PROTOZOAN PREDATION AND O-ANTIGEN DIVERSITY AMONG *SALMONELLA*

by

Hans Karl Wildschutte

B.S., University of Pittsburgh, 1997

M.S., Duquesne University, 2000

Submitted to the Graduate Faculty of

School of Medicine, Molecular Virology and Microbiology in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2006

UNIVERSITY OF PITTSBURGH

School of Medicine, Department of Molecular Virology and Microbiology

This dissertation was presented

by

Hans Wildschutte

It was defended on

August 24, 2006

and approved by

Dr. JoAnne Flynn, Associate Professor, Molecular Virology and Microbiology

Dr. Graham Hatfull, Professor, Biological Sciences

Dr. Timothy Mietzner, Associate Professor, Molecular Virology and Microbiology

Dr. Charles Walsh, Associate Professor, Biological Sciences

Dissertation Advisor: Jeffrey Lawrence, Associate Professor, Biological Sciences

PROTOZOAN PREDATION AND O-ANTIGEN DIVERSITY AMONG *SALMONELLA*

Hans Wildschutte

University of Pittsburgh, 2006

Extensive genetic variability at particular loci is observed among many bacteria because alleles confer higher fitness advantages under certain situations. Extensive diversity is observed at the *Salmonella rfb* locus, encoding enzymes responsible for synthesis of the O-antigen polysaccharide. Historically, diversity at the *rfb* locus was thought to be caused by selective pressures from the immune system and maintained by frequency dependent selection (FDS). This hypothesis works well for pathogens like *Haemophilus influenzae* and *Neisseria meningitis*, which alter their O-antigens during the course of an infection. In contrast, *Salmonella* does not alter its O-antigen. More importantly, *Salmonella* shows host-serovar specificity, whereby strains bearing certain O-antigens cause disease primarily in specific hosts; this is inconsistent with FDS. Alternatively, selective pressure may originate from the host intestinal environment itself, wherein diversifying selection (DS) mediated by protozoan predation allows for the continued maintenance of *rfb* diversity and the survival of *Salmonella*. To test if predation may be a selective pressure influencing O-antigen diversity, amoebae were isolated from separate intestinal environments and shown that these amoebae recognize antigenically diverse *Salmonella* with different efficiencies. More importantly, it was demonstrated that feeding preferences are upheld when *Salmonella* differ only by their O-antigen. Thus, protozoan predation may be the selective pressure influencing O-antigen diversity. For extensive genetic diversity to be maintained by DS, a particular O-antigen should confer a higher fitness in a

certain environment. To test this hypothesis, amoebae were isolated from the intestines of fish, tadpoles, lizards, and turtles and their feeding preferences were determined. As expected, related amoeba from the same host share preferences. Strikingly, unrelated amoebae from the same intestinal environment also had significantly similar feeding preferences, and related amoebae isolated from different environments showed no similarity in prey choice. This demonstrates that amoebae from an environment share feeding preferences. In concert, O-antigen variability may result from selective pressures of predation and subsequently may be maintained by DS whereby a certain O-antigen confers a higher fitness advantage depending on its residing environment. This makes sense of the serovar-host specificity and the clonality of O-antigens among *Salmonella* that were not explained by previous hypotheses.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to extend special thanks to my wife for her support and understanding during the long hours spent both in and outside of the lab. Without the encouragement of your best friend, things are unnecessarily difficult. I also want to thank my family and friends, who are the most important people in my life. I thank my mother for her kindness and for providing me a quiet and friendly environment to write my thesis, and my father for his support during stressful times and the relaxing CA getaways. I would like to acknowledge my mentor Jeffrey Lawrence for all his guidance during my graduate career. If it wasn't for my bike, I may have gone crazy.

Publications

Wildschutte H and Lawrence JG. Differential *Salmonella* survival against communities of intestinal amoebae. Submitted to Microbiology 2006.

