

# Student Abstracts

## Undergraduate Biomathematics Day

Buffalo and Geneseo SUNY Colleges

April 18-19, 2008

*NOTE: Some of the oral presentations also have a poster*

### Friday Evening

7:15pm–8:05pm

#### *Bioinformatic Studies of the Short-Chain Oxidoreductase Enzymes*

Robert Huether

SUNY at Buffalo, Hauptman Woodward Medical Research Institute

The short-chain oxidoreductase (SCOR) enzymes are one of the most ancient and widely spread protein superfamilies. A subset of the SCOR superfamily (TGYK), that have a semi-conserved sequence pattern of TGxxxGxG...YxxxK, has over 10,500 members in the gene bank. There is at least one member of the TGYK family in all genomes sequenced to date. Including all bacterial and eukaryotic species as well as some viruses. They catalyze the oxidation, reduction and epimerization of >300 substrates. We have identified a highly conserved fingerprint (>70% ID) of 40 residues in the TGYK subfamily. The 40 fingerprint residues are critical to catalysis, cofactor binding, protein folding and oligomerization. SCOR crystal structures and sequence alignments reveal that an aspartic acid or an arginine residue in adjacent positions in the sequence determine NAD and NADP binding, respectively. This makes it possible to predict cofactor binding and to infer oxidative or reductive functional preference in 77% of the 10,500 proteins of the TGYK family. The remaining 23% of the family are primarily bacterial and use at least five alternate molecular means of NAD or NADP recognition. There are 35 members of the TGYK family in the human genome. TGYK members play essential roles in the body to maintain key metabolic processes for sugars and steroids. Various members of the human TGYK family are implicated in diseases including cancer, diabetes, hypertension, pseudohermaphroditism, Alzheimer's and polycystic kidney disease. To better understand the relationship between the disease and the protein we are subdividing the 10,500 members in the TGYK family on 1) the bases of patterns of co-variation in quasi-conserved residues in 3 loops that contact the cofactor and substrate and 2) characteristic variation in the lengths between the 40 residues that contain information about folding. Through the subgroups, we can identify substrates and correlated sequence changes in the subgroups with specific diseases as well as trace the course of their evolution.

---

#### *Algorithmic Approach to Understating the Function and Evolution of the Genetic Code*

Sanjay Connare

Hauptman Woodward Medical Research Institute

The short-chain oxidoreductase or SCOR family, has over 10,000 members, and catalyzes over 300 substrates. Every species whose entire genome has been sequenced to date has from 4 to 50 members of this family present in its genome, which covers 2 billion years of evolution. Of the 10,500 members, only 1,000 are annotated with a known substrate and some of these annotations are incorrect. Current bioinformatic techniques are not equipped to deal with the intricacies of annotating the other 9,500. We have developed a set of bioinformatic tools that apply various in-house analysis techniques on this family. The 3 primary and most powerful tools are what we call the "Codon Analysis Program (CAP)", "Prosit Covariance Program (PCP)" and the "Prosit Statistical Program (PSP)". These programs are all part of a recursive process to subdivide the 10,500 SCOR proteins into species/substrate dependent families. While the order in which these programs are run against protein families varies, PSP is typically run first. The PSP expands upon the limitations of the Prosit tool. The

Prosite tool aligns sequences that match a pattern but it is unable to report the amino acid usage in each position for all of the aligned proteins. For instance in the SCORs, position 152, 96% of these proteins have a Tyrosine while the other 2% have a Threonine. This further breaks down families and allows us to continuously improve our search vectors when scanning the genomic databases. The PCP applies a unique method of studying how the genetic makeup depends directly on the amino acid in each position. The PCP allows the user to gather a set of aligned sequences, and then see how different combinations of positions co-vary with one another. This is very powerful in seeing selecting and grouping of proteins by their substrate binding amino acids. Lastly, the CAP analyzes proteins by their nucleotide content. It reports the codon and amino acid usage for every frame in a nucleotide sequence. We have data to suggest that the most ancient proteins have DNA sequences that retain more than one open reading frame (ORF) and that proteins with multiple ORFs also have a severe codon bias towards the use of GC rich codons. The CAP has allowed us to find supporting data showing the longer the species has been existence the more severe its GC codon bias is. As this process repeats recursively, proteins divide into substrate specific subgroups, eliminating false positives. This generates a precise alignment of orthologs. We can then compare these observed patterns and see exactly where insertions or deletions and mutations have occurred in this family throughout the course of evolution. The impact of these bioinformatics tools is significant because it pushes the limits of current thinking in regards to evolution and it takes a different approach to analyzing available data. The anticipated results from this research include helping to predict and pinpoint mutations and malfunctions of proteins that cause diseases such as cancer, Alzheimer's, and high blood pressure. The results from this research will also help in studying changes in bacteria and viruses that mutate frequently and are becoming immune to first line drugs. This will not only show how the initial drugs became ineffective but will help provide information on how to develop new drugs to address said mutations.

