, Jost, K, Davis, OB, Dela Rosa, D. (eds). (2007) Fighting Superbugs: Are disease-resistant bacteria becoming unstoppable? *CQ Researcher*, 17 (29): 673-696.

Collins, A, Love, RJ, Pozo, J, Smith, SH, McOrist, S (2000) Studies on the *ex vivo* survival of Lawsonia intracellularis. *Swine Health and Production*, 8: 211-125.

Heepngoen, P, Sajjaphan, K, Ferguson, JA, Sadowsky, MJ (2008) Genetic and physiological characterization of oxytetracycline-resistant bacteria from giant prawn farms. *Journal of Microbiology and Biotechnology*, 18: 199-206.

Henriques, IS, Fonseca, F, Alves, A, Saavedra, MJ, Correia, A (2008) Tetracycline-resistance genes in Gram-negative isolates from estuarine waters. *Letters in Applied Microbiology*, 47: 526-533.

Hu, JY, Shi, JC, Chang, H, Li, D, Yang, M, Kamagata, YC (2008) Phenotyping and genotyping of antihiotic-resistant Escherichia coli isolated from a natural river basin. *Environmental Science & Technology,* 42 (9):3415-3420.

Kim, SC, Carlson, K (2006) Occurrence of ionophore antibiotics in water and sediments of a mixed-landscape watershed. *Water Research,* 40: 2549-2560.

Kim, SC, Carlson, K (2007a) Quantification of human and veterinary antibiotics in water and sediment using SPE/LC/MS/MS. *Analytical and Bioanalytical Chemistry,* 387: 1301-1315.

Kim, SC, Carlson, K (2007b) Temporal and spatial trends in the occurrence of human and veterinary antibiotics in aqueous and river sediment matrices. *Environmental Science & Technology,* 41: 50-57.

Koike, S, Krapac, IG, Oliver, HD, Yannarell, AC, Chee-Sanford, JC, Aminov, RI, Mackie, RI (2007) Monitoring and source tracking of tetracycline resistance genes in lagoons and groundwater adjacent to swine production facilities over a 3-year period. *Applied and Environmental Microbiology*, 73 (15):4813-4823.

Kolpin, DW, Furlong, ET, Meyer, MT, Thurman, EM, Zaugg, SD, Barber, LB, and Buxton, HT (2002) Pharmaceuticals, hormones, and other organic wastewater contaminants in US streams, 1999-2000: A national reconnaissance. *Environmental Science & Technology,* 36: 1202-1211.

Kümmerer, K, and Henninger, A (2003) Promoting resistance by the emission of antibiotics from hospitals and households into effluent. *Clinical Microbiology and Infection,* 9: 1203-1214.

Lozupone C and Knight R. (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, 71: 8228-8235.

Lozupone C, Hamady M, and Knight R. (2006) UniFrac - An Online Tool for Comparing Microbial Community Diversity in a Phylogenetic Context. *BMC Bioinformatics*, 7: 371.

Lozupone, C, Hamady, M, Kelley ST and Knight, R (2007) Quantitative and qualitative (beta) diversity measures lead to different insights into factors that structure microbial communities. *Applied and Environmental Microbiology*, 73: 1576-1585.

McKinney, CW, Loftin KA, Meyer, MT, and Pruden A (2009) Antibiotic Resistance Genes (ARGs) in Livestock Lagoons: Effect of Livestock Operation Type, Treatment Approach, and Seasonality, and Antibiotics on tet and sul. *Environmental Science & Technology,* in preparation.

Pei, R, Kim, S-C, Carlson, KH, and Pruden, A (2006) Effect of River Landscape on the sediment concentrations of antibiotics and corresponding antibiotic resistance genes (ARG). *Water Research,* 40: 2427-2435.

Pruden, A, Pei, RT, Storteboom, H, Carlson, KH (2006) Antibiotic resistance genes as emerging contaminants: Studies in northern Colorado. *Environmental Science & Technology,* 40: 7445-7450.

Roberts, MC (1994) Epidemiology of tetracycline-resistance determinants. *Trends in Microbiology*, 2: 353-357

Roberts, MC (2005) Update on acquired tetracycline resistance genes. *FEMS Microbiology Letters,* 245: 195-203.

Saitou N, and Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4:406-425.

Talbot, GH, Bradley, J, Edwards, JE, Gilbert, D, Scheld, M, Bartlett, JG (2006) Bad bugs need drugs: an update on the development pipeline from the Antimicrobial Availability Task Force of the Infectious Diseases Society of America. *Clinical Infectious Disease,* 42 (5), 657-668.

Tamura K, Dudley J, Nei M, and Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24:1596-1599.

Thorsten, C, Schneider, RJ, Färber, HA, Skutlarek, D, Meyer, MT, and Goldbach, HE (2003) Determination of Antibiotic Residues in Manure, Soil, and Surface Waters. *Acta hydrochimica et hydrobiologica,* 31: 36-44.

WHO, World Health Organization. (2000) WHO annual report on infectious disease: Overcoming antimicrobial resistance. available online at < http://www.who.int/infectious-disease-report/2000/

Yang, S, and Carlson, K (2003) Evolution of antibiotic occurrence in a river through pristine, urban and agricultural landscapes. *Water Research,* 37: 4645-4656.

Yang, S, Carlson, K (2004) Routine monitoring of antibiotics in water and wastewater with a radioimmunoassay technique. *Water Research,* 38: 3155-3166.

Yu, ZT, Michel, FC, Hansen, G, Wittum, T, and Morrison, M (2005) Development and application of real-time PCR assays for quantification of genes encoding tetracycline resistance. *Applied and Environmental Microbiology,* 71: 6926-6933.

TABLES

Table 2.1: Description of Environments

Environment	Site Name	Site ID	Details
AFO Lagoons	Beef Feedlot 1	LBF1	18,000 head cattle
	Beef Feedlot 2	LBF2	28,000 head cattle
	Large Organic Dairy	LOD	1800 cows
	Large Conventional Dairy	LCD	6000 cows
	Small Organic Dairy	SOD	150 cows
	Small Conventional Dairy	SCD	800 cows
WWTPs	Facility 1	WWTP1	0.5 million gallons per day
	Facility 2	WWTP2	14.3 million gallons per day
	Facility 3	WWTP3	1.5 million gallons per day
Hatchery	Fish Rearing Unit	FRU	50,000 trout
Pristine River	Cache La Poudre River, Site 0a	PR0a	Elephant Rock, Roosevelt National Forest
	Cache La Poudre River, Site 0b	PR0b	Profile Rock, Roosevelt National Forest
	Cache La Poudre River, Site 1	PR1	Greyrock trailhead, Roosevelt National Forest

Table 2.2: Frequency of detection of thirteen antibiotic resistance genes in AFOs, WWTPs, and the pristine river environments

	tet efflux genes				tet ribosomal protection genes							sul genes	
Environment	tet(B)	tet(C)	tet(E)	tet(H)	tetB(P)	tet(M)	tet(O)	tet(Q)	tet(S)	tet(T)	tet(W)	sul(I)	sul(II)
AFO Lagoons	0.04	0.77	0.28	0.89	0.04	1.00	0.85	0.87	0.49	0.96	0.96	1.00	0.94
WWTPs	0.00	0.91	0.45	0.36	0.00	0.91	0.91	0.45	0.00	0.45	1.00	1.00	0.82
Fish Rearing Unit	0.00	0.75	0.00	0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.25	1.00	0.25
Pristine River	0.00	0.08	0.00	0.04	0.00	0.25	0.08	0.00	0.00	0.04	0.33	0.04	0.17

FIGURES

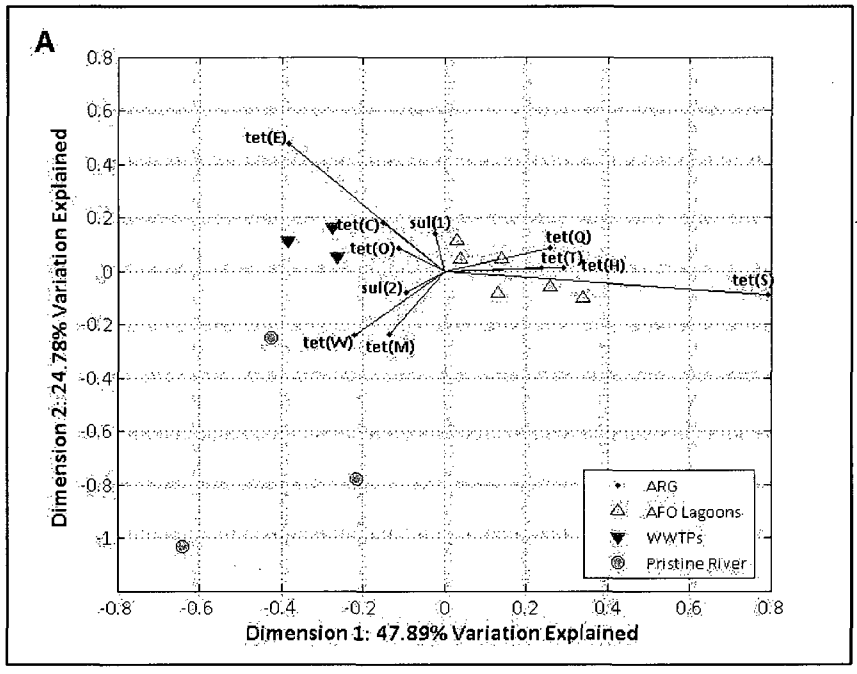

Figure 2.1: Correspondence analysis of ARG probabilities. (A)The number of positive detections of each gene within the AFO, WWTP, and Pristine Poudre River (PR) environments was subject to correspondence analysis. The x-axis represents the dimension (Dimension 1) explaining 47.89% of the total variation; the y-axis represents the second principal component (Dimension 2), explaining 24.78% of the total variation for 72.67% cumulative explained variability. Lines from the origin to the ARG points indicate the loading of the presence/absence of eleven ARG (tet(B) and tetB(P) were excluded from further analysis because each gene was detected in only one AFO lagoon). A classification tree based on the PCs was calculated and is overlaid on this plot. (B) Plot of ARG variables FOD of tet(H), tet(Q), tet(S), tet(T) vs FOD of tet(C), tet(E), tet(O).

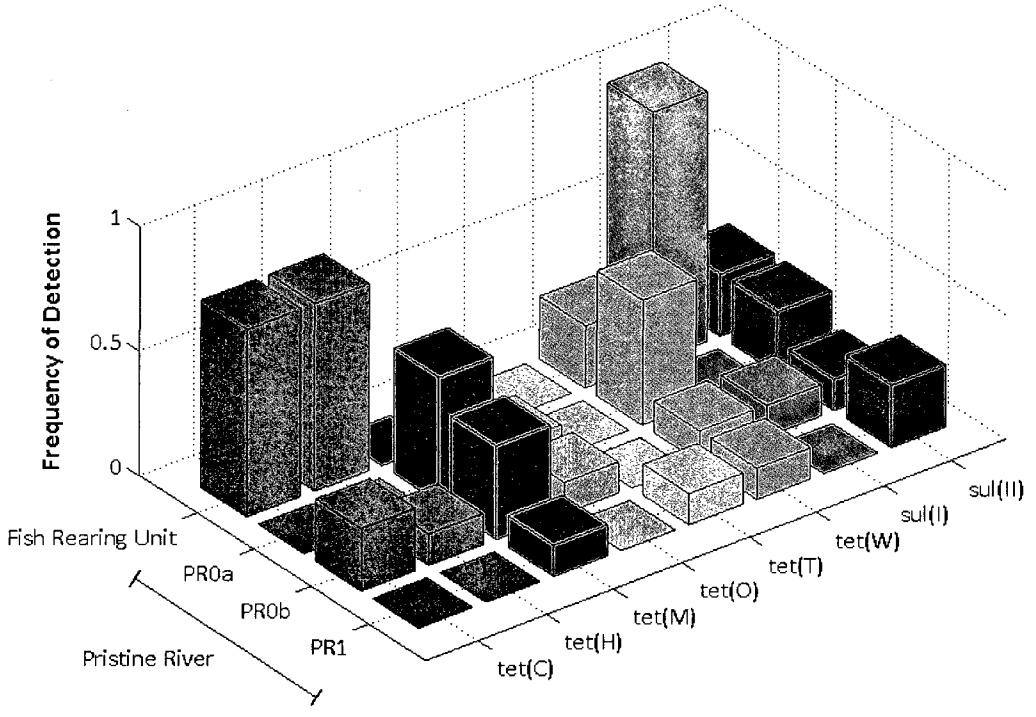

Figure 2.2: Comparison of total frequency of detection of tet(C), tet(H), tet(M), tet(O), tet(T), tet(W), sul(I), sul(II) in upstream PR sites and the FRU. The total frequency of detection was calculated as the sum of all detections of one ARG divided by the total number of possible detections.

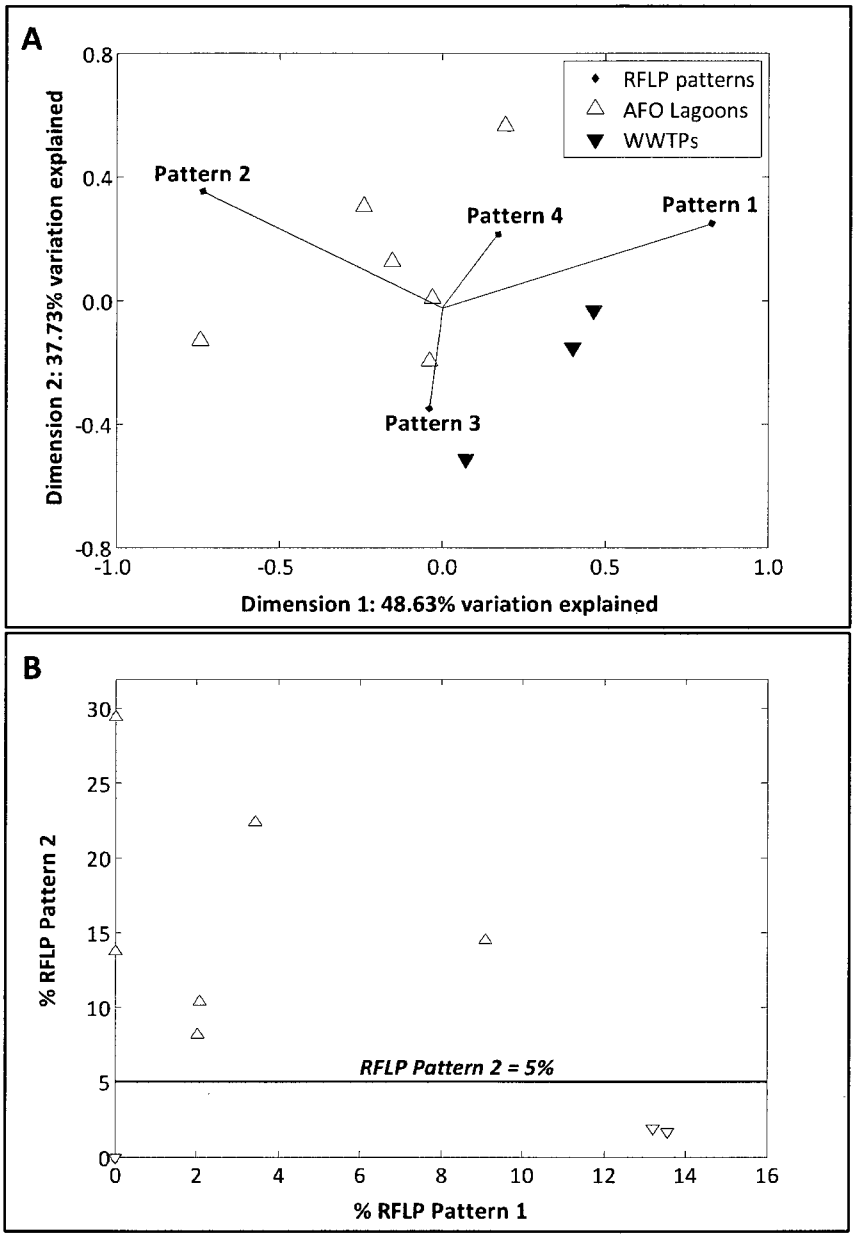

Figure 2.3: Correspondence analysis of RFLP patterns of tet(W) clones from AFO Lagoons and WWTP Activated Sludge. (A) The RFLP pattern counts of clones from the AFO and WWTP environments were subject to correspondence analysis. The x-axis represents the first principal component explaining 48.63% of the total variation; the y-axis represents the second principal component, explaining 37.73% of the total variation for 86.36% cumulative explained variability. Arrows indicate the loading of the presence/absence of the four conserved RFLP patterns. (B) Plot of patterns identified as useful for classification in correspondence analysis.