Wildschutte H, Wolfe DM, Tamewitz A, Lawrence JG. Protozoan predation, diversifying selection, and the evolution of antigenic diversity in *Salmonella*. PNAS 2004 Jul 20;101(29):10644-9.

Guillot J, Demanche C, Norris K, Wildschutte H, Wanert F, Berthelemy M, Tataine S, Dei-Cas E, Chermette R. Phylogenetic relationships among *Pneumocystis* from Asian macaques inferred from mitochondrial rRNA sequences. Mol Phylogenet Evol. 2004 Jun;31(3):988-96.

Norris KA, Wildschutte H, Franko J, Board KF. Genetic variation at the mitochondrial large-subunit rRNA locus of *Pneumocystis* isolates from simian immunodeficiency virus-infected Rhesus macaques. Clin Diagn Lab Immunol. 2003 Nov;10(6):1037-42.

Dykhuizen, D. E., and L. Green. 1991. Recombination in *Escherichia coli* and the definition of biological species. J Bacteriol **173:**7257-7268.

Rabsch, W., H. L. Andrews, R. A. Kingsley, R. Prager, H. Tschape, L. G. Adams, and A. J. Baumler. 2002. *Salmonella enterica* serotype Typhimurium and its host-adapted variants. Infect Immun **70:**2249-2255.

1.0 SELECTIVE PRESSURES AND DIVERSITY AMONG BACTERIA

1.1 THE EVOLUTION OF NATURAL SELECTION

Charles Darwin was the first to describe the concept of phenotypic traits among individuals and the importance of diversity in relation to natural selection. In his lengthy argument entitled "The Origin of Species by Means of Natural Selection or the Preservation of Favored Races in the Struggle for Life," he says that phenotypic differences naturally occur among populations. He further speculates that that if these observed phenotypes were advantageous and heritable then these traits would lead to speciation if acted upon over long enough periods of time (37). Through these ideas, Darwin proposed natural selection - the process by which individual organisms with favorable traits are more likely to survive and reproduce than those with unfavorable traits. For instance, he hypothesized that natural selection acted upon native finches of the Galapagos Islands resulting in specific advantageous characteristics like a particular beak shape and size due to environmental adaptations. It is these types of advantageous characteristics that Darwin thought contributed to finch speciation. Besides differences observed among animals in nature, Darwin also speculated that domesticated animals like pigeons differed from wild ones due to artificial selection imposed by domestication itself and thus were also on the path to speciation. According to this proposed neo-Darwinism concept, phenotypic diversity in a population should be low because, first, beneficial traits would be passed on to offspring and

every individual would express this trait and, second, disadvantageous traits would be lost thus removing diversity which is detrimental. Although neither concepts of speciation nor origin of phenotypic diversity were understood, Darwin recognized a critical role of within-species variation and laid down the foundation of evolution through natural selection.

The discovery of genetics and molecular biology helped mold neo-Darwinism to explain speciation and the basis for both between-species and within-species diversity. Speciation arises from divergent evolution whereby genes shared among individuals diverge overtime due to different selective pressures eventually leading to reproductive isolation. The effects of reproductive isolation and different selective pressures acting on separate populations results not only in the gain and loss of different genes among species but also effects differences of shared genes between species. Thus, high phenotypic and high genotypic diversity between species is expected. Because high phenotypic diversity was correlated to high genotypic diversity between different species, low within-species phenotypic diversity was thought to reflect low genotypic diversity as proposed by Darwin. Because a gene is beneficial, it sweeps throughout a population; every individual expresses the gene and no diversity should exist. However, it was becoming apparent that differences among proteins and DNA did occur within individuals of a species (106, 112); these observations lead to a reshaping and a more comprehensive understanding of natural selection and conspecific diversity.