---

## **Saturday Morning**

8:30am–9:45am

### ***Degree-correlated scale-free networks and epidemics***

Eamon O'Dea  
SUNY Geneseo

Mathematical models provide insight into the way a disease spreads through a population. The structure of the contact network can dramatically affect the dynamics of these models. We investigate scale-free networks in which negative correlations in the connectivity of certain classes of hosts leads to more hosts becoming infected. We will discuss how such results follow from the discrete time stochastic and percolation models used, as well as the practical significance of the results.

---

### ***Island biogeography with active dispersal***

Temitope Brotherson<sup>1</sup>, Karina Aliaga<sup>1</sup>, and Gareth Russell<sup>1,2</sup>

1. New Jersey Institute of Technology
2. Rutgers University

MacArthur and Wilson's theory of island biogeography models the effect of distance and area of an island on the immigration and extinction rates of species, and therefore on equilibrium species richness. Like many classic spatial models in ecology, it assumes 'passive' dispersal and population dynamics. Here, we introduce the concept of emigration as a choice-based disappearance mechanism whereby species play an active role in deciding whether to leave, or remain on, an island. We link the emigration rate to both the area and distance of an island, using a variety of functional forms, and demonstrate how this function alters traditional species-area and species-distance relationships. We show how our modified model explains some otherwise puzzling patterns found in real data.

---

***The presence of cross-hybridized DNA duplexes in pools is determined using the relationship between SYBR Green fluorescent absorbance and temperature***

Ellen Schmidt and Sarah Orbesen

SUNY Geneseo

Watson-Crick (WC) duplexes, joining complementary sequences of DNA in opposite orientations, are stable, energetically favorable double stranded segments of nucleic acid. The presence of non-WC, or cross-hybridized (CH) duplexes, produces less favorable, or more positive free energies ( $\Delta G$ ). CH duplexes can be detected within a pool of hybridized DNA using the fluorescent dye SYBR Green I. By binding to the minor groove of the DNA helical structure, fluorescent emission is increased in the presence of CH duplexes. This separates pools into two groups that are distinguished by the presence or absence of a positive fluorescent signal, and are labeled positive and negative, respectively. We determine a systematic method to determine whether pools are positive or negative, based on the relationship between temperature and absorbance. Using R-Squared to determine linearity; we can analytically predict the presence of CH duplexes within pools. We then use a group testing method to identify CH duplexes from the information gained from the pooling experiment.

---

**Saturday Morning**

10:00am–11:15am

***Sequence analysis of the NAA subset of TGYK SCORs***

Dana Hogan

Hauptman Woodward Medical Research Institute

Short-Chain Oxidoreductase (SCOR) enzymes catalyze oxidation, reduction, and epimerization reactions. We have been able to identify 10,500 members of the largest subgroup in the SCOR family (the TGYK subgroup) that are essential to human growth and development and have been implicated in Breast cancer, Alzheimer's disease, and other afflictions. They are present in every organism, for which genomes have been sequenced (over 600), and are one of the most ancient and largest families of enzymes. Although it is possible to identify members of the TGYK group, the primary function and substrate of the majority of them and their role in health and disease remain unknown or ambiguous. Our goal is to fully characterize and annotate all members of the TGYK group according to function (oxidation or reduction), cofactor binding (NAD, NADP), substrate (sugars, alcohols, flavones, steroids, etc.) functional form (dimer, tetramer) and details of catalytic mechanism (proton and hydride transfer). This requires not only maximized true hits and excluded false hits but accurate alignment of all members of the TGYK family. To do this, many search vectors based on the TGYK family of SCORs fingerprint were used. A variant of the NAG sequence that interacts with the cofactor in which the glycine is replaced by an alanine, NAA is present in 549 sequences of the overall 10,500 sequences in the TGYK subfamily. The distribution of cofactor recognition residues Arginine(R) and Aspartic Acid (D) in the NAA set is 43% R and 42% D, indicating fairly even NADP and NAD binding, respectively. NAD is generally associated with dehydrogenase reactions while NADP is associated with a reductase activity. In the NAA subset, the population of residues that commonly control dimer formation is 25% F-N-F, 15% W-N-F, and 9% L-N-F. Following convention, this percentage of aromatics at the dimer interface is common. Of the 9 specific residues that commonly define substrate, one sequence appears ten times (FMELVVPMQ). However, the first five residues have 9 sequences with ten or more hits. NAA is mostly found in betaproteobacteria (22.7%) and alphaproteobacteria (21.8%). I am currently analyzing the other 60% for patterns. In conclusion, I hope to further analyze my results and fully categorize the NAA subset.