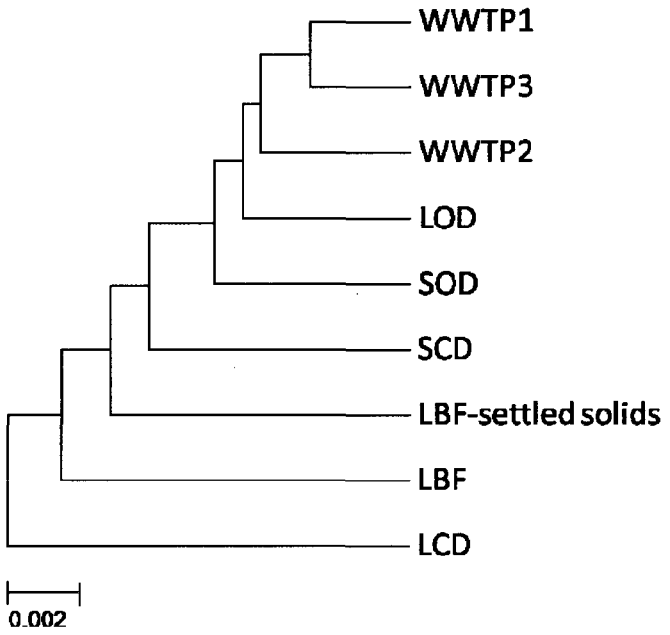

Figure 2.4: Jackknife cluster tree of the AFO and WWTP environments based on sequencing of 1152-bp amplicon of the tet(W) gene. Cluster tree of the clone libraries generated from the AFO and WWTP environments. Percentages at nodes represent the jackknife cluster support for each node.

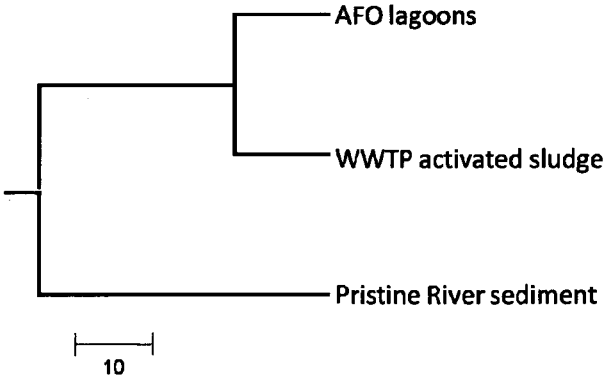

Figure 2.5: Jackknife cluster tree of the Pristine, AFO and WWTP environments based on sequencing of 167-bp amplicon of the tet(W) gene. Cluster tree of the Pristine, AFO and WWTP environments constructed with UniFrac. There was >99% jackknife cluster support for tree nodes.

CHAPTER 3: ANTIBIOTIC RESISTANCE GENES (ARG): A LANDSCAPE PERSPECTIVE[1]

Abstract

Antibiotics and their corresponding antibiotic resistance genes (ARG) have been found in many environmental matrices, including soils, groundwater, and surface water and sediments, but the dominant sources and mechanisms governing their transport are still unknown. Samples from the Cache la Poudre River (PR) and South Platte River (SPR) in Northern Colorado were characterized with respect to the distribution, levels, and diversity of their ARG profiles. On the basis of ARG indicator variables derived in previous work, most river samples were classified as wastewater treatment plant (WWTP) influenced using the methods of discriminant analysis. Classification and regression tree (CART) analysis was used to determine the relationship between spatial explanatory variables and the frequency of detection of ARG indicator variables. There was good agreement between the classification of river sites according to spatial variables and source indicator variables, demonstrating the effectiveness of these indicators in source-tracking ARG. Multivariate linear regression of sul(I) showed significant correlation with inverse-distance weighted (IDW) animal feeding operation (AFO) count (R^2 = 0.83, p<0.0003), whereas tet(W) was not correlated with any explanatory variable tested. When compared to an existing clone library of tet(W) genes from AFOs and WWTPs, Site 4 in the PR (PR4) was significantly different from AFO environment (p<0.05); Site 3 in the SPR (SPR3) was not significantly different from either environment. The PR4 environment was most similar to that of WWTP activated sludge, while the SPR confluence (SPR3) showed equal similarity with AFOs and WWTPs. Thus, we establish a link between ARG indicator variables and

[1]This manuscript will be submitted to *Nature* with the following co-authors: M. Arabi, J.G. Davis, and A. Pruden.

spatial indicators and demonstrate that the ARG profiles of river samples were more similar to WWTPs than AFO lagoon or pristine river environments. Based on this work, transport of ARG from sources may be the dominant mechanism for ARG proliferation in riverine environments relative to selection of native bacteria.

3.1 Introduction

While it has been shown that more environments with heavy anthropogenic influence tend to have elevated levels and more unique types of ARG than undisturbed environments (Pei et al. 2006, Pruden et al. 2006, Aurbach et al. 2007, Koike et al 2007, Hu et al. 2008), the primary mechanism involved in the proliferation and spread of ARG has not been determined. We suggest two mechanisms as driving mechanisms for elevated ARG levels observed in impacted surface waters: (1) The presence of low levels of antibiotics, metals, and disinfectants causes native bacteria carrying ARG to be selected and/or causes an increase in the horizontal gene transfer (HGT) of ARG; or (2) Non-native bacteria possessing ARG are transported to surface waters via point-source pollution, such as WWTP effluents or runoff from feedlots, or non-point source pollution, such as runoff agricultural fields where manure or biosolids have been applied (Pruden et al. 2006, Baquero et al. 2008).

In our previous work (Chapter 2), we demonstrated that the presence of tet(C), tet(E), and tet(O) were more dominant in WWTPs than in AFO lagoons. Similarly, the frequency of detection of tet(H), tet(Q), tet(S), and tet(T) were strongly correlated; these genes were detected with greater frequency in AFO lagoons than in WWTPs. The sul(I) gene was present in 100% (n=62) of source samples (WWTPs, AFO lagoons, Fish Hatchery and Rearing Unit (FRU) discharge), but was detected just once in (n=24) pristine river samples assayed. The tet(M) and tet(W) genes were detected in nearly half of all pristine river sites and were not useful for

indicating influence by a source. However, due to its ubiquity and its potential for demonstrating a link showing selection of ARG from the native river, the tet(W) gene was amplified from source and pristine environments. In AFO lagoons, the tet(W) phylogenetic profile was diverse and distinct from the profile of WWTPs. It was indicated from an analysis of a smaller amplicon of tet(W) in pristine river sediments that the pristine river environment was unique from both source environments (Chapter 2).

Thus, we identified the unique attributes of the ARG profiles of pristine river, AFO lagoons, and WWTPs environments and demonstrated that these attributes could be used to successfully classify these environments (Chapter 2). The aim of the present study is to build on this work by determining the sources and mechanisms involved in the spread of ARG to a river environment. To this end, we address three objectives: (1) characterize and classify impacted river samples according to the distribution, frequency of detection and phylogenetic diversity of their ARG profiles using the methods developed in Chapter 2; (2) create a geospatial database of the study area to assist in developing variables that reflect the measure of urban and agricultural inputs; and (3) determine which spatial indicators are valuable in explaining the variations in the ARG profiles of river samples.

3.2 Experimental Section

3.2.1 Site description

The Cache La Poudre River (PR) is a highly-suitable model system for characterizing the sources and processes involved in the transport of ARG. Its source arises from snow melt in the Rocky Mountains. It has few tributaries and is well-zonated between pristine, agriculture and urban sources. Levels of many classes of antibiotics including tetracyclines, beta-lactams, macrolides, sulfonamides, and ionophores have already been characterized in the Poudre River

system (Yang & Carlson 2003, 2004a, 2004b; Kim & Carlson 2006, 2007a, 2007b). The PR is a tributary of the South Platte River (SPR) in northeastern Colorado. In contrast to the PR, the SPR is classified as an impaired water body according to the Environmental Protection Agency. The SPR receives major influences from wastewater (WW) discharges as it flows through the Denver metropolitan area; during low-flow, WW effluent can approach 90% of the river flow. As the SPR flows northeast away from Denver, the land use of the watershed becomes dominated by concentrated animal feeding operations (CAFOs) and cropland (Figure 3.1). A summary of the river sites analyzed in this study is presented in Table 3.1.

3.2.2 Sample collection of river water and sediments

Five sites along the PR and three sites along the SPR were selected as river environments representing various degrees of anthropogenic influence. Of these impacted river sites, four sites: PR2, PR3 (referred hereto as PR3b), PR4 and PR5, were described previously (Pei et al. 2006, Yang and Carlson 2003). An additional site (PR3a) directly downstream of the Mulberry WWTP, between PR2 and PR3b, was also sampled. Pristine river sites PR0a, PR0b, and PR1 were characterized in depth in Chapter 2 and were subject to further analysis by quantitative polymerase chain reaction (Q-PCR) in this work.

The sites sampled along the SPR were chosen from a nutrient study conducted by the National Water Quality Assessment (NAWQA) Program, starting from immediately downstream of the Denver Metro Water Reclamation Facility discharge and continuing to just past the PR confluence with the SPR in Kersey, CO. To capture seasonal and matrix variability, river water and sediment (sed) were collected three times during low flow (October 2006, February 2007, and October 2007) and once during high flow conditions in May 2007.

Water was collected from the center of the river in sterile 1-L Nalgene bottles and kept on ice until further analysis. A metal spade was used to collect ~0.5 kg of sediment from the top 5 cm of the river floor. The sample was mixed and a portion was collected in a sterile 50-mL centrifuge tube that was kept on ice until further analysis.

3.2.3 Sample preparation and DNA extraction

Suspended sediment from water samples was collected via filtration with 0.45 µm filter membranes (Millipore). Filters were placed in sterile 10 mL Petri dishes and frozen prior to DNA extraction. Filters were subsequently thawed, cut into small pieces using aseptic technique, and added directly to extraction tubes. DNA extraction was carried out using the UltraClean Soil DNA Kit (MoBio Laboratories, Inc.) according to manufacturer's instructions. 1 g portions of the sediment sample were extracted in duplicate using the same kit and methodology. The quality of the DNA was verified by agarose gel electrophoresis. DNA from duplicate extractions of river sediment was pooled. All DNA extracts of river samples were aliquoted and stored at -80°C. Aliquots were diluted 1:10 with sterile deionized water for qualitative and quantitative PCR assays.

3.2.4 Distribution of ARG in river samples determined by Polymerase Chain Reaction (PCR) assays

PCR assays were used to determine the presence of eleven tet genes, tet(B), tet(C), tet(E), tet(H), tet(M), tetB(P), tet(Q), tet(S), tet(T), tet(W), and two sul genes, sul(I) and sul(II), in river samples. PCR assays of river samples were carried out in tandem with the PCR assays of source environments (AFOs and WWTPs) described in Chapter 2., Sections 2.4 and 3.1. Primers, annealing temperatures, and reaction conditions were described previously (Aminov et al. 2001,

Aminov et al. 2002, Pei et al. 2006, Koike et al. 2007) and are also described in detail in the Experimental Section (2.2) of Chapter 2.

3.2.5 ARG quantification with Quantitative Real-Time Polymerase Chain Reaction (Q-PCR)

The quantities of the 16S rRNA gene and the ARGs sul(I) and tet(W) present in DNA extracts were determined using Q-PCR as described in Suzuki et al. (2000) and Pei et al. (2006), respectively. Briefly, the reaction mixture (10 μL) for quantifying the 16S rRNA gene consisted of 1x Power SybrGreen master mix (Applied Biosystems), 5 μM each of 1396F and 1492R primers and 5 μM of the 16S probe. The reaction mixture (10 μL) for quantifying the ARGs sul(I) and tet(W) consisted of 1x Power SybrGreen master mix (Applied Biosystems) and 0.2 μM final concentration of forward and reverse primers. The 7300 Real-Time PCR System (Applied Biosystems, Foster City, CA) was used for the Q-PCR assays, with a reduced denaturing step of 10 min and annealing temperature of 65°C for sul(I) and 60°C for tet(W) (Pei et al. 2006). Serially diluted standards spanning six orders of magnitude were used as calibration curves for the analysis. Standards for sul(I), tet(W), and the 16S rRNA gene were generated from environmental samples that were cloned and sequenced for identity verification. PCR amplicons of the cloned sequences were used as the template in control reactions. Sterile deionized water was used as a blank to ensure quantification was not affected by contamination.

3.2.6 Genetic profiling of tet(W) with shot-gun cloning and Restriction Fragment Length Polymorphism (RFLP) analysis

An 1152-bp region of the tet(W) gene was amplified via PCR as described by Koike et al. (2007) with primer (2pM) and Mg^{2+} (4mM) concentration optimized for the DNA matrices of all samples. PCR products were then cloned into the pCR1™ vector using the Topo-TA™ cloning kit (Invitrogen). A BstUI restriction digest designed previously (Chapter 2.2) was used for screening

the diversity of clones prior to sequencing. BstUI digests were carried out with 20 µL of the typical reaction mixture (2 µL m13 PCR product, 18 µL New England Biolabs Buffer 2, 1 unit BstUI restriction enzyme) and were carried out at 60°C for 1 hr and visualized on a 2% agarose gel. Selected clones were sequenced in both directions using the T3 and T7 primers at Colorado State University's Proteomics and Metabolomics facility. Large amplicons were assembled using ChromasPro 1.41 software (Technelysium Pty Ltd.). Phylogenetic and molecular evolutionary analyses, including sequence alignments and tree construction, were carried out using *MEGA* version 4 (Tamura et al. 2007).

A neighbor-joining tree was generated using MEGA 4 (Tamura et al. 2007) and rooted to the sequence of the fusA gene from *Aquifex aeolicus.* The fusA gene codes for translation elongation factor EF-G and is believed to be the ancestor of tet(W) and other ribosomal protection tet genes (Aminov and Mackie 2007).

UniFrac is designed to compare communities based on the combined shared branch length of the phylogenies and is useful for understanding trees generated with large numbers of sequences (Lozupone and Knight 2005).

An existing clone library of 10 clones of a 167bp region of the tet(W) gene amplified from river site PR0b-sed was compared to the clone libraries of PR4 and SPR3. To reduce bias that could have been introduced by the use of a different reverse primer, long sequences with mismatches within the priming region (148-167bp) were eliminated from analysis. To ensure utmost sequence reliability when comparing a small conserved region of the gene, regions at the beginning of the alignment where sequences can exhibit some variability from the sequencing process were clipped, leaving 458 clones over a 113-bp region of the tet(W) for analysis. The sequences were aligned and trees created as described above. The PR0b-sed

clone library was completely identical (12/12 clones) compared to the PR2-ss library with 3 unique sequences (of 21 clones total), and the PR3b-ss library with 3 unique sequences (of 9 clones total). The PR0b consensus sequence was blasted against the clone library in this study and against The National Center for Biotechnology Information (NCBI) database. No sequences sharing 100% identity were found. A neighbor-joining tree of all the clones from the PR sequences and the PR4-ss and SPR3-ss clone libraries was drawn, mid-point rooted and tested using the Interior Branch test of phylogeny using MEGA 4 (Saitou and Nei 1987)

3.2.7 South Platte River Basin site analysis

The Envirofacts Warehouse tool available from the EPA's website (<http://www.epa.gov/enviro/>) was queried to obtain a list of National Pollutant Discharge Elimination System (NPDES) permits issued in the state of Colorado. At the time the database was queried (April 2007), 2078 NPDES permits had been issued in Colorado. 1119 of the permits were missing location coordinates (eg. latitude and longitude). Mailing addresses were given on most permits, although the variability of the structure of these permits did not allow an automated technique such as an address locator program to be used. Instead, the coordinates were determined manually using Google Earth 4 (Google Inc., Mountain View, CA).

Due to the manual labor required to geocode 1000+ permits, the database was restricted to only those within the South Platte River Basin and with Standard Industrial Classification (SIC) codes designated as either Agricultural (0211-0921), Animal product processing (2011-2023), Water and Sanitation (4941-4959), Not Elsewhere Classified (9999), or had no SIC code assigned (total of 363 entries). Between two hundred and three hundred of these entries were geocoded using Google Earth. The entries that were not geocoded either were duplicate entries, such as a reissued permit, or did not have sufficient locational information for identification. The

collected information was later used to generate shapefiles of WWTPs, Fish Hatcheries, Feedlots, and Dairies within the study area watershed. At this time in the state of Colorado there was no law in effect requiring animal feeding operations (AFOs) to obtain NPDES permits unless they were directly discharging to surface water. Thus, the NPDES permits only captured a handful of (AFOs) within the watershed and were inadequate to fully understand the influence of AFOs on the study area.

Beef feedlots and dairies can be easily recognized via satellite imagery; the two combined represent >90% of all animal feeding operations along Colorado's front range. Beef feedlots and dairies that were located within the watershed were manually digitized as polygons using the Environmental Risk Assessment Management System (eRAMS) tool (Department of Civil Engineering, Colorado State University) available online: <http://ids-yampa.engr.colostate.edu/erams/index.psp >. Polygon area was calculated using ArcGIS version 9.2 (ESRI, Inc., Redlands, CA). One limitation in using manual digitization is potentially missing smaller farms. This could result in variability in the detection of small farms across the watershed. Because of this known limitation, digitized polygons with areas less than 15 acres were excluded from analysis. It is inherent that larger CAFOs and AFOs will generally have more animals and thus represent larger potential contributors of manure that can impact the natural environment than small farms. Therefore, exclusion of small facilities or farms is unlikely to have a significant effect on the dataset.