1.2 THE NEUTRAL THEORY AND WITHIN SPECIES DIVERSITY

Motoo Kimura proposed the Neutral theory in the late 1960s which further expanded the ideas of neo-Darwinism and within-species diversity. Kimura hypothesized that genetic diversity is

observed among loci within species due to near-neutral genetic mutations, ones which are not advantageous or detrimental (111-113). He admitted that most mutations are lethal and lost from the population but suggested that other mutations result in substitutions which are nucleotide base changes that occur and may or may not change an encoding amino acid. These substitutions are usually either slightly detrimental or beneficial and give rise to slightly different loci, or alleles, that persist in the population. Over time, these alleles may increase and sweep the population or go extinct due to stochastic events of allelic frequencies. In fact, Kimura mathematically demonstrated that alleles randomly increase and decrease in the population over time thus giving rise to small variation at most genes within a species (57, 112, 172). The Neutral theory together with neo-Darwinism provides explanations for alleles observed in populations and furthers our understanding of intra-species diversity.

Soon after Kimura proposed the neutral theory, studies with *Drosophila melanogaster* suggested the maintenance of large genetic differences among individuals which did not completely fit the ideas of natural selection and the Neutral theory (182). Here, one advantageous gene did not sweep the population as predicted by natural selection, and genetic differences appeared to provide fitness advantages so this did not follow the neutral theory. For instance, studies showed that in a population of *D. melanogaster*, a rare male genotype had a higher fitness and thus more likely to mate compared to the common genotype of this species (4, 183, 225). Other studies revealed rare male fitness advantage in different *Drosophila* species suggesting the maintenance of genetic diversity occurs among other species and may even happen in different populations (3, 4, 179, 225, 244, 245). Therefore, extensive genetic diversity, above that predicted by the Neutral theory, was selected for and maintained (183, 184). This did not violate the ideas of natural selection because no one genotype was advantageous all the time. Different

alleles were selected for because they were beneficial in certain instances. Along with natural selection and the Neutral theory, the maintenance of extensive genetic diversity further added to the understanding of genetic diversity within a species.

The phenomenon of maintaining genetic diversity within a species is accepted today and observed in many organisms. More recently, diversity among deer mice was observed at the albumin locus encoding a major plasma protein providing a higher fitness at certain elevations (238). In the prokaryotic bacterium *Haemophilus influenza*, genetic diversity is observed at many loci which encode outer membrane proteins contributing to increase fitness during the course of an infection (232). The maintenance of diversity through selection has been observed in numerous organisms and is thought to play a pivotal role in the adaptation of an organism to its environment (8, 24, 62, 88, 134, 152, 159, 198, 199, 249). Extensive diversity occurs in both eukaryotic and prokaryotic organisms with recombination during mitosis giving rise to most intra-species diversity of the former. In this body of work I will focus on phenotypic and genotypic diversity of prokaryotic organisms. These organisms are haploid and do not undergo mitosis. The diversity in bacteria is generated through other mechanisms which will be discussed in the following chapters.

1.3 THE CAUSE AND MAINTENANCE OF BACTERIAL GENETIC DIVERSITY

Extensive phenotypic diversity among bacteria is usually observed with structures that appear on the outer surface of the cell. These structures include but are not limited to outermembrane proteins, O-antigens, and flagella [Figure 1 (9, 24, 72, 104)]. Because outermembrane structures are the first components to interact with a bacterium's physical surrounding environment, it is

believed that phenotypic diversity with these components provides a means of adaptation to fluctuating environmental conditions. In most instances, bacterial phenotypic diversity reflects genotypic diversity. If within a species one structure did well all the time, then there would be little phenotypic or genotypic differences among individuals as explained by the neutral theory and neo-Darwinism. However, because selective pressures arise from within particular environments which influence bacterial evolution, there should be a link between ecological factors and genetic diversity observed in the genes that encode these products. Tying together both genetic diversity with ecological factors will shed light upon how a bacterium interacts within its habitat and help identify the forces that influencing its evolution.