---

## ***Sequence and Substrate Determining Analysis of the AAG Subset of TGYK SCORs***

Jimmitti Teysir

Hauptman Woodward Medical Research Institute

The Short-Chain Oxidoreductase (SCOR) enzymes are a superfamily of proteins that are between one and three billion years old. SCORs are virtually ubiquitous within all species, catalyzing many different substrates and playing significant roles in bodily growth and development. Furthermore, the malfunction of these enzymes has been linked to various human diseases including Alzheimer's, breast cancer and polycystic kidney disease. The majority of SCORs have unknown cofactors and substrates and so a key objective is to evaluate the entire superfamily and determine the role, specifically the substrate, for each enzyme. We focused our study on the TGYK subfamily to which 70% of SCORs belong. The TGYK subfamily is characterized by amino acids in key positions that define protein functions. The common amino acid sequence "VNNAG" was examined and variants of the last three amino acids were noted. It is hypothesized that a correlation exists between these variants and cofactor specificity, substrate specificity and dimer formation in the TGYKs. 236 SCORs containing the AAG variant were found in the Swiss-Prot/Trembl and PDB databases. SCORs generally use NAD to catalyze oxidation reactions and NADP for reduction reactions. Previous studies have shown that an Aspartic Acid (D) in position 37 determines NAD binding while an Arginine (R) in position 38 determines NADP binding. The population of D and R residues indicates that 59.7% of the AAG set bind NAD and 28.6% bind NADP. The data also implies that these enzymes came about later in evolution as D and R are generally found in eukaryotic SCORs. Strong patterns also existed with the residues that govern dimer formation in  $\alpha$ -helix 5. Approximately 35% of AAG SCORs contain the aromatic group W-N-F and 13% contain F-N-F at the dimer interface. Finally, 9 residues in 3 Hypervariable loops form the substrate binding pocket in TGYKs. The AAG set did not have its own characteristic set of substrate determining residues. However, certain combinations that arose, such as TD-NGQ-FMTV, are known substrate fingerprints which allow us to identify them with correct annotations rather than those more ambiguous annotated in the gene bank. Other combinations were found frequently, including LM-RGM-SMQL, which occurred 24 times, and QE-GIC-YMVL, which occurred 10 times. These numbers suggest that such occurrences are not random and identify a substrate specific group. Further examination revealed that these combinations were species-specific and that the enzymes also had a similar set of cofactor determining residues and dimer residues.

---

## ***Mathematical Modeling of the BMP4 and FGF Signaling Pathways during Neural and Epidermal Development in Xenopus***

Binh Nguyen, Tung Bui, Marsida Lisi, and Rashmi SubbaRao

University of Houston, Downtown

During embryonic development, ectodermal cell fate in *Xenopus* is determined by the mitogen-activated protein (MAP) kinase and bone morphogenetic protein-4 (BMP-4) signaling pathways. It has been observed that an increase in MAPK kinase activity in the dorsal ectoderm inhibits the activity of the BMP-4 pathway in this region of the embryo. This inhibition induces neural development in the dorsal ectoderm by preventing the SMAD 1/4 transcription factor complexes from migrating into the nucleus. In an attempt to further understand the interaction of these two pathways, a mathematical model consisting of nonlinear ordinary differential equations is developed. Initial simulations of the model aim to reproduce previously known results from biological experiments. By adjusting the key biochemical parameters in the simulations, the model will be used to make experimentally testable predictions on the threshold of MAPK activity required for neural induction and the robustness of this signaling cross-talk.