In December 1, 2006, the Colorado Department of Public Health and the Environment (CDPHE) began accepting CAFO discharge permits. CAFOs are only required to obtain a permit if they are discharging to surface water. However, if a CAFO indirectly discharges to surface waters, for example during a rainfall event, that CAFO is liable for penalties associated with an unpermitted discharge according to the Clean Water Act (CDPHE 2006, CDPHE 2009). A list of

CAFO permits including the type of CAFO and the number of animals at each facility was compiled and graciously provided by Stewart Environmental, Inc. (Fort Collins, CO). CAFO permits were spatially joined in ArcGIS version 9.2 (ESRI, Inc., Redlands, CA) with the AFO polygon layer created with eRAMS. Permitted CAFOs constituted ~50% of the digitized polygons. The digitized areas were regressed with counts to generate expected counts for the remaining polygons. To provide the closest fit to the data, regressions were done separately for small dairies, large dairies, small feedlots, and large feedlots (Figure A2.1, Appendix A2). When the summation of predicted counts (for those missing this information) and actual counts (for the portion that had CAFO permits) were plotted against area, there was a strong correlation (Figure A2.2, Appendix A2).

3.2.8 Construction of spatial variables

Based on the assumption that sources closer to a river site would be more influential than sources farther away, a method of inverse distance weighting was used in calculating the total capacity of WWTPs or total animal counts of AFOs for each river sampling point (Shepard 1968). Distances from sources to river sites were calculated using a two-step method. First, sources were plotted as points within the watershed and the Proximity tool in ArcGIS was used to determine the closest access point to the river and an estimate of overland flow. These access points were plotted as XY coordinates along the river line shapefile. Then the River layer was split into multiple small lines using Split at Vertices tool in ArcToolbox and distances from the source to the nearest outlet (river site) were calculated using a method of additive selection and statistical summation. Distances between each connected river site were also determined and used to calculate the distance from a source to each downstream river site. Plotting and calculations were performed using ArcGIS version 9.2 (ESRI, Inc., Redlands, CA). Equation 3.1 was used to determine the inverse distance weighted (IDW) count where w is the inverse-

distance weight, C is the count or capacity, and d is the sum of the overland and channel distance from a source to a river location (Shepard 1968). The count or capacity at a particular farm or treatment plant was multiplied by the IDW. Then these weighted counts were summed for each river site. Only those sources within the watershed area of a river site were included in the summation. In the text, the IDW capacity at a river site will be referred to as an IDW count.

$$\sum_{i=1}^{n} w_i \; C \quad \text{where} \quad w_i = \frac{d^{-1}}{\sum_{j=1}^{n} d_j^{-1}} \qquad \text{(Equation 3.1)}$$

We used the Soil and Water Assessment Tool (SWAT) interface to delineate the watershed. The underlying procedure in SWAT is borrowed from ArcHydro extension in ArcGIS 9.2 (ESRI, Inc., Redlands, CA) and requires the processing of a 30-m Digital Elevation Model (DEM) from National Elevation Dataset (NED) that is administered by the U.S. Geological Survey (USGS).

3.2.9 Classification of river sites by ARG profiles

All statistical procedures described below were performed in Matlab R2009a (The MathWorks Inc., Natick, MA). In Chapter 2, tet(C), tet(E), and tet(O) were identified with WWTP environments and tet(H), tet(Q), tet(S), tet(T) with AFO environments. These gene combinations were shown to be effective for distinguishing between classifying between pristine, WWTP, and AFO environments (Chapter 2). Here, the frequency of detection (FOD) of these ARG groups in source environments were used as the training dataset for classification of river sites using the methods of classification and regression tree analysis (CART) and linear classification by discriminant analysis.

Similarly, the clone libraries of AFO lagoons and WWTPs (Chapter 2) were used as the training dataset for classification of PR4 and SPR3 by RFLP pattern distribution. Sequences generated from the clone libraries of PR4 and SPR3 were aligned and a neighbor-joining tree

was created using MEGA 4 (Tamura et al. 2007). The clone libraries of PR4 and SP3 were aligned both with and without sequences from the clone libraries developed in Chapter 2, Section 2.4. The trees were rooted to the elongation factor gene, fusA, of *Aquifex aeolicus* and used to perform Principal Component Analysis (PCA), UniFrac significance tests, and to calculate an environmental distance matrix; these phylogenetic analyses were carried out in UniFrac (available online: <http://bmf2.colorado.edu/unifrac/index.psp>) (Lozupone and Knight 2005, Lozupone et al. 2007).

3.2.10 Classification of river sites by spatial variables

Multivariate linear regression of sul(I)/16S rRNA and tet(W)/16S rRNA was performed using total counts and IDW counts as explanatory variables. Models were selected using the process of backward selection. The R^2 coefficient and the F-statistic were used to assess the model fit. Confidence intervals of parameter coefficients were used to determine the significance of explanatory variables with intervals including zero indicating non-significance.

CART analysis of spatial variables was also performed using IDW counts and the interaction of inverse distance with IDW counts of AFOs, WWTPs, and FRUs as explanatory variables and ARG groups 1 and 2 and sul(I)/16S rRNA as response variables. To determine the most critical explanatory variables of those tested, bootstrapping and aggregation of trees was performed using the TreeBagger method in Matlab R2009a. 1000 trees were generated and compared on the basis of mean squared error. This method involves changing each explanatory variable and determining how it affects the mean squared error (classification error). The variables resulting in the highest misclassification when permutated represent the most important features in classifying the tree. 1000 trees were then grown from the reduced set of explanatory variables, and the mean squared error was compared with the full set. If the error of the reduced model

was similar to that of the full model, then those variables were used for construction of a simplified tree.

3.3 Results and Discussion

3.3.1 Presence and distribution of ARG determined by PCR

A total of sixty-four river samples from sites PR2, PR3a, PR3b, PR4, PR5, SPR1, SPR2, SPR3, were screened for the presence of thirteen ARG by PCR. The tet(B) and tetB(P) genes were not detected in any river sample. This is not surprising given that in Chapter 2, these genes were only detected in AFO lagoons and were detected just twice out of 64 possible detections. These genes were not used in downstream calculations or analyses. The remaining eleven ARG were detected in a minimum of three samples. The total FOD in SPR samples was slightly higher (0.33) compared to PR samples (0.26). On average, ARG were detected with greater frequency in suspended sediments than sediments in the SPR (0.44 vs. 0.22) and the PR (0.36 vs. 0.16). The most frequently detected ARGs were sul(I) and sul(II), being found in 73% and 63% of all samples, respectively. Of the tet genes, tet(C) was the most frequently detected (45%).

When the distribution of ARG in PR and SPR samples was compared with previously characterized pristine PR samples (Table 3.2), the FOD of the tet(C), sul(I), and sul(II) genes were found to be markedly higher in the impacted river (non-pristine river sites). The tet(C) gene was detected in nearly half of impacted river samples analyzed compared with less than 10% detection in the pristine river samples. Similarly, the FOD of sul(II) was nearly 4-fold higher in the impacted vs. pristine river environment. The sul(I) gene was found in 73% of impacted river samples whereas sul(I) was detected in just 1 of 24 pristine PR samples. In contrast, the FOD of both tet(W) and tet(M) were similar between pristine and impacted sites. This supports the

results of previous work that demonstrated similar distributions of tet(W) and tet(M) in pristine and source environments (Chapter 2).

3.3.2 Classification of river sites with ARG indicator variables

The tet genes tet(C), tet(E), and tet(O) will be referred to in the remainder of the text as tet group CEO and the tet genes tet(H), tet(Q), tet(S), and tet(T) as tet group HQST. The major findings from Chapter 2 were as follows: presence of sul(I) indicated an impacted environment, tet group CEO were the most dominant genes in WWTPs, and tet group HQST were found in greater frequency in AFO lagoons than in WWTPs. The overall frequency of detecting ARGs in an environment can distort classification when the between-group variance is high. Therefore, the FOD of source and river environments were first scaled to the overall frequency of detection for each environment (source or river) prior to classification. The FOD of tet group CEO and tet group HQST in pristine river and source environments were plotted and used for training in discriminant analysis (Figure 3.2a). The resulting plot was used to classify river samples as most similar to pristine PR, WWTPs, or AFOs. SPR3 was classified with the AFO group and all other PR and SPR sites with the WWTPs. SPR3 is located just downstream of the confluence of the PR with the SPR and is highly impaired from anthropogenic influence (Figure 3.1). It has many WWTPs and AFOs within its watershed, but immediately upstream of SPR3 is a heavy agricultural area (Figure 3.1a and 3.1c). SPR1 is heavily influenced by urban WW (Figure 3.1) and has the highest normalized FOD of tet group CEO, the indicator variable for WWTPs. When sul(I) is also included in the analysis (Figure 3.2b), there appears to be a trend of increasing FOD of sul(I) at more downstream sites. Also, PR2 shifts closer to pristine river samples; this could be because it is nearest to the pristine region of the river and has few known sources within its watershed. From these findings, it is probable that spatial variables could be useful in explaining variation in the ARG profiles of river samples.

3.3.3 Classification and Regression Tree (CART) analysis of ARG profiles

CART analysis is particularly useful when analyzing data that does not meet the assumptions of linear regression, such as homogeneity of variance, normal distribution, and independence of explanatory variables. Furthermore, when bootstrapping and aggregation (bagging) analysis of CART is performed, variables that are most important for inclusion in the CART model can be easily identified without the multiple steps of forward or backward model selection used with linear regression (Lewis 2000). Feature importance is related to the increase in mean squared error (error resulting from misclassification) when the variable is permutated. Variables that result in greater error and misclassification when altered are more necessary for classification and regression of the data (Lewis 2000). Variables with higher feature importance were used in a final bagging analysis to verify that model reduction did not increase the misclassification error of the tree ensembles.

Bagging analysis of CARTs (Figures A2.3 – A2.5, Appendix 2) was performed using the nine spatial indicators as explanatory variables (Figure 3.3a) and ARG variables as the response variables (Figure 3.4a); bagging was used to select a reduced model (Figure 3.3b-d) for generating a final tree. A total of nine spatial variables were used in CART analysis (Table 3.3, Figure 3.3a). CART analysis was performed separately for each ARG indicator variable: sul(I), tet group CEO, and tet group HQST. For each ARG variable, 1000 trees were grown and bagged to determine feature importance of the explanatory variables.

For sul(I), five of the nine variables had similar feature importance including the IDW discharge (WWTP), the IDW cattle count (AFO), and the interaction of IDW trout count (Figure 3.3b). In our previous work, we detected sul(I) in just 1 of 24 pristine samples but in 100% (62 total) source samples (Chapter 2). Thus, it was hypothesized that sul(I) would be a good

indicator of impact from source environments, but its ubiquity in those sources made it ineffective for classification between sources. The reduced model selected by bagging of CARTs (Figure 3.3b) included at least one spatial variable related to each source and the feature importance was similar between variables, again demonstrating impact, but not a particular source environment.

IDW WWTP discharge and the ratio of WWTP discharge to number of cattle were the most important variables of the nine tested by CART bagging analysis of tet group CEO (Figure 3.3c). This is in line with the working hypothesis that tet group CEO could be an indicator of WWTP influence on a watershed. When the same analysis was performed with tet group HQST, three variables were chosen for the reduced model: IDW WWTP discharge, an interaction between AFOs and WWTPs, and the ratio of WWTP discharge to AFO cattle count (Figure 3.3d). It seems that the variables relating to WWTPs may be as useful in classification of tet group HQST as those relating to AFOs.

The reduced models were then used to create a single CART, or tree, for each ARG response variable. Trees were pruned to reduce noise and misclassification (Figure 3.4). Figure 3.4a is included as a reference of the actual FODs of the three ARG response variables at each river site. IDW WWTP discharge and IDW cattle count were both used for classification of sul(I) in river sites (Figure 3.4b). There is a trend of increasing impact (by both WWTPs and AFOs) from left to right of the grouped river sites (members). The FOD of sul(I) also increases from left to right and closely matches the range of actual FOD for each member group. This strengthens our previous claim that the presence of sul(I) indicates the presence of a source in the watershed (Chapter 2), and further shows that increasing impact of various sources within a watershed correspond with an increasing detection of sul(I). The IDW WWTP discharge is used as the only variable in the classification of river sites by tet group CEO (Figure 3.4c). The increasing impact from WWTPs,

from left to right, corresponds with an increase in FOD of tet group CEO. Again, there is an agreement with our previous assumption that tet group CEO is an indicator of WWTP influence (Chapter 2). The FOD of tet group HQST is explained by measures of both AFO and WWTP influence (Figure 3.4d); an interaction term of WWTP and AFO influence and the IDW WWTP discharge are the two spatial variables used for classification of tet group HQST in river sites. Some measure of AFO influence was used in the regression; however, AFO influence alone is not sufficient for explaining the frequency of detection of tet group HQST.

One of the limitations of using the tet group CEO and tet group HQST as variables is that, with the exception of tet(S), all of the genes were detected in both source environments, but at varying frequencies. This combined with the small sample size of the data set could limit the ability to make distinctions between river sites that are simultaneously influenced by both source environments. Given the presence of these gene groups in both sources, it could be that WWTPs are the main source of tetHQST, but in sites where agricultural influence is high, AFOs provide an additional source of these genes.

3.3.4 Multivariate linear regression of quantities of sul(I) and tet(W)

While CART is a useful method of non-linear regression, until recently (Stanimarova et al. 2008) its use has been limited in the scientific literature (Lewis 2000, Stanimarova et al. 2008, Titonell et al. 2008). Therefore, a simple example of multivariate linear regression is also offered. Multivariate linear regression using the ordinary least squares (OLS) method was performed number of copies of sul(I) and tet(W) present in the samples (quantified by Q-PCR and presented in Figure A2.6) as response variables and the following spatial variables as the explanatory variables: the IDW counts for WWTPs, AFOs, and FRUs, and an interaction term

(IDW WWTP*IDW AFO). Confidence intervals of parameters were used as the criteria for backwards model selection.

The actual and predicted quantities of sul(I) for the full model (Figure 3.5a) are strongly correlated (R^2 = 0.98, p <0.0003). Of the parameters for the full model, only the IDW cattle count was significant. In the reduced model, IDW cattle count was used as the sole predictor; there was also a good fit between the actual and predicted data (R^2 = 0.83, p <0.0003). The same trend emerges: higher concentrations of sul(I) at more impacted river sites.

Both in previous work and in this study, tet(W) was detected with similar frequencies in pristine, source, and impacted river environments (Chapter 2). It would follow then, that quantities of tet(W) would not be correlated with spatial indicators of sources. The predicted and actual tet(W) quantities were not significantly correlated in the full model (R^2 = 0.11, p = 0.95) and no parameters were significant (Figure 3.5c). For visual comparison, a reduced model using AFO IDW cattle count as the only explanatory variable (R^2 = 0, p = 0.85) is also shown (Figure 3.5d).

It is fully acknowledged that some assumptions may be violated in these simplistic regression models. Nonetheless, the results of linear regression are in agreement with conclusions made in both previous (Chapter 2) and present work; sul(I) is detected with greater frequency and increases in magnitude at more impacted sites and tet(W) is not correlated with spatial source indicators.

3.3.5 Phylogenetic analysis of tet(W) profiles of river sites PR4 and SPR3

A 1152-bp region of the tet(W) gene was amplified in suspended sediment from PR4 and SPR3 for phylogenetic analysis. Clone libraries of PR4 and SPR3 were generated and digested with BstUI to determine the distribution of restriction fragment length polymorphism (RFLP)

patterns (Figure 3.6). Based on the BstUI RFLP analysis of tet(W) clone libraries of WWTPs and AFOs, four patterns were dominant and easily identified from sample to sample. These four patterns were each identified in the clone libraries of PR4 and SPR3. Only one other unique pattern was identified in PR4. No unique patterns were identified in SPR3. In particular, pattern c1 and c2 were useful in classifying clones from AFOs from those in WWTPs (Chapter 2). The distribution of these patterns was used for classification based on discriminant analysis of PR4 and SPR3. SPR3 was classified with AFO lagoons and PR4 with WWTPs (Figure 3.6).

100% of patterns c1, c2, and c4 and 50% of clones of pattern c3 were sequenced, assembled and aligned with the existing clone libraries of 5 AFO lagoons and 3 WWTPsThe tree generated from the alignment of the sequences was analyzed using the online software, UniFrac (Lozupone and Knight 2005), to compare the libraries from the river environments to the existing libraries of source environments.

Pairwise comparisons were made between the following environments: PR4, SPR3, WWTPs (sequences from three libraries) and AFO lagoons (sequences from six libraries) for a total of 5 pair-wise comparisons (WWTPs and AFOs have been compared previously (Chapter 2)). Sequences that were absent from either environment being compared were excluded, and the resulting sequences were tested to determine significance at the p=0.05 level (Lozupone and Knight 2005). The clone library of PR4 was significantly different from AFO lagoons ($p < 0.05$),but not from WWTPs (p=0.15); SPR3 was not significantly different from either AFOs (p=1.0) or WWTPs (p=0.3).