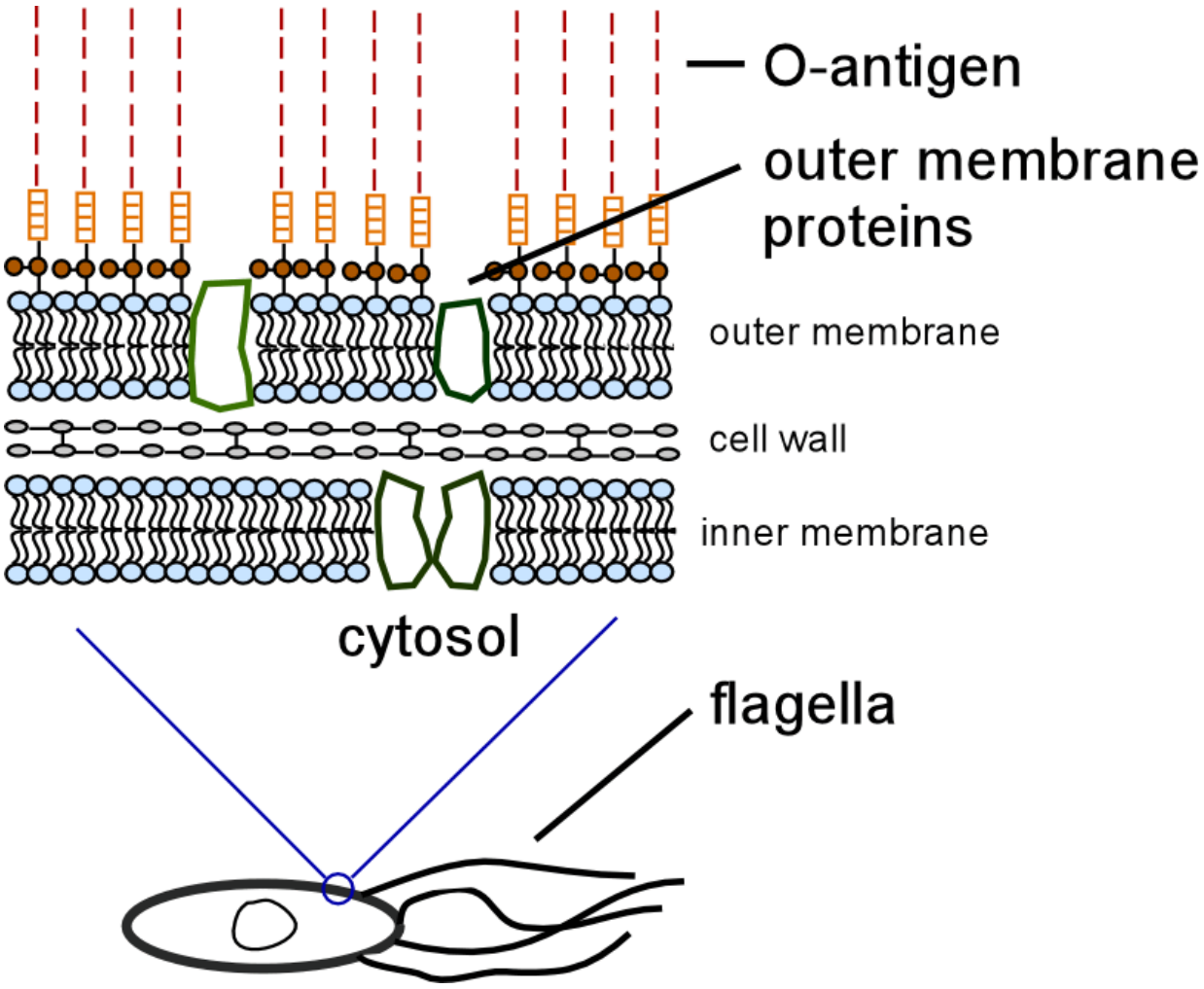

Figure 1. Outer membrane architecture of a Gram-negative bacterium.