---

## **Saturday Morning**

11:30am–12:45pm

### ***Detecting the driving forces behind species diversity patterns: a modeling approach***

Colin Kremer<sup>1</sup> and E. Binney Girdler<sup>2</sup>

1. SUNY Geneseo
2. Kalamazoo College

Community ecologists seek to understand the forces driving patterns of species diversity within ecological communities. Traditional niche theory suggests that diversity patterns are explained by the adaptation of species to best fill a particular ecological niche. A different theory explaining diversity patterns, called neutral theory, has arisen in the last decade, and suggests that many diversity patterns can be reproduced entirely as the result of dispersal and chance mortality, with the assumption that all species have identical competitive abilities.

To compare these two theories, a variety of new multivariate statistical methods have been developed to estimate the relative contributions of niche and neutral processes to determining spatial patterns of species diversity. To verify these methods, I have created an individual-based computer model in which plant species compete for resources on a grid and reproduce and colonize new habitats via dispersal. By controlling the shape of the dispersal kernels, as well as the species' competitive abilities, I can control the relative strength of niche and neutral forces driving diversity patterns within my model. Simulation data are then subjected to these new statistical methods to determine how well niche and neutral patterns are detected.

---

### ***A mathematical model on colony collapse disorder of the honeybee***

Christopher Milazzo and Adam Cesar

Buffalo State College

Colony collapse disorder is a phenomenon in which worker honeybees from beehives abruptly disappear. Our objective is to formulate a mathematical model and search for possible explanations for colony collapse disorder of honeybees in the United States. We use two contributing factors to the rapid disappearance of the entire colonies to develop the model and seek for possible answers. The first factor is the honey bees loss of direction resulting in the inability to navigate back to their hives once they have left; and the second is known viruses among honey bees; for example, acute paralysis virus and deformed wing virus.

---

### ***Human Papilloma virus***

Alison Green

Canisius College

Human Papillomavirus (HPV), a common sexually transmitted virus, is a collection of more than 100 viruses, some of which (called "high-risk" oncogenic or carcinogenic HPV) are associated with certain types of cancer. HPV 16 and 18 cause approximately 70% of all cervical cancer. An estimated 11,150 new cases of cervical cancer will develop in the U.S. in 2007. In 2006, the FDA approved the first vaccine to prevent cervical cancer, designated for females ages 9-26, called Gardasil. Studies have indicated Gardasil has a 95-100% success against HPV types 6, 11, 16, and 18. Many studies have shown that there exists a lack of HPV awareness, as well as knowledge of causes, effects and preventive measures. In this study we compare two strategies for controlling the spread of carcinogenic HPV in the population, while minimizing cost. The research is conducted via a cost analysis of a mandatory vaccination policy vs. a mass-media awareness campaign, each individually modeled with a system of differential equations. Mandatory (but not universal) vaccination includes individual-based education at the time of vaccination only, while the mass-media campaign is assumed to be ongoing. In both cases the education influences females to get vaccinated and/or reduce their sexual activity, but is of limited duration. We use qualitative analysis to derive the respective control reproductive numbers, and

numerical analysis to obtain the total costs of vaccination, education, high sexual activity, and expected cancer treatment costs for infected females in both models. Our results support the conclusions of a 2005 study analyzing the epidemiology of HPV with a potential vaccination that even in the presence of a vaccine, the infective population will remain large due to a high transmission rate. Our results also support the conclusion that a high transmission rate and a high reproductive rate require a high efficacy and high vaccine coverage to eliminate the epidemic.