The relatedness of each environment to all other environments can be expressed as an environmental distance (based on sequence similarity between environments). Small environmental distances represent greater similarity (shared branch length) between

environments (Lozupone and Knight 2005). River sites could be impacted by both urban and agricultural sources. To assign a magnitude to the relatedness of river sites to each source environment, a ratio of the distance from the two source environments to a river site can be used (Figure A2.8, Appendix A2). The distance between PR4 and the AFO environment was 1.8 times greater than the distance between PR4 and the WWTP environment. In contrast, the distances between sources and SPR3 were nearly equal: a ratio of 1.03:1 (AFO:WWTP). Thus, the tet(W) profile of PR4 shares much greater similarity with that of WWTPs than of AFOs, but the profile of SPR3 is equally similar to AFOs and WWTPs.

The limited scope of the phylogenetic study precludes a regression or classification analysis of river sites with spatial indicators; however, these spatial variables can guide quantitative interpretation and discussion of the differences in tet(W) profiles of PR4 and SPR3. Site PR4 is located in a rural area in Northern Colorado, between the small urban centers of Fort Collins (population 130,000) and Greeley (60,000). Eight WWTPs, including four major dischargers, have permits for discharge to the PR or its tributaries upstream of PR4. An average of 28 million gallons per day (MGD) is discharged upstream of PR4. In fact, one of the WWTPs discharges directly to the river only 4km upstream of the P4 sampling point. The watershed of PR4 also includes several AFOs (Figure 1a).

SPR3 is located just downstream of the confluence of the PR. The upstream region of the SPR is dominated by WWTPs (Figure 3.1a and 3.1b) but as the SPR flows north away from the Denver metro area, it enters a heavy agricultural area (Figure 3.1a and 3.1c). Two large beef CAFOs (total permitted capacity of 54,000 head) are located adjacent to the river, less than 4km upstream of the SPR3 sampling point.

The phylogenetic analysis of tet(W) suggests that WWTPs are major players in influencing the communities of ARG in river environments. In areas influenced by large concentrated AFOs, such as SPR3, these AFOs also play a prominent role.

Though presence of tet(W) in pristine environments has been documented, low copy number and matrix effects prevented amplification of the 1152-bp region of the gene in upstream river samples. To determine whether the selection of native tet(W) genes from native river sediments due to selection pressures could explain the results found, amplicons of a smaller overlapping region of the tet(W) gene were also analyzed using UniFrac. An existing clone library of 10 clones of a 167bp region of the tet(W) gene amplified from river site PR0b-sed was compared to the clone libraries of PR4 and SPR3. The 0b-sed clones and 2 clones from the PR2-ss library form one cluster separate from the upstream river cluster. The clustering of the upstream PR river (0b-sed clones and 2 clones from the PR2-ss) separately from the further downstream PR-ss suggests that the tet(W) community in native river sediments is different from the communities seen at more impacted sites. Clones from the PR2-ss clone library are spread between both clusters. As PR2 only receives minor influences, it is possible that the community is balanced between tet(W) originating from native river sediments and those genes that are transported there by point and nonpoint sources.

3.4 Summary

This is the first study to compare the ARG profiles of pristine and impacted river environments with multiple source environments and to utilize geospatial analysis as a tool to explain variation in the ARG profiles of riverine samples. Urban (WWTPs), agricultural (AFOs), and river environments within the same watershed were characterized simultaneously to obtain several lines of evidence in order to elucidate whether selection of native ARG or transport from

WWTPs or AFOs is the predominant mechanism of ARG persistence and propagation in river environments. The general consensus of the various analyses of the ARG profiles, WWTPs appear to be the dominant source influencing the presence, quantities, distribution, and diversity of ARGs. However, in areas with intense agricultural land use, such as SPR3, AFOs may also be influential. Selection of native ARG due to antibiotic pollution of rivers from both urban and agricultural sources has been a persistent theory of the mechanism of ARG propagation in river environments. However this work indicates that transport from source environments, particularly WWTPs, is the more likely scenario. In this work, we demonstrated correlation between spatial indicators of source environments and their corresponding ARG variables. Furthermore, the tet(W) profiles of river sites shared greater similarity with source environments than with more pristine river sites. From these findings, we suggest that selection of ARG at source environments, followed by transport to river environments is the most probable mechanism driving the elevated levels and distribution of ARGs in the South Platte River Basin.

REFERENCES

Aminov, RI, Chee-Sanford, JC, Garrigues, N, Teferedegne, B, Krapac, IJ, White, BA, Mackie, RI (2002) Development, validation, and application of PCR primers for detection of tetracycline efflux genes of gram-negative bacteria. *Applied and Environmental Microbiology,* 68: 1786–93.

Aminov, RI, Garrigues-Jeanjean, N, and Mackie, RI (2001) Molecular ecology of tetracycline resistance: Development and validation of primers for detection of tetracycline resistance genes encoding ribosomal protection proteins. *Applied and Environmental Microbiology*, 67: 22-32.

Baquero F, Martinez JL, Canton R (2008) Antibiotics and antibiotic resistance in water environments. *Current Opinion in Biotechnology*, 19: 260-265.

CDPHE (2006) Colorado Department of Public Health. CAFO Fact Sheet: Status of Federal and Colorado CAFO Permit Regulations and Requirements. Available online: <http://www.coloradolivestock.org>.

CDPHE (2009) Colorado Department of Public Health. Regulation 81: What does it mean for Non-Permitted Large CAFOs? Available online: <http://www.cdphe.state.co.us/oeis/eap/cafo.html>.

Hu, JY, Shi, JC, Chang, H, Li, D, Yang, M, Kamagata, YC (2008) Phenotyping and genotyping of antihiotic-resistant Escherichia coli isolated from a natural river basin. *Environmental Science & Technology,* 42 (9): 3415-3420.

Lewis, RJ (2000) An Introduction to Classification and Regression Tree (CART) *In* Annual Meeting of the *Society for Academic Emergency Medicine* proceedings. San Francisco, California.

Kim, SC, Carlson, K (2006) Occurrence of ionophore antibiotics in water and sediments of a mixed-landscape watershed. *Water Research,* 40: 2549-2560.

Kim, SC, Carlson, K (2007a) Quantification of human and veterinary antibiotics in water and sediment using SPE/LC/MS/MS. *Analytical and Bioanalytical Chemistry,* 387: 1301-1315.

Kim, SC, Carlson, K (2007b) Temporal and spatial trends in the occurrence of human and veterinary antibiotics in aqueous and river sediment matrices. *Environmental Science & Technology*, 41: 50-57.

Koike, S, Krapac, IG, Oliver, HD, Yannarell, AC, Chee-Sanford, JC, Aminov, RI, Mackie, RI (2007) Monitoring and source tracking of tetracycline resistance genes in lagoons and groundwater adjacent to swine production facilities over a 3-year period. *Applied and Environmental Microbiology*, 73 (15):4813-4823.

Kolpin, DW, Furlong, ET, Meyer, MT, Thurman, EM, Zaugg, SD, Barber, LB, and Buxton, HT (2002) Pharmaceuticals, hormones, and other organic wastewater contaminants in US streams, 1999-2000: A national reconnaissance. *Environmental Science & Technology,* 36: 1202-1211.

Kümmerer, K, and Henninger, A (2003) Promoting resistance by the emission of antibiotics from hospitals and households into effluent. *Clinical Microbiology and Infection,* 9: 1203-1214.

Lozupone C and Knight R. (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, 71: 8228-8235.

Lozupone C, Hamady M, and Knight R. (2006) UniFrac - An Online Tool for Comparing Microbial Community Diversity in a Phylogenetic Context. *BMC Bioinformatics*, 7: 371.

Lozupone, C, Hamady, M, Kelley ST and Knight, R (2007) Quantitative and qualitative (beta) diversity measures lead to different insights into factors that structure microbial communities. *Applied and Environmental Microbiology*, 73: 1576-1585.

Pei, R, Kim, S-C, Carlson, KH, and Pruden, A (2006) Effect of River Landscape on the sediment concentrations of antibiotics and corresponding antibiotic resistance genes (ARG). *Water Research,* 40: 2427-2435.

Pruden, A, Pei, RT, Storteboom, H, Carlson, KH (2006) Antibiotic resistance genes as emerging contaminants: Studies in northern Colorado. *Environmental Science & Technology,* 40: 7445-7450.

Saitou N, and Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4: 406-425.

Stanimirova, I, Kubik, A, Walczak, B, Einax JW (2008) Discrimination of biofilm samples using pattern recognition techniques. *Analytical and Bioanalytical Chemistry*, 390: 1273–1282.

Tamura K, Dudley J, Nei M, and Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24: 1596-1599.

Tittonell, P, Shepherd, KD, Vanlauwe, B, and Giller, KE (2008) Unravelling the effects of soil and crop management on maize productivity in smallholder agricultural systems of western Kenya—An application of classification and regression tree analysis. *Agriculture, Ecosystems and Environment*, 123: 137–150.

Yang, S, and Carlson, K (2003) Evolution of antibiotic occurrence in a river through pristine, urban and agricultural landscapes. *Water Research*, 37: 4645-4656.

Yang, S, Carlson, K (2004) Routine monitoring of antibiotics in water and wastewater with a radioimmunoassay technique. *Water Research*, 38: 3155-3166.

TABLES

Table 3.1 Poudre and South Platte River sampling locations

River	Site Name	Site ID	Location
Cache La Poudre	Site 0a	PR0a	Elephant Rock, Roosevelt National Forest
	Site 0b	PR0b	Profile Rock, Roosevelt National Forest
	Site 1	PR1	Greyrock Trailhead, Roosevelt National Forest
	Site 2	PR2	Shields St. Bridge, Fort Collins
	Site 3a	PR3a	Mulberry St. Bridge, Fort Collins
	Site 3b	PR3b	Drake Reclamation Facility, Fort Collins
	Site 4	PR4	95th Avenue Bridge, Weld County
	Site 5	PR5	Greeley Municipal Airport, Greeley
South Platte	Site 1	SPR1	Clear Creek Confluence Park, Commerce City
	Site 2	SPR2	County Road 54 Bridge, Evans
	Site 3	SPR3	Poudre River Confluence, Kersey

Table 3.2 Frequency of detection of eleven ARG in the impacted river environment compared to the pristine river environment

Environment	tet(C)	tet(E)	tet(O)	tet(H)	tet(Q)	tet(S)	tet(T)	tet(M)	tet(W)	sul(I)	sul(II)	Total
Pristine River	0.08	ND	0.08	0.04	ND	ND	0.04	0.25	0.33	0.04	0.17	0.09
Impacted River	0.45	0.06	0.23	0.08	0.22	0.05	0.02	0.31	0.39	0.73	0.63	0.29

Table 3.3 Spatial explanatory variable dataset used in regression analyses

variable type	IDW counts			IDW count*distance^{-1}			IDW count	IDW cnt*dist^{-1}	ratio
Source	WWTP	AFO	FHRU	WWTP	AFO	FHRU	WWTP*AFO	WWTP*AFO	WWTP:AFO
Feature ID	1	2	3	4	5	6	7	8	9
factor				$\times 10^{-6}$		$\times 10^{-6}$		$\times 10^{-6}$	
PR0a	0	0	0	-	-	-	0	-	1.0086
PR0b	0	0	50000	-	-	6.14	0	-	1.0086
PR1	0	0	50000	-	-	0.94	0	-	1.0086
PR2	0.1	1399	841513	0.64	0.123	72.41	70	0.08	-0.0005
PR3a	7.0	2930	829236	84.26	0.190	63.56	20384	16.00	0.0011
PR3b	6.5	2897	812296	73.33	0.134	54.56	18867	9.81	0.0011
PR4	2.6	6494	755994	92.99	0.218	32.81	16665	20.29	0.0002
PR5	10.3	7048	728438	85.20	0.165	24.58	72374	14.08	0.0002
SPR1	57.5	0	0	95.91	-	-	0	-	-379.5
SPR2	3.6	8915	0	49.76	0.159	-	32495	7.93	0.0015
SPR3	5.3	8997	721667	45.65	0.147	22.73	47719	6.73	0.0006

FIGURES:

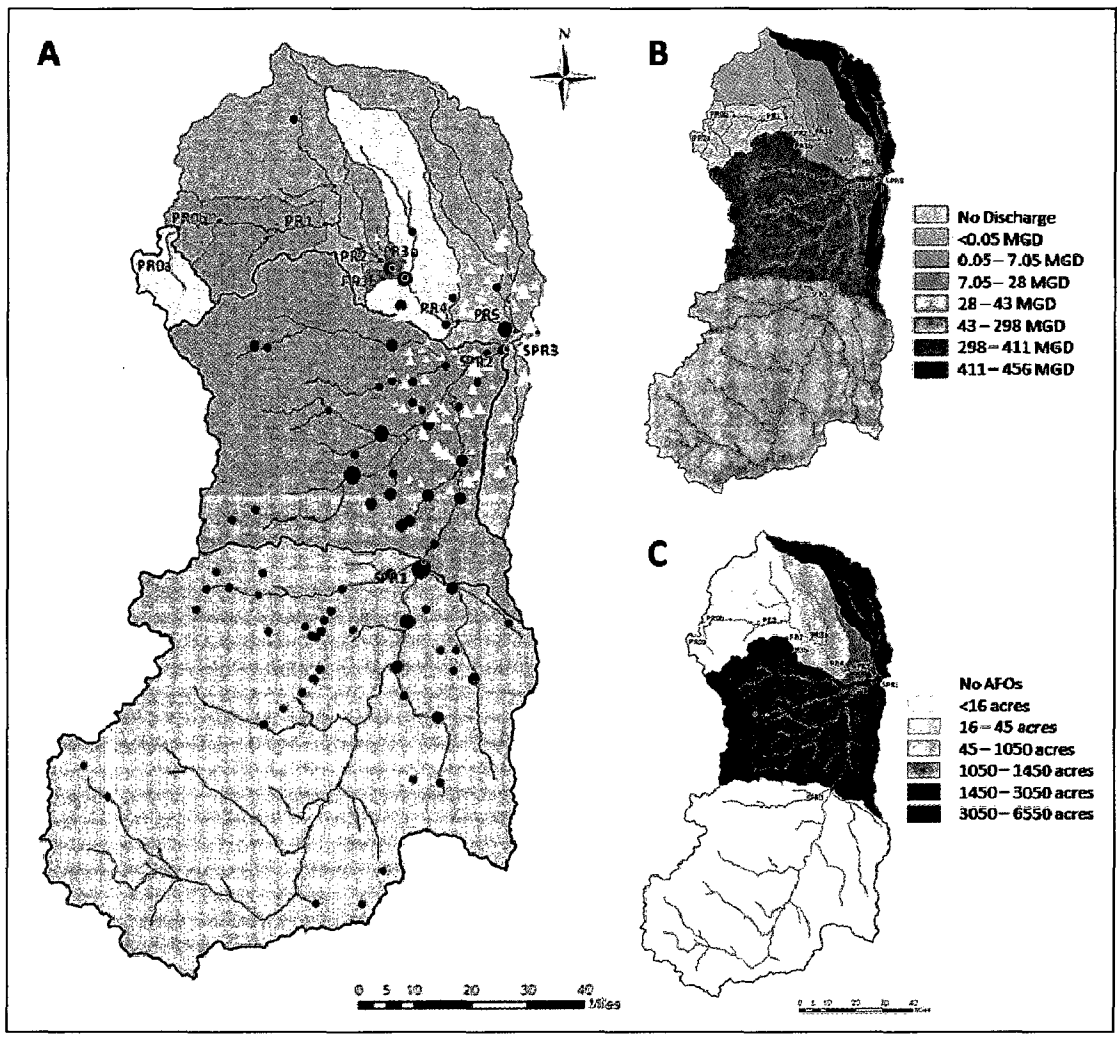

Figure 3.1: Maps of the study region. (A) Map of rivers, sampling sites, sub-watersheds and potential sources of antibiotics and ARG: sub-watersheds for each river sampling location are outlined in dark grey and an alternating grey-scale background; site ID is labeled in bold face type; WWTPs are depicted by red circles with increasing circle size corresponding to greater average design flow; AFOs are depicted as yellow triangles with scales representing larger farms (by both count and area); hatcheries are shown as green diamonds. (B) Map of cumulative average design flow of WWTPs for each watershed. (C) Map of cumulative AFO acreage for each watershed.