Phenotypic diversity is usually observed among structures that appear on the outer surface of a Gram-negative bacterium such as the O-antigen, outer membrane proteins, of flagellar proteins.

1.3.1 The maintenance of bacteria diversity

The maintenance of extensive genetic diversity within a bacterial species is not predicted by the neutral theory instead diversity can be explained by balancing selection. Two models, frequency dependent selection (FDS, Figure 2a) and diversifying selection (DS, Figure 2b), are used to describe this hypervariable maintenance. FDS states that rare alleles confer a higher fitness advantage but as this once rare allele increases in the population, fitness diminishes (7, 129). For example, the immune system has been viewed as the selective pressure causing genetic diversity among pathogens [Figure 1 (23, 122, 132, 196)]. Innate and adaptive systems comprise a harsh changing environment that bacterial invaders must battle to survive. One strategy by which pathogens combat changing environments is by randomly switching outer membrane structures through various genetic mechanisms. This diversity may result in rare structures in the bacterial population and thus avoid immune recognition (29, 163, 210). However, as these structures become more frequent, bacterial fitness diminishes because this once rare structure is now identified and targeted as foreign (Figure 2a). Both bacterial phenotypic and genetic diversity is retained because it confers a higher fitness advantage while interacting with its environment during the time of an infection.

DS is an alternative model which also explains the maintenance of bacterial genetic diversity. With prokaryotes - which are haploid and only contain one copy of a chromosome as opposed to diploid organisms – a single allele confers fitness depending on the environment an organism resides (Figure 2b) and not at the frequency an allele occurs in the population (Figure 2a). For example, the maintenance of diversity with the outer surface protein *ospC* among

Bordetella burgdorferi was proposed to be maintained because certain alleles confer a higher fitness advantage against different types of host immune systems (24). Since *B. burgdorferi* infects a wide host range during its lifecycle, between species host immune systems represent different environments; extensive variability is maintained because no single *ospC* allele does well in all host environments. Both models can explain the maintenance of diversity but the difference between the two is the requisites they define. In some cases one model fits while the other does not.