---

## **Saturday Afternoon**

1:45pm–3:00pm

### ***Bone fragility is increased in mice lacking the ribosomal protein L29***

Laura G. Sloofman, Xiaozhou Zhou, Liyun Wang, and Catherine B. Kirn-Safran  
University of Delaware

Ribosomal proteins (RPs) play an important function in the maintenance of a normal protein synthetic rate. Our group generated the first viable mouse mutant model lacking an individual ribosomal protein. In these mutants, the absence of RPL29 resulted in global skeletal growth deficiencies that persist during adulthood. In the current study, we used RPL29 knockout mice to evaluate the importance of high volume protein synthesis on bone structure and rigidity. The bone microstructure of both null and wild type six-month-old femurs was analyzed at mid-shaft using micro-computed tomography. Considerable differences in the bone geometry were found between the null and control mice. Specifically, using a Student t-test we found that RPL29-deficient femoral diaphyses exhibit a significant decrease in cortical area (11 % and 19% for males and females, respectively;  $p < 0.05$ ) and polar moment of inertia (39% and 32% for males and females, respectively,  $p < 0.05$ ) when compared to control bones. Interestingly, partial preservation of cortical thickness was observed in null males only. Three-point bending tests were performed on the femurs to calculate the bending rigidity (EI), ultimate force, stiffness, and Young's elastic modulus. Our results show that RPL29 deficiency is associated with increased bone material properties associated with increased bone brittleness. These results are likely due to the decrease of bone quality in null mutant bones. In conclusion, these studies provided genetic evidence that the rate of protein synthesis is an important determinant for the establishment of normal bone integrity. This project is supported by the University of Delaware NIH-COBRE P20 RR016458-06.

---

### ***Dispersal mechanisms and processes of spores of the plant pathogenic fungus *Magnaporthe oryzae****

Kyle Stern  
University of Delaware

*Magnaporthe oryzae* is a fungus that causes a devastating disease called 'rice blast'. Each year, the disease destroys enough of the world's rice, barley, and wheat crops to feed more than sixty million people. *M. oryzae*'s premier source of infective inoculum is its asexual conidiospores, borne on stalks called conidiophores. In this project we investigate the dispersal process of *M. oryzae* conidiospores, which is a key part of the epidemiology and survival of the fungus. To date, there is little data on whether this fungus deploys its spores via an active or passive mechanism, either of which can have a significant bearing on control of the spread of this fungus. The project focuses on determining whether *M. oryzae* spores display an active or passive spore dispersal pattern. Spore dispersal is analyzed in a controlled setting, namely a Petri dish placed in a fungal growth chamber. After allowing seven to ten days for the spores to be released from their stalks, the distance the spores traveled was measured using Nikon Elements, a computer software package that analyzes microscope data. In parallel, we have been using measurable parameters, differential equations, and computer simulations to model spore behavior. Data from our experiments suggest that spores utilize an active dispersal process; however the mathematical model suggests that spores are passively dispersed. This difference may be explained simply by the extremely tenuous connections between the spores and their stalks. Any handling of

the spores may cause their release, and therefore, the experimental data may be significantly vulnerable to external variables. Data, models and simulations will be presented and discussed.

---

### ***Phase Response Curves of a Bursting Neuron***

Ikemefuna Agbanusi

New Jersey Institute of Technology

The phase response curve (PRC) is a graph that indicates the sensitivity of an oscillator to the timing of stimulus i.e. which phase of oscillation the stimulus occurs. Thus the PRC can be constructed for any oscillator. Bursting neurons differ from other neuronal oscillators in that they have an active period characterized by fast spiking activity and a silent phase. We adapt a 3-variable model capable of simulating the bursting activity observed in our experimental work on the PD neuron of the Stomatogastric Ganglion (STG) of the crab *Cancer borealis*. This model is also used in our numerical experiments and theoretical considerations. PRCs have been studied extensively for spiking neurons however, the assumptions typically made in utilizing the PRCs do not hold for bursting neurons. This is as a result of the phenomena of spike addition, deletion and burst truncation. Our goal is to understand the mechanisms behind these phenomena. We also develop a way of approximating the PRC by ‘dissecting’ it into the active and silent phases independently. We then compute the approximate burst PRC by using the PRC of spiking activity in the active phase.

---

### **Saturday Afternoon**

3:15pm–4:30pm

#### ***Epidemiology Modeling***

Andrew Davis, Taoufik Youbi, and Joe Skufca

Clarkson University

We explore a number of epidemiological models to study spread of disease on a network. The primary focus of this talk will explore how compartment models of disease transmission can be applied within network context. We will describe some of our research results which have focused upon a susceptible-infected-susceptible (SIS) model for the spread of an infectious disease. This model explores disease propagation in a population of individuals moving along a network, where disease transmission is possible when individuals are collocated on the network. We study this model within the context of lattices as well as small world and scale free networks. We consider the effect of adaptive behavior of the population. Analysis shows how different adaptive behaviors affect the propagation of the disease in the population. We find that even modest adaptations may significantly reduce the impact of the disease on the population.