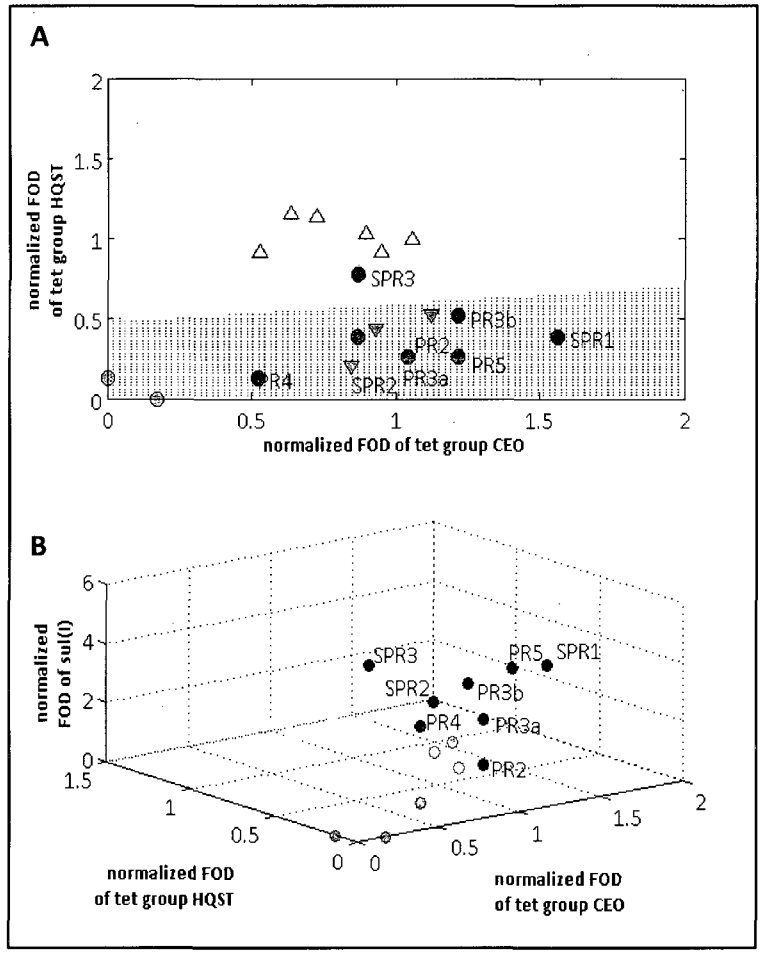

Figure 3.2: Discriminant analysis of ARG profiles of river sites by normalized frequency of detection of ARG indicator variables. The average frequency of detection of tet group CEO and tet group HQST were used as the variables for discriminant analysis of river sites (A). Frequency of detection was normalized to the overall frequency of detection of all genes to reduce the bias that total frequency of detection can have in classification. The plot is extended to 3-D with the addition of the normalized FOD of sul(I) as an indicator variable (B).

A

Feature ID	Source	Mathematical Representation	Variable description	units
1	WWTP	xw	IDW average discharge	(MGD)
2	AFO	xa	IDW cattle count	(no. cattle)
3	FHRU	xf	IDW fish count	(no. trout)
4	WWTP	$xw*dw^{-1}$	interaction	$(MGD)(m^{-1})$
5	AFO	$xa*da^{-1}$	interaction	$(no. cattle)(m^{-1})$
6	FHRU	$xf*df^{-1}$	interaction	$(no. trout)(m^{-1})$
7	WWTP&AFO	xw*xa	interaction	(MGD)(no. cattle)
8	WWTP&AFO	$(xw*dw^{-1})*(xa*da^{-1})$	interaction	(MGD)(no. cattle)
9	WWTP&AFO	xwt*xat	ratio of total counts	unitless

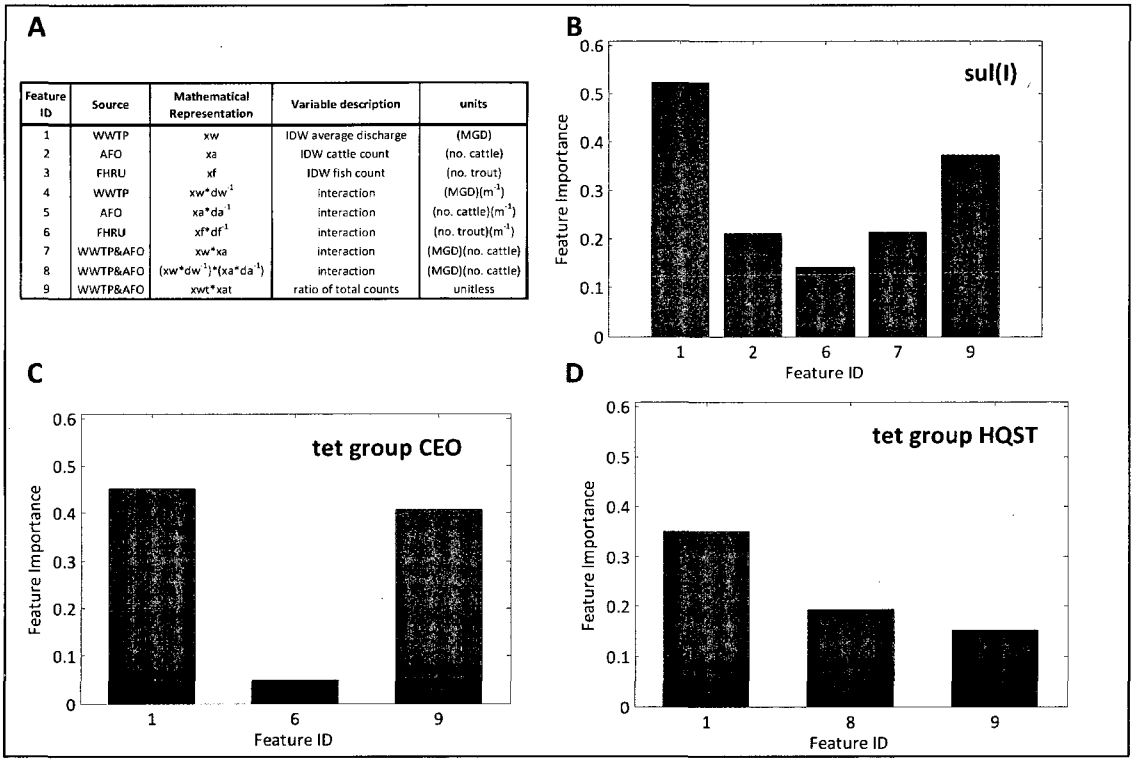

Figure 3.3: Bootstrapping analysis of Classification and Regression Tree (CART) Analysis of ARG variables with spatial variables. Regression tree analysis was performed using nine spatial indicators (A) as explanatory variables. Analyses were performed for three response variables: the frequency of detection (FOD) of sul(I), the FOD of tet group CEO, and the FOD of tet group HQST. Ensembles of 1000 bootstrapped trees were tested. All nine variables were analyzed and those with greater importance in the full model were chosen for further bootstrapping analyses. For details of the feature importance of each of the nine variables and the associated classification mean squared error (MSE) refer to figures A2.3, A2.4, and A2.5 in the Appendix.

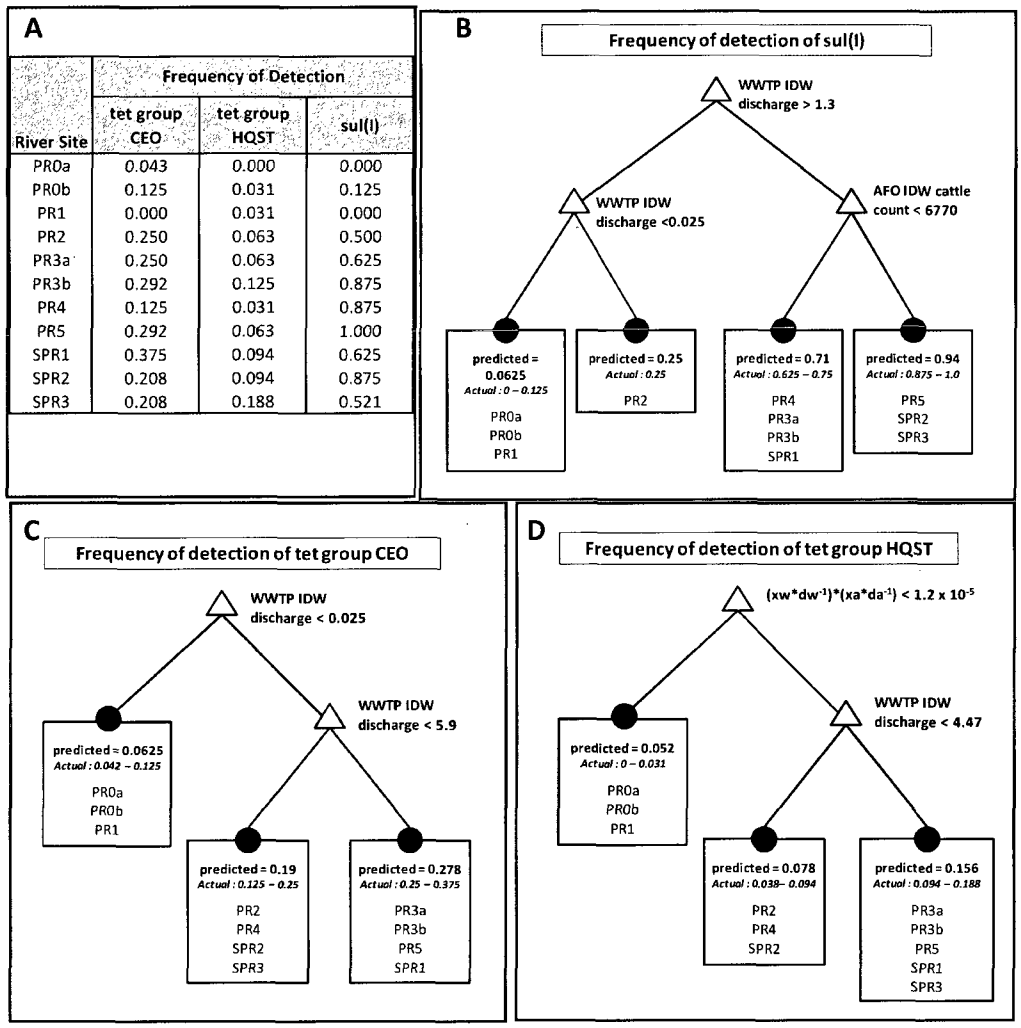

Figure 3.4: Regression trees for ARG indicator variables classified by spatial variables. (A) summary of frequency of detection of ARG indicator variables for each river site; (B) regression tree for FOD of sul(I); (C) regression tree of the FOD of tet group CEO; (D) regression tree of the FOD of tet group HQST. Trees shown were pruned from full trees to reduce noise and misclassification.

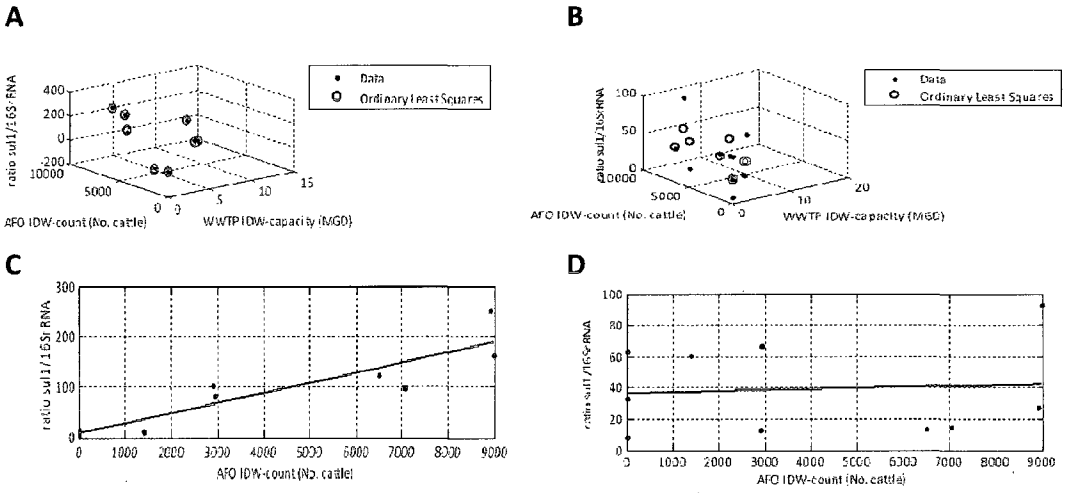

Figure 3.5: Linear regression of the average quantity of sul(I) and tet(W) in riverine samples with spatial variables. Multivariate linear regression was performed using the average normalized copies of the sul(I) (plots A and C) and tet(W) (plots B and D) as response variables and the following spatial indicators as explanatory variables: the IDW counts of WWTPs, AFOs, and FRUs and the interaction of IDW counts of WWTPs and AFOs. Shown are 3-D plots of the raw data and predicted data (ordinary least squares) for the full model for sul(I) (A) and tet(W) (B). In the reduced model, the AFO IDW count was used as the single explanatory variable for sul(I) (C) and tet(W) (D). ARG were normalized to the copies of the 16SrRNA gene. [Note: the choice of using AFO IDW count for the reduced tet(W) model was arbitrary since no parameters in the full model were significantly different from zero.]

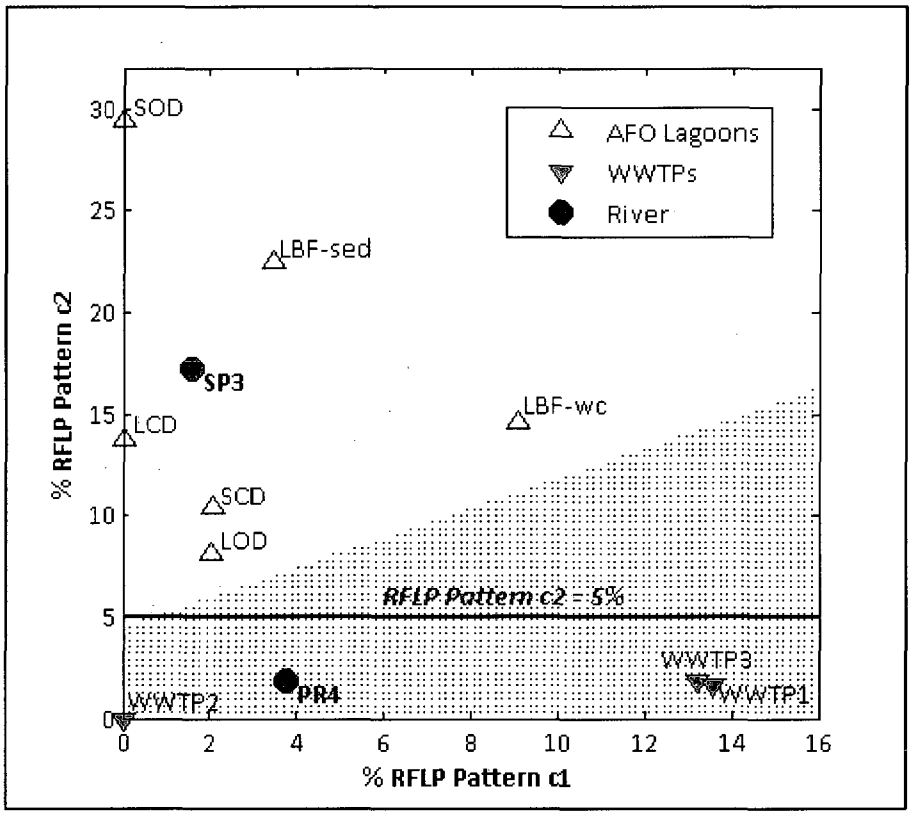

Figure 3.6: Classification of river sites PR4 and SP3 by RFLP patterns. The percentage of RFLP patterns c1 and c2 were used to classify river samples PR4 and SP3 as either WWTP or AFO influenced. Shaded red and yellow areas represent the area classified as WWTP influenced and AFO influenced, respectively. The classification tree is overlaid over the plot, defining the space above the line (samples with RLFP pattern c2 representing more than 5% of all clones) as AFO influenced.

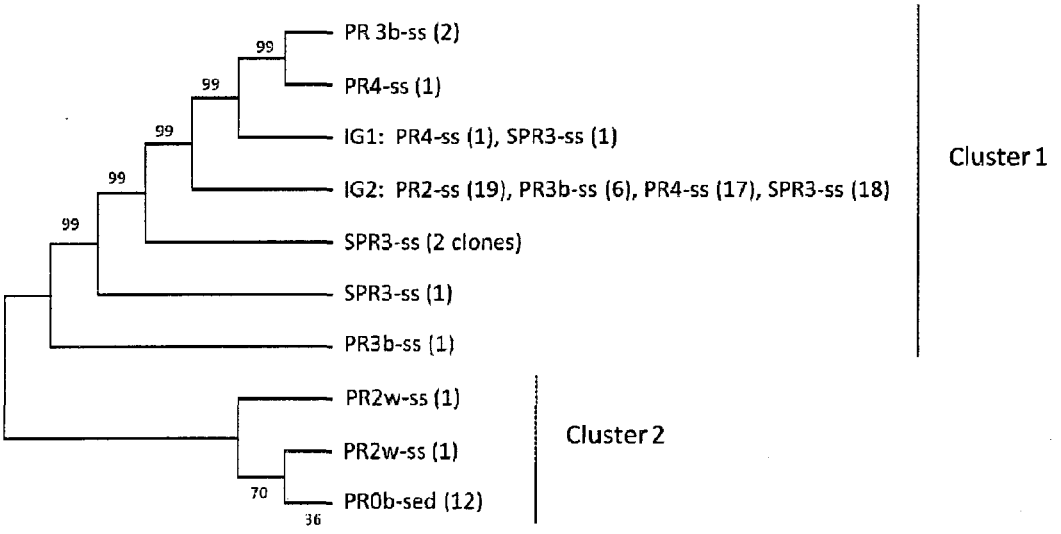

Figure 3.7: Evolutionary relationships of the tet(W) gene clone libraries of river samples. The evolutionary history of the functional tet(W) gene was inferred using the Neighbor-Joining method (Satiou and Nei 1987). The optimal tree with the sum of branch length = 0.07522124 is shown. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There were a total of 113 positions in the final dataset. Phylogenetic analyses were conducted in MEGA4 (Tamura et al. 2007). Numbers above nodes represent bootstrap values for those nodes.

CHAPTER 4: CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

In this study, ARG profiles of pristine river and putative source environments were characterized and found to be unique. Methods were developed to classify river samples according to the distribution, frequency of detection, and phylogenetic diversity of tet and sul ARG. Extensive site analysis of the study region was then used in classification and regression analyses to identify spatial variables explaining the variability in the ARG profiles of river samples. This approach was used to identify ARG profiles that were particularly effective for classification. Specifically, sul(I), tet(C), tet(S), as well as the groups of tet genes: tet group CEO and tet group HQST, were recognized as having value for future study.

Quantities of ARG, as determined by Q-PCR, showed initial success in that they could be correlated with spatial variables in linear regression analysis. Additional Q-PCR analyses could be performed, focusing for example on tet(C), which was more dominant in WWTPs and tet(S) which was only detected in agricultural samples. These may provide a more realistic measure of impact in an environment, given that they can vary over much greater orders of magnitude than frequency of detection.

The results of this analysis also indicated that transport of ARG from source environments may be the more dominant mechanism of ARG proliferation in river environments, as compared to selection of resistant native river bacteria. Additional mesocosm and field studies could be carried out to determine how both antibiotic resistant bacteria and naked ARG can be transported through rivers. One potential field study that could be performed is the monitoring

of sites downstream of WWTPs at varying distances from the discharge point. The tet genes (C), (E), and (O) may be good choices for quantification of ARG impact in WWTP discharge. For this proof-of-concept study, where source classification is as crucial, sul(I) may also be a good ARG to target. The spatial distribution of ARG quantities at various lengths from the discharge point could be spatially correlated and used to determine how distance affects the quantity of ARG in downstream river sites.

Additionally, the phylogenetic aspect of this work provided some intriguing results, but the methods are hampered by the cost of materials and extensive time required for the analysis. Profiling methods such as denaturing gradient gel electrophoresis (DGGE) and single stranded conformation polymorphism (SSCP) have been used for functional gene analysis and could be extended to the study of phylogeny of ARG. A more straightforward approach is to develop the restriction fragment length polymorphism (RFLP) analyses presented here into an automated terminal restriction fragment length polymorphism analysis. Since the RFLP analysis did show the same trends as phylogenetic analysis, this would serve to reduce the labor and cost involved and allow for more sites to be characterized.

Another outcome of this work is the creation of the spatial database for the study region of the South Platte River basin. In the past, sites were chosen based on visual characterization of a study area, but the geospatial database could allow a more appropriate distribution of sampling sites to be determined which may improve the power of these methods in future analyses.

Additionally, since WWTPs were implicated as a dominant source of ARG, this points to the need for more studies to determine what treatment approaches are effective in destroying not only intracellular ARG, but also extracellular ARG. Also, more studies need to examine the differences between ARG as either intracellular or extracellular entities. It is crucial to consider

the biology of microorganisms when developing these approaches since often times typical treatment techniques enhance the transport and spread of mobile genetic elements such as ARG.

APPENDIX 1: SUPPLEMENTAL TABLE AND FIGURES FOR CHAPTER 2

Table A1.1: PCR primers used in this study

Gene	Primer	Annealing temperature (°C)	Amplicon length (bp)	Sequence (5' →3')	References
Sulfonamide resistance genes					
sul(I)	FW	55.9	163	CGCACCGGAAACATCGCTGCAC	Pei et al. 2006
	RV			TGAAGTTCCGCCGCAAGGCTCG	
sul(II)	FW	60.8	191	TCCGGTGGAGGCCGGTATCTGG	Pei et al. 2006
	RV			CGGGAATGCCATCTGCCTTGAG	
Tetracycline efflux resistance genes					
tet(B)	FW	61	206	TACGTGAATTTATTGCTTCGG	Aminov et al. 2002
	RV			ATACAGCATCCAAAGCGCAC	
tet(C)	FW	70	207	GCGGGATATCGTCCATTCCG	Aminov et al. 2002,
	RV			GCGTAGAGGATCCACAGGACG	Koike et al. 2007*
tet(E)	FW	61	199	GTTATTACGGGAGTTTGTTGG	Aminov et al. 2002
	RV			AATACAACACCCACACTACGC	
tet(H)	FW	66	185	CAGTGAAAATTCACTGGCAAC	Aminov et al. 2002,
	RV			ATCCAAAGTGTGGTTGAGAAT	Koike et al. 2007*
Tetracycline ribosomal protection protein resistance genes					
tet(M)	FW	55	171	ACAGAAAGCTTATTATATAAC	Aminov et al. 2001
	RV			TGGCGTGTCTATGATGTTCAC	
tet(O)	FW	60	171	ACGGARAGTTTATTGTATACC	Aminov et al. 2001
	RV			TGGCGTATCTATAATGTTGAC	
tetB(P)	FW	46	169	AAAACTTATTATATTATAGTG	Aminov et al. 2001
	RV			TGGAGTATCAATAATATTCAC	
tet(Q)	FW	63	169	AGAATCTGCTGTTTGCCAGTG	Aminov et al. 2001
	RV			CGGAGTGTCAATGATATTGCA	
tet(S)	FW	50	169	GAAAGCTTACTATACAGTAGC	Aminov et al. 2001
	RV			AGGAGTATCTACAATATTTAC	
tet(T)	FW	46	169	AAGGTTTATTATATAAAAGTG	Aminov et al. 2001
	RV			AGGTGTATCTATGATATTTAC	
tet(W)	FW	64	168	GAGAGCCTGCTATATGCCAGC	Aminov et al. 2001
	RV			GGGCGTATCCACAATGTTAAC	
	1194RV	60	1152	CGACAGCAAAGCGGAAACA	Koike et al. 2007

* Different PCR conditions from original publication. In this study, the most recent conditions were used.

Table A1.2a: Presence of 13 ARG in pristine PR and FRU

				Tetracycline Efflux Resistant Genes				Tetracycline RPP Resistant Genes							Sulfonamide Resistant Genes	
	Site	Media	Date	tet(B)	tet(C)	tet(E)	tet(H)	tet(M)	tet(O)	tetB(P)	tet(Q)	tet(S)	tet(T)	tet(W)	sul(1)	sul(2)
pristine PR	PR0a	sed	Oct 2006	-	-	-	-	-	-	-	-	-	-	+	-	+
		ss		-	-	-	-	+	+	-	-	-	-	-	-	-
		sed	Feb 2007	-	-	-	-	-	-	-	-	-	-	-	-	-
		ss		-	-	-	-	-	-	-	-	-	-	+	-	-
		sed	May 2007	-	-	-	-	+	-	-	-	-	-	+	-	-
		ss		-	-	-	-	+	-	-	-	-	-	+	-	-
		sed	Oct 2007	-	-	-	-	-	-	-	-	-	-	-	-	-
		ss		-	-	-	-	+	-	-	-	-	-	-	-	+
	PR0b	sed	Oct 2006	-	-	-	-	-	-	-	-	-	-	-	-	-
		ss		-	+	-	-	+	+	-	-	-	-	-	-	-
		sed	Feb 2007	-	-	-	-	+	-	-	-	-	-	+	-	-
		ss		-	-	-	-	-	-	-	-	-	-	-	+	-
		sed	May 2007	-	-	-	-	+	-	-	-	-	-	-	-	+
		ss		-	-	-	-	-	-	-	-	-	-	-	-	-
		sed	Oct 2007	-	-	-	-	-	-	-	-	-	-	-	-	-
		ss		-	+	-	+	-	-	-	-	-	-	-	-	-
	PR1	sed	Oct 2006	-	-	-	-	-	-	-	-	-	-	-	-	-
		ss		-	-	-	-	-	-	-	-	-	-	-	-	-
		sed	Feb 2007	-	-	-	-	-	-	-	-	-	+	+	-	-
		ss		-	-	-	-	-	-	-	-	-	-	-	-	-
		sed	May 2007	-	-	-	-	+	-	-	-	-	-	-	-	+
		ss		-	-	-	-	-	-	-	-	-	-	-	-	-
		sed	Oct 2007	-	-	-	-	-	-	-	-	-	-	-	-	-
		ss		-	-	-	-	-	-	-	-	-	-	-	-	-
FHRU	FHRU-L	sed	Oct 2008	-	+	-	+	-	-	-	-	-	-	-	+	+
		eff		-	+	-	-	-	-	-	-	-	-	-	+	-
	FHRU-S	sed		-	+	-	+	-	-	-	-	-	-	+	+	+
		eff		-	-	-	+	-	-	-	-	-	-	-	+	-

Table A1.2b: Presence of 13 ARG in WWTPs and AFO Lagoons

				Tetracycline Efflux Resistance Genes				Tetracycline RPP Resistance Genes							Sulfonamide Resistant Genes	
	Site	Media	Date	tet(B)	tet(C)	tet(E)	tet(H)	tet(M)	tet(O)	tet(B/P)	tet(Q)	tet(S)	tet(T)	tet(W)	sul(1)	sul(2)
WWTPs	WWTP1	as	Apr 2008	-	+	+	-	+	+	-	+	-	-	+	+	+
		bs		-	-	-	+	+	+	-	-	-	+	+	+	-
		as	Apr 2009	-	+	+	-	+	+	-	-	-	+	+	+	+
		bs		-	+	-	+	+	+	-	-	-	+	+	+	+
		eff		-	+	-	-	+	+	-	+	-	-	+	+	+
	WWTP2	as	Apr 2008	-	+	+	-	+	-	-	-	-	-	-	+	-
		as	Apr 2009	-	+	-	+	+	+	-	-	-	-	+	+	+
		eff		-	+	-	-	+	+	-	+	-	-	+	+	+
	WWTP3	as	Apr 2008	-	+	+	-	+	+	-	+	-	-	+	+	+
		as	Apr 2009	-	+	+	+	+	+	-	-	-	+	+	+	+
		eff		-	+	-	-	+	+	-	+	-	+	+	+	+
AFO Lagoons	BF1	sed	Oct 2006	-	+	-	+	-	+	-	+	-	+	+	+	+
		wc		-	-	-	+	-	+	-	+	-	-	+	+	+
		sed	Feb 2007	-	-	-	+	+	+	-	+	-	+	+	+	+
		wc		-	+	-	+	+	+	-	+	-	+	+	+	+
		sed	May 2007	-	+	-	+	+	+	-	+	-	+	+	+	+
		wc		-	-	-	+	+	+	-	-	-	+	+	+	+
		sed	Oct 2007	-	+	-	+	+	-	-	+	+	+	+	+	+
		wc		-	-	-	+	+	-	-	+	-	+	+	+	+
	BF2	sed	Oct 2006	-	+	-	+	+	+	-	+	+	+	+	+	-
		wc		-	-	-	+	+	+	-	+	+	+	+	+	+
		sed	Feb 2007	+	-	-	+	+	+	-	+	+	+	+	+	-
		wc		-	+	-	+	+	+	-	+	+	+	+	+	-
		sed	May 2007	-	-	-	+	+	+	-	+	+	+	+	+	+
		wc		-	+	-	+	+	+	-	+	-	+	+	+	+
		sed	Oct 2007	-	+	-	+	+	-	-	+	+	+	+	+	+
		wc		+	+	-	-	+	+	-	+	-	+	+	+	+
	LCD	sed	Oct 2006	-	+	-	+	+	+	-	+	+	+	+	+	+
		wc		-	+	+	+	+	+	-	+	+	+	+	+	+
		sed	Feb 2007	-	+	-	-	+	+	-	+	-	+	+	+	+
		wc		-	+	-	+	+	-	-	+	+	+	+	+	+
		sed	May 2007	-	+	-	+	+	+	-	+	-	+	+	+	+
		wc		-	+	+	+	+	+	-	-	-	+	+	+	+
		sed	Oct 2007	-	+	-	-	+	+	-	+	-	+	+	+	+
		wc		-	+	+	+	+	+	-	-	-	+	+	+	+
	LOD	sed	Oct 2006	-	+	-	-	+	+	-	+	+	+	+	+	+
		wc		-	-	+	+	+	+	-	+	+	+	+	+	+
		sed	Feb 2007	-	+	+	+	+	+	-	+	-	+	+	+	+
		wc		-	+	-	+	+	+	-	+	+	+	+	+	+
		sed	May 2007	-	+	+	+	+	+	-	+	-	+	+	+	+
		wc		-	+	+	+	+	+	-	+	-	+	+	+	+
		sed	Oct 2007	-	+	+	+	+	+	-	+	-	+	+	+	+
		wc		-	+	-	+	+	+	-	-	-	+	+	+	+
	SCD	sed	Oct 2006	-	+	-	+	+	+	-	+	+	+	+	+	+
		wc		-	+	-	+	+	+	+	+	+	+	+	+	+
		sed	Feb 2007	-	+	+	+	+	+	+	+	+	+	+	+	+
		sed	May 2007	-	+	-	+	+	+	-	+	-	+	+	+	+
		wc		-	+	-	+	+	-	-	-	-	+	+	+	+
		sed	Oct 2007	-	+	-	+	+	-	-	+	+	+	+	+	+
		wc		-	-	-	+	+	+	-	+	+	+	+	+	+
	SOD	sed	Oct 2006	-	+	-	-	+	+	-	+	+	+	+	+	+
		wc		-	+	+	+	+	+	-	+	+	+	+	+	+
		sed	Feb 2007	-	-	-	+	+	+	-	+	+	-	+	+	+
		wc		-	+	-	+	+	+	-	+	+	+	+	+	+
		sed	May 2007	-	+	+	+	+	+	-	+	-	+	+	+	+
		wc		-	+	-	+	+	+	-	+	-	+	+	+	+
		sed	Oct 2007	-	+	+	+	+	-	-	+	+	+	+	+	+
		wc		-	-	+	+	+	+	-	-	-	+	+	+	+

FIGURES:

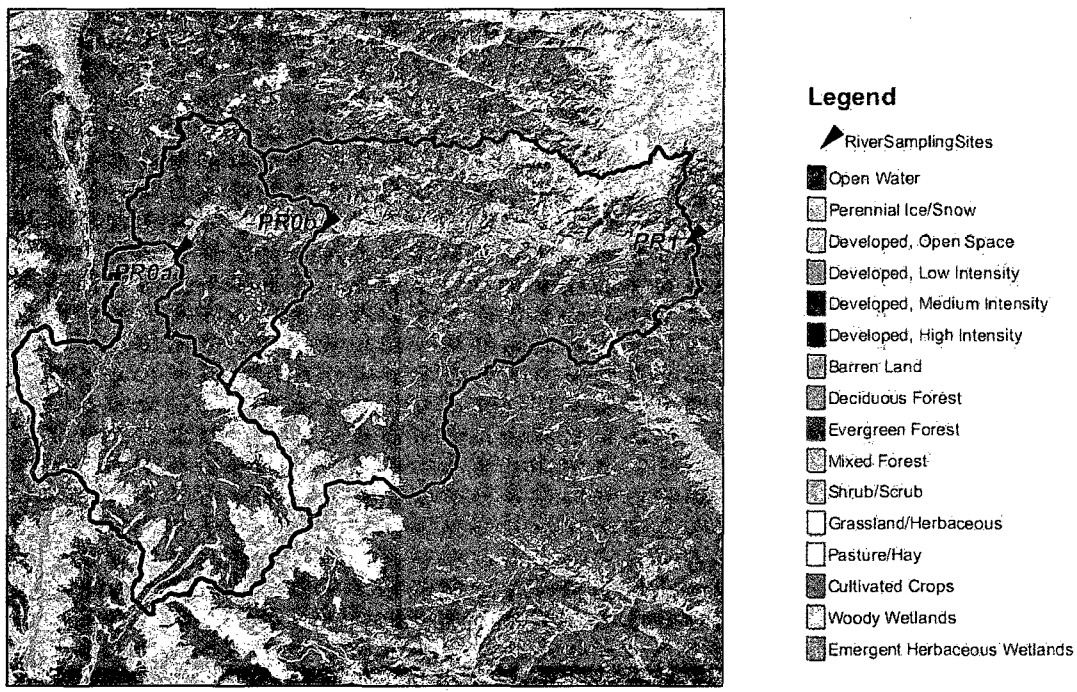

Figure A1.1: Map of pristine Poudre River sampling locations showing surrounding land use.