---

#### ***Tubuloglomerular Feedback-Mediated Dynamics in Three Coupled Nephrons***

Tracy Stepien

University at Buffalo

A model of three coupled nephrons branching from a common cortical radial artery is developed to further understand the effects of equal and unequal coupling on tubuloglomerular feedback. The integral model of Pitman *et al.* (2002), which describes the fluid flow up the thick ascending limb of a single, short-looped nephron of the mammalian kidney, is extended to a system of three nephrons through a model of coupling proposed by Pitman *et al.* (2004). Analysis of the system, verified by numerical results, indicates that stable limit-cycle oscillations emerge for sufficiently large feedback gain magnitude and time delay through a Hopf bifurcation, similar to the single nephron model, yet generally at lower values. Previous work has demonstrated that coupling induces oscillations at lower values of gain, relative to uncoupled nephrons. The current analysis

extends this earlier finding by showing that asymmetric coupling among nephrons further increases the likelihood of the model nephron system being in an oscillatory state.

---

### ***Cellular Automata Models for Population Expansion of *Spartina Alterniflora****

Lee Canning and Mike Dixon

Buffalo State College

Biological invasion and ecosystem engineering has been considered in a number of scientific works recently. Its impact has been felt in water and land areas, and is recognized as one of the top environmental problems, which may influence future economical and social development. Mathematical modeling helps predict and manage its impact. In our work, we use Cellular Automata (CA) techniques and present models, implemented on a Linux cluster, that will help understand the changes in the ecosystem produced by the invasion of *Spartina Alterniflora*, a perennial deciduous grass, which is generally found in intertidal wet lands. The spread of *Spartina alterniflora* has got attention in several regions, including the US and China. *Spartina alterniflora* has a better chance to spread to new areas than native species and once established, it spreads more rapidly than natives. *Spartina alterniflora* is considered to be an ecosystem engineer because of its capability to seize territory from native species, drastically affect sediment dynamics, tidal wave energy, habitat structure for wetland animals and benthic invertebrate populations. CAs are models of physical systems, where space and time are discrete and interactions are local. They have been extensively used as models for complex systems. CAs have also been applied to several physical problems, where local interactions are involved. In spite of the simplicity of their structure, CAs exhibit complex dynamical behavior and can describe many physical systems and processes.

---

**Poster Abstracts** (NOTE: some of the posters are also presentations, and so the abstract is above)

---

### **Analysis and Characterization of the CAG subset in the Short-Chain Oxidoreductase Enzymes**

Tyler Kirsch

Hauptman-Woodward Medical Research Institute

The Short-Chain Oxidoreductase (SCOR) enzymes are characterized by a quasi-conserved forty amino acid fingerprint. We have identified 10,500 SCOR enzymes present in the TGYK sub-family. An NAG (Asparagine, Alanine, Glycine) trimer is found in positions 87-89 in 60% of all SCOR enzymes. We have found 551 SCORs with the Asparagine (N) replaced by a Cysteine (C) creating a CAG subset. This subset was analyzed for the presence of specific residues in specific sequence positions that define cofactor, substrate preference, dimer and tetramer assembly, and species correlation. The sequence specific Aspartate (Asp) and Arginine (Arg) residues predict a ratio of NADP binding to NAD binding of 7 to 3. Analysis of residues in one of three substrate binding loops using a covariance program revealed that 45% of the CAG SCORs have a QXGQ pattern and 5% have a NXGQ pattern. Biochemical studies have shown that several of the enzymes containing the QXGQ sequence are 3-OH Aceto-Acetyl CoA dehydrogenases and dozens of enzymes containing the NXGQ pattern are Beta-Keto Acetyl carrier protein reductases. Dimer Formation was then studied in this CAG SCOR set. Covariance analyses of residues on the dimer formation face show the patterns of F-N-F and Y-N-F and W-N-F in a ratio of 88% to 7% to 2%. All members of the CAG QGQ subset were found to contain a cofactor determining D for NAD binding and an F-N-F pattern suggesting a well-defined dimer of specific stereochemistry. Previous studies of other subgroups of the SCOR family have shown that the last four residues after the carbonyl terminal contribute to the details of tetramer formation. Examination of the carbonyl terminal of the QXGQ subset revealed a highly conserved pattern of DGA[IL]RMstop. Further