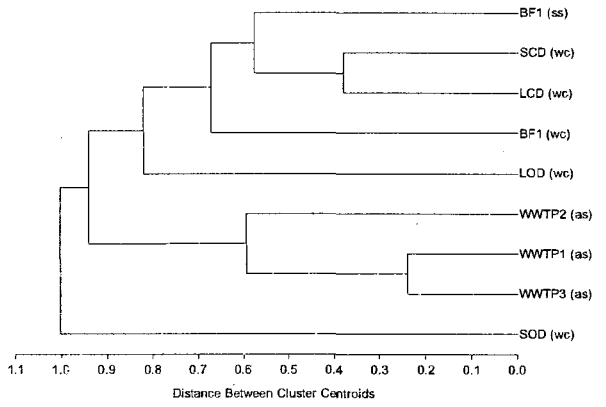

Figure A1.2: Cluster generated from analysis of RFLP patterns of the entire tet(W) clone libraries.

APPENDIX A.2: SUPPLEMENTARY TABLES AND FIGURES FOR CHAPTER 3

TABLES:

Table A2.1 Presence of 13 ARG in PR and SPR samples

			Tetracycline Efflux Resistant Genes				Tetracycline RPP Resistant Genes							Sulfonamide Resistant Genes	
Site	Media	Date	tet(B)	tet(C)	tet(C)	tet(M)	tet(M)	tet(O)	tet(B/P)	tet(Q)	tet(S)	tet(T)	tet(W)	sul(1)	sul(2)
PR2	sed	Oct 2006	-	-	-	-	-	-	-	-	+	-	-	-	-
		Feb 2007	-	-	-	-	-	-	-	-	-	-	-	-	-
		May 2007	-	-	-	-	-	-	-	-	+	-	+	-	-
		Oct 2007	-	-	-	-	-	-	-	-	+	-	-	-	-
	ss	Oct 2006	+	-	-	-	-	-	-	-	-	-	-	-	-
		Feb 2007	+	-	+	-	-	-	+	+	+	+	+	-	-
		May 2007	+	-	-	-	-	-	-	-	-	-	-	-	-
		Oct 2007	+	+	-	-	-	+	-	-	+	+	+	-	-
PR3a	sed	Oct 2006	-	-	-	-	-	-	-	-	-	+	-	-	-
		Feb 2007	-	-	-	-	-	-	-	-	-	-	+	-	-
		May 2007	-	-	-	-	-	-	-	-	-	+	-	-	-
		Oct 2007	-	-	-	-	-	-	-	-	-	+	+	-	-
	ss	Oct 2006	+	-	-	-	-	-	-	-	-	+	+	-	-
		Feb 2007	+	-	-	-	-	+	-	-	-	-	-	-	-
		May 2007	-	-	-	-	-	-	-	+	-	+	+	-	-
		Oct 2007	+	-	-	-	-	-	-	+	-	+	+	-	-
PR3b	sed	Oct 2006	-	-	-	-	-	-	-	-	+	+	-	-	-
		Feb 2007	-	-	-	-	-	-	-	-	-	+	-	-	-
		May 2007	+	-	-	-	-	-	-	+	-	+	+	-	-
		Oct 2007	-	-	-	-	-	-	-	-	-	+	+	-	-
	ss	Oct 2006	+	-	-	-	-	-	+	+	-	+	-	-	-
		Feb 2007	+	-	+	-	-	+	+	+	-	+	-	-	-
		May 2007	+	-	-	-	-	-	-	-	-	+	+	-	-
		Oct 2007	+	-	-	-	-	-	-	+	-	+	+	-	-
PR4	sed	Oct 2006	-	-	-	-	-	-	-	-	-	+	+	-	-
		Feb 2007	-	-	-	-	-	-	-	-	-	-	-	-	-
		May 2007	-	-	-	-	-	-	-	-	-	-	-	-	-
		Oct 2007	-	-	-	-	-	-	-	-	-	+	+	-	-
	ss	Oct 2006	-	-	-	-	-	-	+	-	+	+	+	-	-
		Feb 2007	+	-	-	-	-	+	+	+	+	+	+	-	-
		May 2007	-	-	-	-	-	+	-	+	-	+	+	-	-
		Oct 2007	+	+	-	-	-	+	-	+	-	+	+	-	-
PR5	sed	Oct 2006	-	-	-	-	-	-	-	-	-	+	+	-	-
		Feb 2007	-	-	-	-	-	-	-	+	+	+	+	-	-
		May 2007	-	-	+	-	+	-	-	-	-	+	+	-	-
		Oct 2007	-	-	-	-	-	-	-	-	-	-	+	-	-
	ss	Oct 2006	+	-	-	-	-	+	+	-	+	+	+	-	-
		Feb 2007	+	-	+	+	-	+	-	+	+	+	-	-	-
		May 2007	-	-	-	-	-	-	+	+	-	+	+	-	-
		Oct 2007	-	-	-	-	-	-	+	+	-	+	+	-	-
SPR1	sed	Oct 2006	+	-	-	-	-	-	-	+	-	+	+	-	-
		Feb 2007	+	-	-	-	-	-	-	-	-	+	+	-	-
		May 2007	-	-	-	-	-	+	+	+	-	+	-	-	-
		Oct 2007	+	-	-	-	-	-	-	-	+	+	-	-	-
	ss	Oct 2006	+	-	-	-	-	-	-	-	+	+	+	-	-
		Feb 2007	+	-	-	-	-	+	+	+	+	+	+	-	-
		May 2007	+	-	-	-	-	+	+	-	+	+	+	-	-
		Oct 2007	-	-	-	-	-	-	-	-	-	-	-	-	-
SPR2	sed	Oct 2006	-	-	-	-	-	-	-	+	-	-	-	-	-
		Feb 2007	+	-	-	-	-	-	+	-	+	-	+	-	-
		May 2007	+	-	-	-	-	-	-	-	-	+	+	-	-
		Oct 2007	-	-	-	-	-	-	-	-	-	+	-	-	-
	ss	Oct 2006	-	-	-	-	-	-	-	+	-	+	-	-	-
		Feb 2007	+	-	-	-	-	+	+	-	-	+	+	-	-
		May 2007	-	+	-	-	+	-	-	-	+	+	+	-	-
		Oct 2007	-	-	-	-	-	-	-	-	+	-	+	-	-
SPR3	sed	Oct 2006	+	-	-	-	-	-	-	+	+	-	-	-	-
		Feb 2007	+	-	-	-	-	-	-	-	-	-	-	-	-
		May 2007	-	-	-	-	-	-	-	+	-	+	+	-	-
		Oct 2007	-	-	-	-	-	-	-	-	-	+	+	-	-
	ss	Oct 2006	+	-	-	-	-	-	+	+	-	+	-	-	-
		Feb 2007	-	-	+	-	-	+	+	+	+	+	+	-	-
		May 2007	+	+	-	-	+	+	-	+	-	+	+	-	-
		Oct 2007	-	-	-	-	-	-	-	+	-	+	+	-	-

FIGURES:

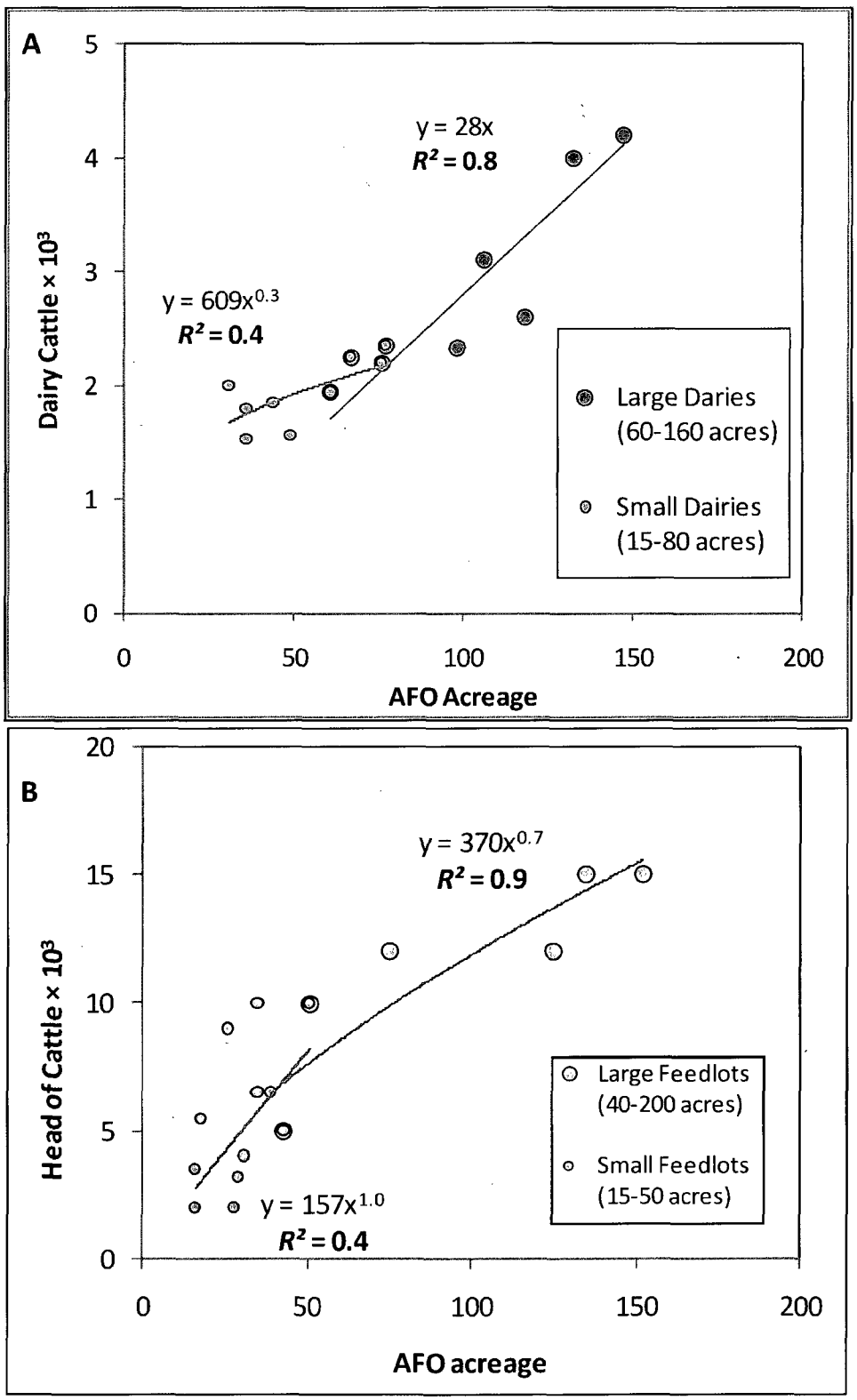

Figure A2.1 Regression of animal counts and AFO area for dairies (A) and beef feedlots (B).

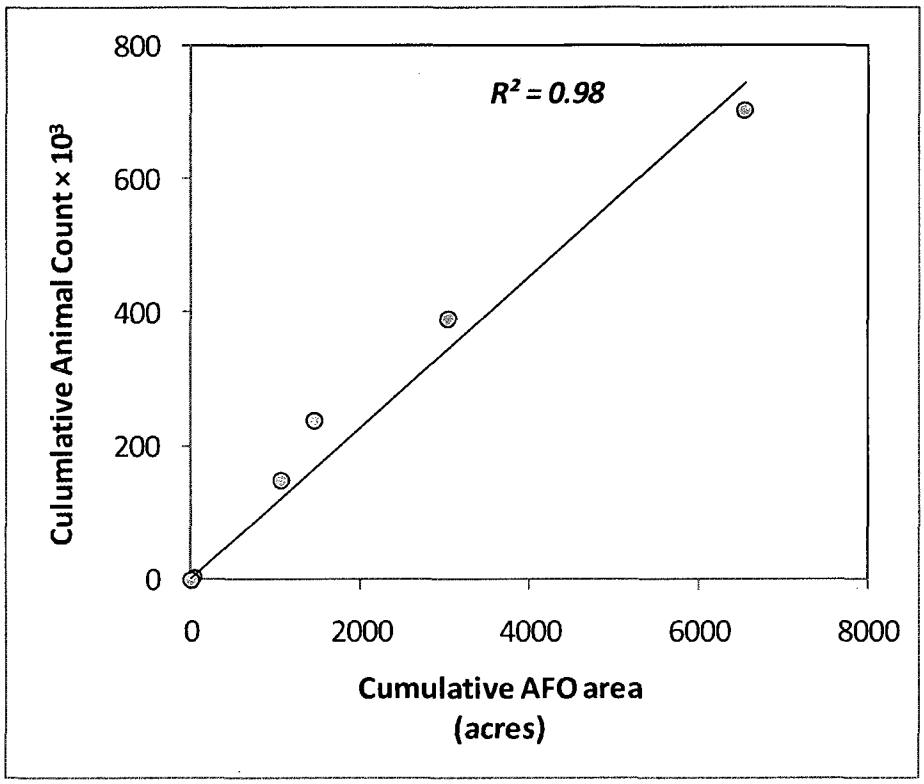

Figure A2.2: Correlation between cumulative AFO area and cumulative animal counts influencing for the sub-watershed of each site.

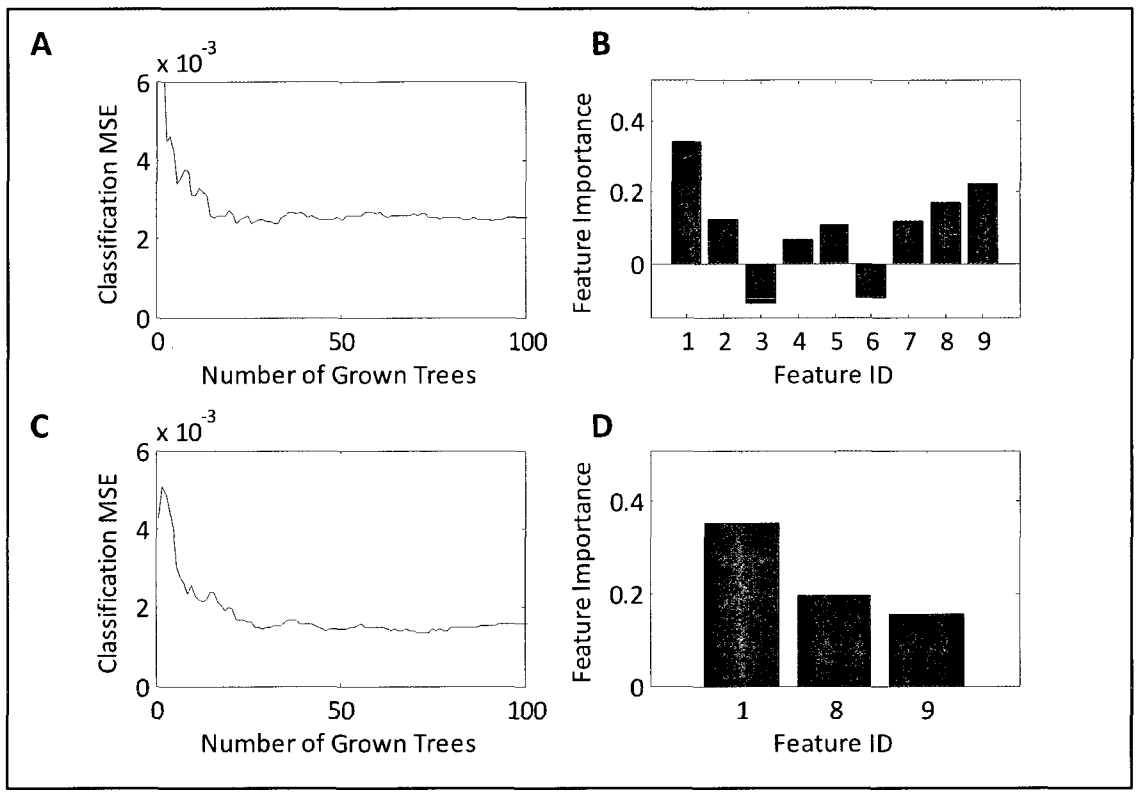

Figure A2.3: Bootstrapping and aggregation of Classification and Regression Tree (CART) Analysis for the overall frequency of detection of tet(H), tet(Q), tet(S) and tet(T). Regression tree analysis was performed using the frequency of detection of tet(H), tet(Q), tet(S) and tet(T) as the response variable and the following spatial indicators as explanatory variables: the inverse distance weighted counts of WWTPs (1), AFOs (2), FRUs (3), and the interaction of the weighted counts with total inverse distance for the same sources (4-6), the interaction between WWTP and AFO for IDW counts (7) and for IDW count*inverse distance(8) and the normalized ratio of cumulative capacity of WWTPs to total count of AFOs (9). One thousand boot-strapped trees were created but classification error is only shown for the first 100 trees. (A) error associated with the number of trees grown for all nine variables; (B) importance of explanatory variables determined by increases in mean-squared error when features undergo permutations; (C) error associated with the number of trees grown for the reduced model (D) feature importance of variables in the reduced model.