analysis of the lambda ( $\lambda$ ) between the conserved TGXXXGXG pattern and the D Residues that recognize NAD in the CAG set, revealed a correlation between the  $\lambda$  values and species. The shorter sequences are associated with the simplest bacteria and the longest spacing associated with eukaryotic organisms. Another specific sequence of VXNR was detected in the substrate defining loops. This pattern has been identified in 87 members of the CAG set. They all contain a cofactor determinant D, which indicates NAD binding and suggests that all 87 are dehydrogenases of the same, as yet undetermined, substrate. They also contain a dimer forming F-N-F sequence. Altogether, the CAG set was found to have important correlations between the substrate specific residues, the cofactor positions, dimer formation, tetramer assembly, and species distribution.

---

### **Modeling a vector-borne disease in horses**

Caitlin Ryan and Colin Kremer  
SUNY Geneseo

Potomac horse fever (PHF) is a potentially fatal, acute intestinal disease that is observed in horses during the summer months across the United States. PHF is a complex, vector-borne disease that is caused by *Neorickettsia risticii*, an obligate intracellular bacterium harbored by a trematode host, *Acanthatrium oregonense*, which parasitizes freshwater snails. Xiphidiocercariae released from snails then infect caddisfly larvae, which serve as the second host in the life cycle of *A. oregonense*. Horses come in contact with *N. risticii* when they accidentally consume infected, deceased adult caddisflies. We create a stochastic matrix model of infected *A. oregonense* populations in different host stages and conduct an elasticity analysis to determine which segments of the vector pathway might be most susceptible to intervention efforts targeted at preventing PHF. We find that the fluke population in larval caddisflies may be the most influential stage of the *A. oregonense* life cycle for the prevention of PHF.

---

### **COMPARATIVE GENOMIC STUDY OF *STREPTOMYCES* THROUGH MICROARRAY ANALYSIS**

Scott Mollison, and Mark A. Gallo  
Niagara University

The project involves the characterization of various strains of *Streptomyces* by using microarray technology. This will be accomplished by amplifying sequences along the length of the *S. avermitilis* and *S. coelicolor* genomes that flank secondary metabolite and antibiotic resistance genes as well as housekeeping genes, which should serve as an internal positive control as all strains should have these sequences. The genomes of *S. avermitilis* and *S. coelicolor* have been sequenced, therefore this study will allow a comparison of the genomic DNA of other *Streptomyces* that produce important medicinally-important compounds as well as laboratory-isolated strains of *Streptomyces* genera to which there has been no molecular characterization.

Microarrays will be made by spotting the previously described PCR products and hybridizing labeled genomic DNA isolated from strains purchased for the project as well as uncharacterized strains isolated from the environment. The comparative genomic hybridization data output should therefore in theory be all-or-none, in which either the gene is present in the genome being examined or it is not. However there is the possibility of homologous and paralogous sequences that may add complications to the analysis. The purpose of the project is thus twofold. The first is comparison of the phylogeny of *Streptomyces* determined using rRNA sequence data with the data obtained by microarray analysis. Secondly, this knowledge it allows for analysis of previously uncharacterized strains for putative secondary metabolite and antibiotic resistant genes by examining them for the presence or absence of characteristic sequences that flank these genes.

---

## ANTIBIOTIC PROFILING OF *STAPHYLOCOCCUS* ISOLATES FROM WHITETAIL DEER

Sarah A. Giacomini, Richard Myers, and Mark A. Gallo  
Niagara University

*Staphylococcus aureus* is a prevalent species of bacteria found in many settings, both clinical and agricultural. *S. aureus* is a common pathogen that has become an increasingly serious issue due to antibiotic resistance. This study looks at the level of antibiotic resistance in *Staph* obtained from the nasal passages of whitetail deer, a source which should not have been exposed to antibiotics. The goals of this research are to gain further information concerning antibiotic resistance and the genes involved, along with the prevalence and distribution of these genes in the isolates obtained from deer. Chromosomal DNA was successfully isolated from each sample. Pulse-Field Gel Electrophoresis (PFGE) was performed to compare the genomic composition of the different isolates. Antibiotic resistance profiling was also performed on the isolates using several common antibiotics.