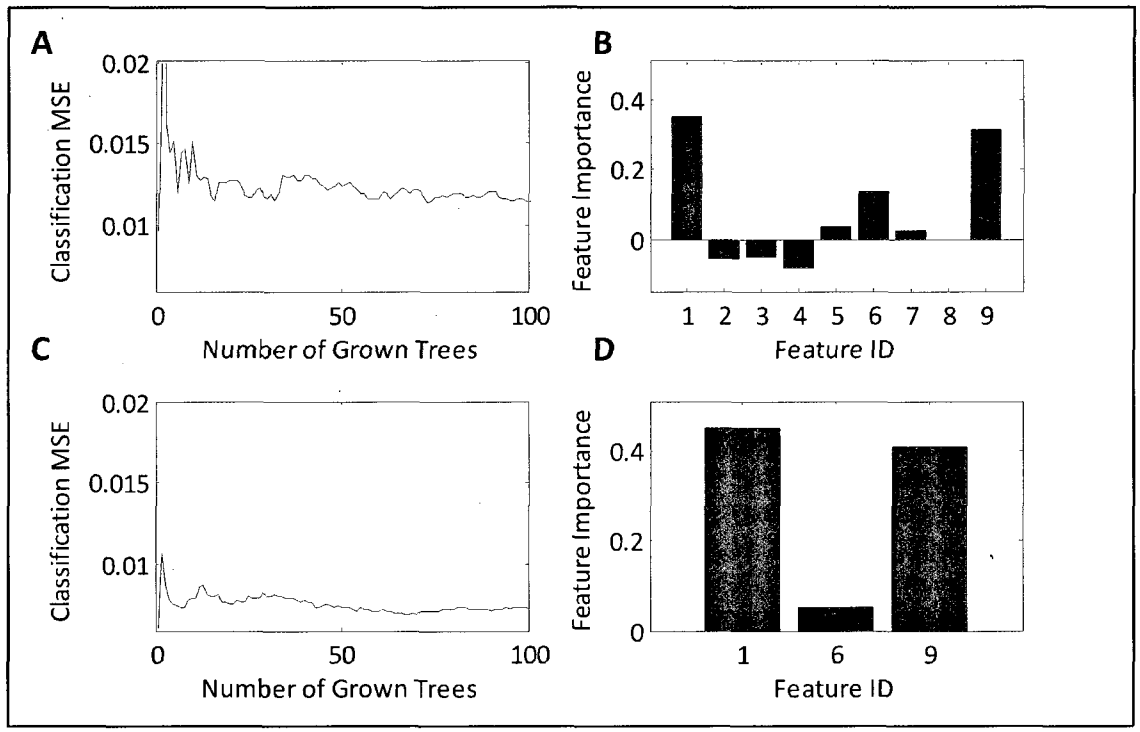

Figure A2.4: Bootstrapping and aggregation of Classification and Regression Tree (CART) Analysis for the overall frequency of detection of tet(C), tet(E), and tet(O). Regression tree analysis was performed using the frequency of detection of tet(C), tet(E), and tet(O) as the response variable and the following spatial indicators as explanatory variables: the inverse distance weighted counts of WWTPs (1), AFOs (2), FRUs (3), and the interaction of the weighted counts with total inverse distance for the same sources (4-6), the interaction between WWTP and AFO for IDW counts (7) and for IDW count*inverse distance(8) and the normalized ratio of cumulative capacity of WWTPs to total count of AFOs (9). One thousand bootstrapped trees were created but classification error is only shown for the first 100 trees. (A) error associated with the number of trees grown for all nine variables; (B) importance of explanatory variables determined by increases in mean-squared error when features undergo permutations; (C) error associated with the number of trees grown for the reduced model (D) feature importance of variables in the reduced model.

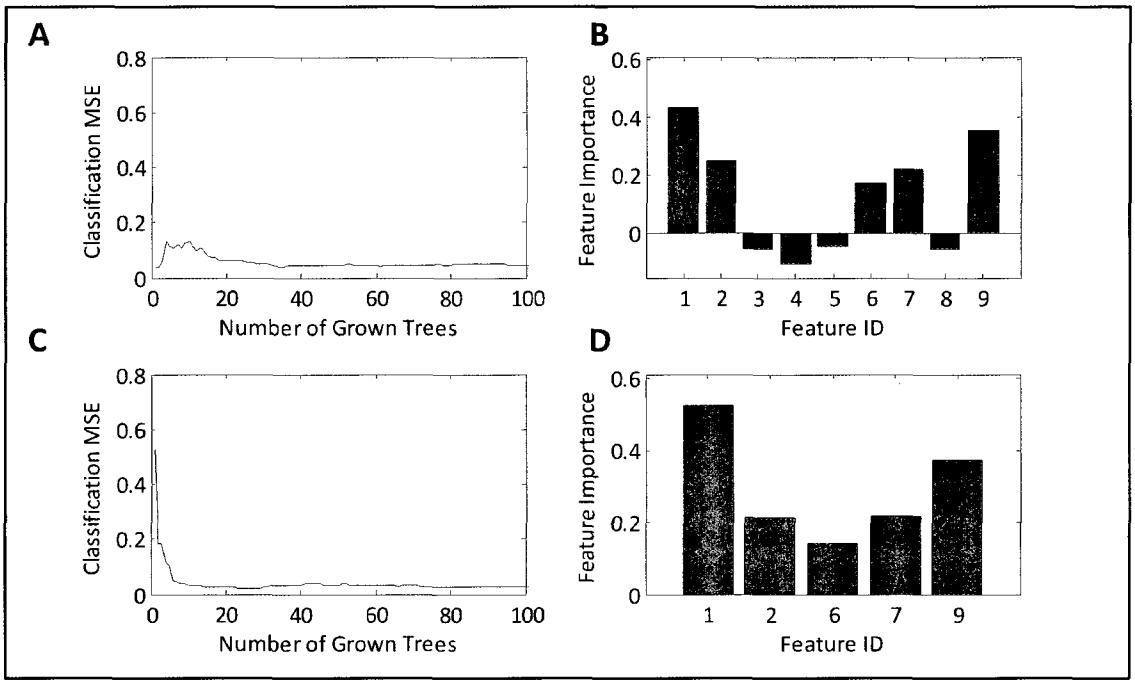

Figure A2.5: Bootstrapping and aggregation of Classification and Regression Tree (CART) Analysis for the frequency of detection of sul(I). Regression tree analysis was performed using the frequency of detection of sul(I) as the response variable and the following spatial indicators as explanatory variables: the inverse distance weighted counts of WWTPs (1), AFOs (2), FRUs (3), and the interaction of the weighted counts with total inverse distance for the same sources (4-6), the interaction between WWTP and AFO for IDW counts (7) and for IDW count*inverse distance(8) and the normalized ratio of cumulative capacity of WWTPs to total count of AFOs (9). One thousand boot-strapped trees were created but classification error is only shown for the first 100 trees. (A) error associated with the number of trees grown for all nine variables; (B) importance of explanatory variables determined by increases in mean-squared error when features undergo permutations; (C) error associated with the number of trees grown for the reduced model (D) feature importance of variables in the reduced model.

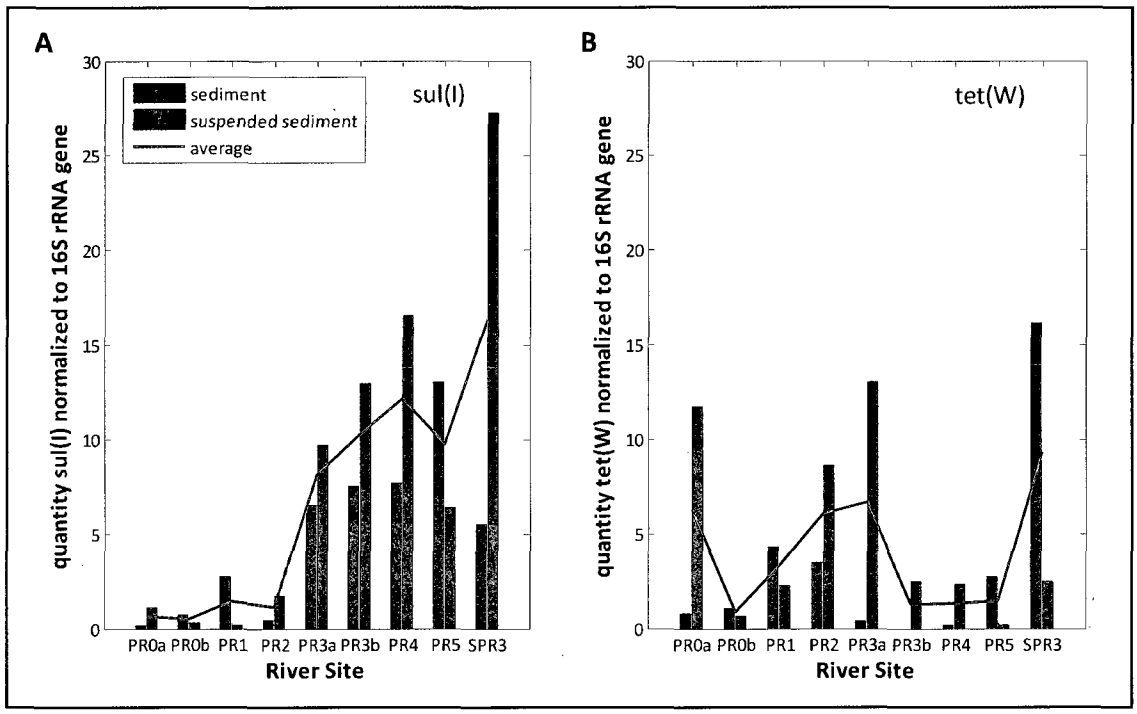

Figure A2.6: Quantities of sul(I) and tet(W) normalized to 16S rRNA quantities in Poudre River sediment and suspended sediment. The abundance of sul(I), tet(W) genes was quantified in river samples with Q-PCR. Absolute copy numbers were normalized to the copies of 16S rRNA to control for variations in extraction efficiency and inhibition. Bars represent the average normalized copies of ARG for all four sampling dates. The line represents the average of all ARG normalized copies of ARG for all sampling dates and media.

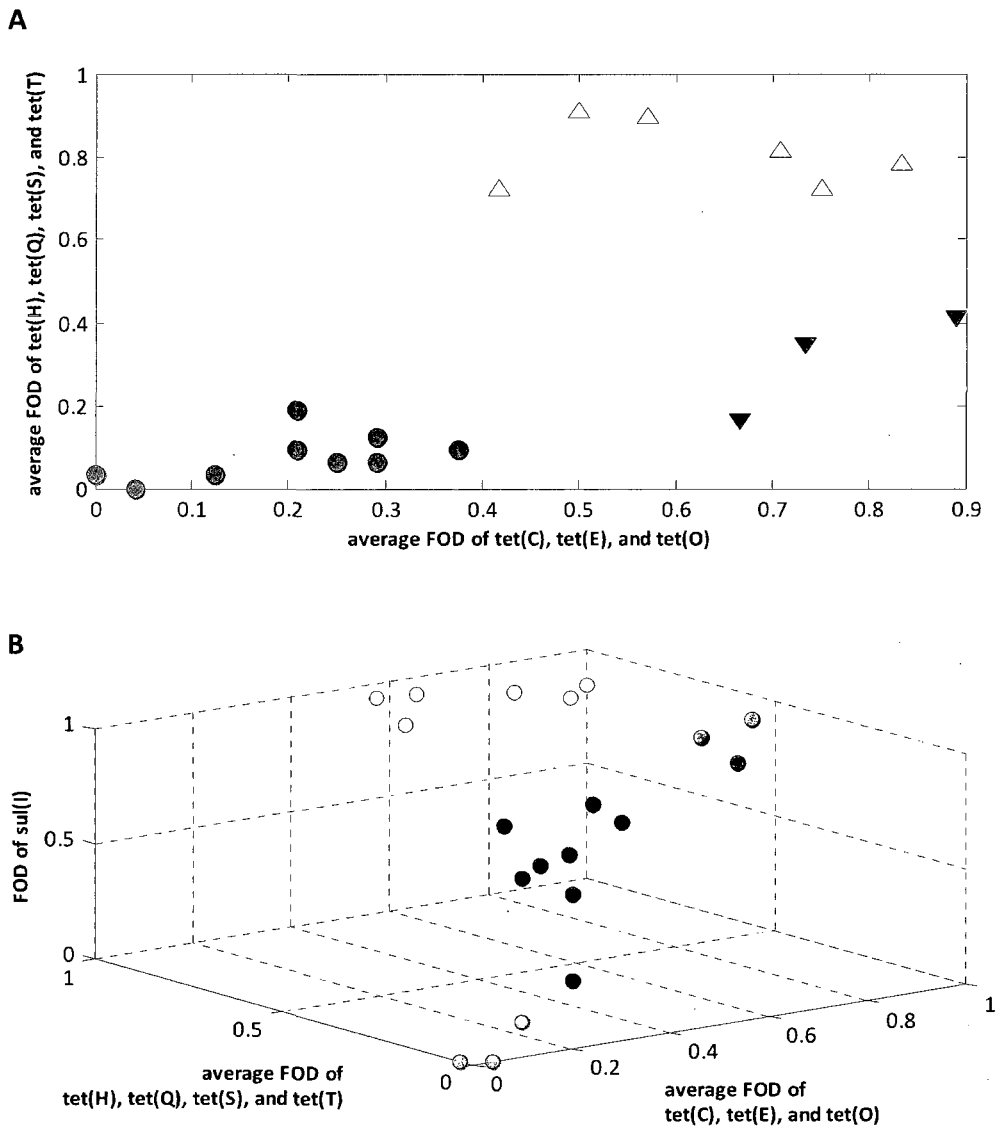

Figure A2.7: Discriminant analysis of ARG profiles of river sites by unnormalized frequency of detection of ARG indicator variables. The average frequency of detection of tet group CEO and tet group HQST were used as the variables for discriminant analysis of river sites (A). The plot is extended to 3-D with the addition of the normalized FOD of sul(I) as an indicator variable (B).

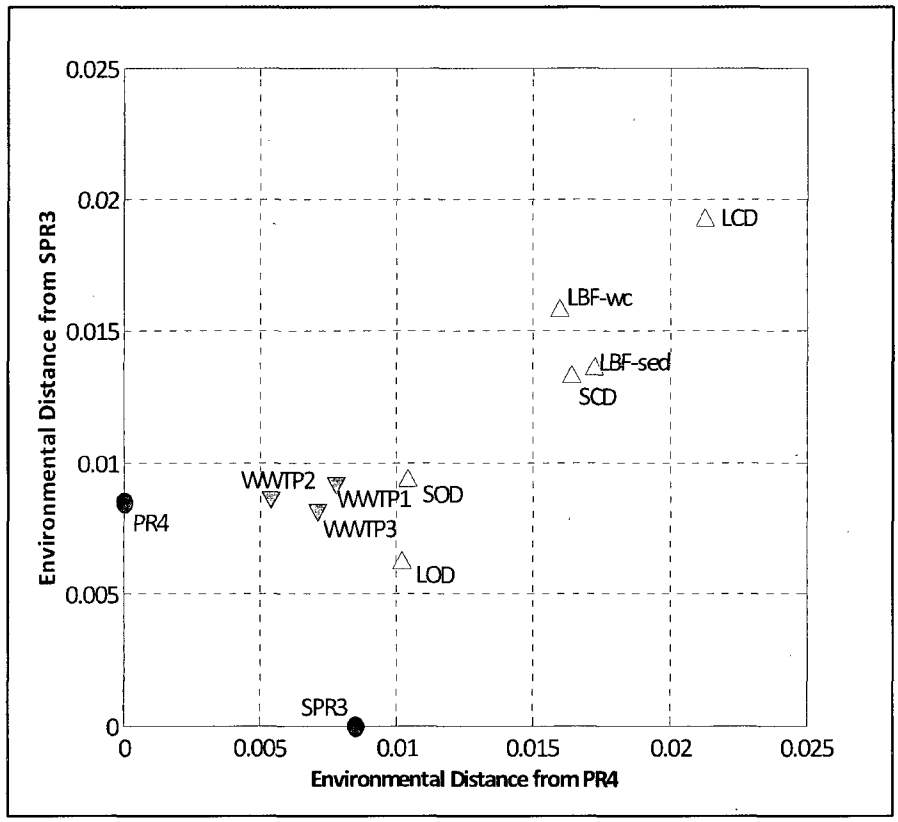

Figure A2.8: Plot of environmental distances of source environments PR4 and SPR3 based on phylogenetic relatedness of tet(W) profiles. Environmental distances were calculated using UniFrac (Lozupone and Knight 2005). The distance of a source from PR4 is equal to the horizontal distance from the Y-axis. Distances from SPR3 are vertical distances from the x-axis. [Note: the plot can only be used to determine distances to PR4 or SPR3, but not between different source environments].