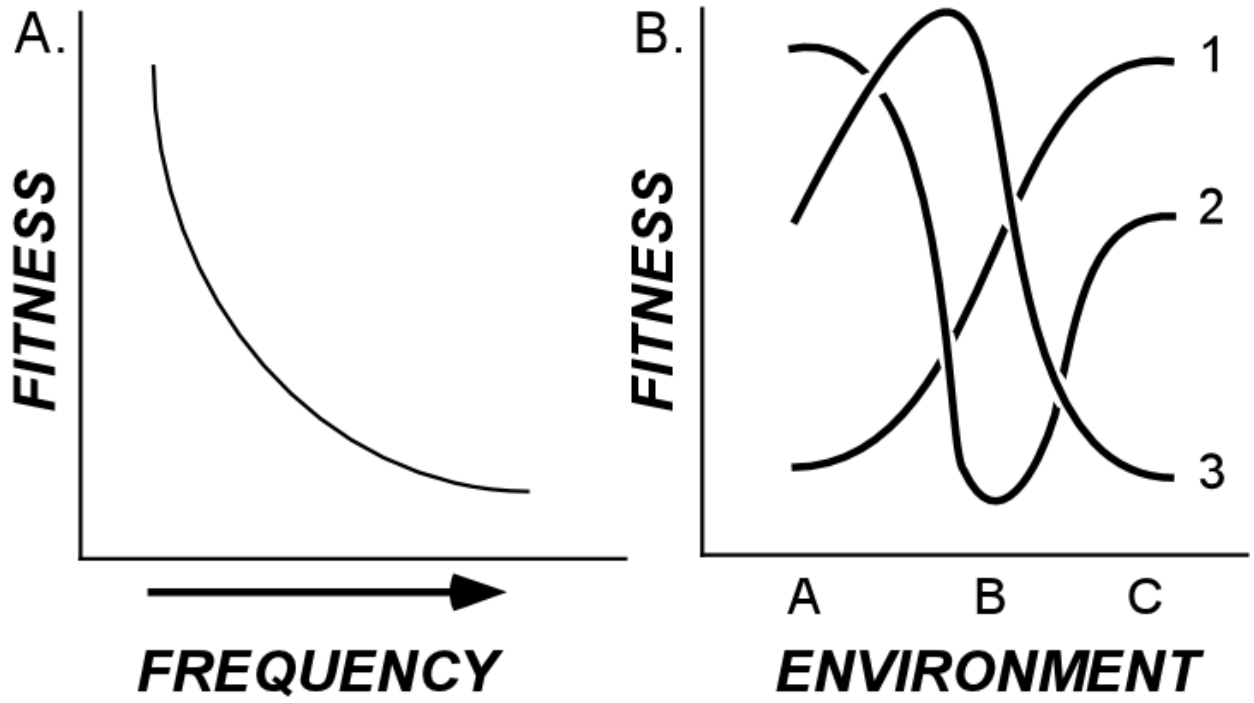

Figure 2. Models used to explain the maintenance of bacterial diversity.

A. Under frequency dependent selection, organismal fitness is highest when alleles are rare. Counter selection of common alleles precludes selective sweeps, maintaining variability at polymorphic loci. **B.** Under diversifying selection, fitness depends on the environment which an organism resides. Here, strains 1, 2, and 3 have different fitnesses in environments A, B, C.

1.3.2 Spatial and temporal regimes

Genetic diversity is influenced by selective pressures that can act either over time or over space. Selective temporal pressures act over a relatively short period of time such as during the course of an infection and influence genetic diversity by imposing rapidly changing environments. For instance, a pathogen must adapt to changing environmental conditions exerted by the immune pressures; if not, the probability of survival is low. In this single host environment, the immune system recognizes cells bearing antigenic profiles which are targeted and killed. However, pathogens like *Neisseria meningitidis* and *Haemophilus influenza* (Chapter 1.4.1) quickly adapt to changing environments by switching their outer membrane structures resulting in higher fitnesses, survival, and bacterial persistence within a single host [Figure 1, (12, 17, 23, 41, 155, 162, 163, 176, 235, 254)]. This selective pressure is defined on a temporal scale because diversity is often generated with every cell division event within one host.

Alternatively, selective pressures can act between environments on a spatial scale. Studies using *Pseudomonas fluorescens* suggest that different environments play a role in generating and maintaining diversity (24, 192, 228). Rainey and Travisano showed that a single *P. fluorescens* strain invaded and adapted to separate environments of standing culture suggesting adaptation to different spatial environments (192, 227, 229, 230). As a result, bacterial diversity is manifested as long term differences meaning that variability is not rapidly generated over a short period of time. When variability does occur by spatial regimes, usually the bacterium is else where as opposed to being in the same host like during an infection. Selective pressures acting on a spatial scale are defined by diversity that occurs between environments

which provides a higher fitness advantage in the new environment. In some cases, such as the maintenance of *rfb* diversity among *Salmonella* which will be described in Chapters 3 and 4, clear distinctions are observed between spatial and temporal modes of pressures which clarify the usage of either FDS or DS to describe the maintenance of genetic diversity.

1.3.3 Predator-prey interactions

Environmental temporal and spatial selective pressures that influence bacterial diversity take the form of predator-prey interactions. These pressures should be great whether the bacterium is a predator or prey. If it is prey, then death usually follows upon capture; if the bacterium is a predator, invading a new habitat or host may depend upon survival. Because predator-prey relations occur in environments where bacteria thrive such as in the water column, soil, and intestines, it is possible that these life threatening interactions have influenced bacterial evolution. I will discuss different forms of predator-prey interactions in which studies have suggested are important selective forces that influence bacterial phenotypic and genetic diversity. (i.) The immune system is a predator which hunts foreign bacterial invaders during infections. Because human individuals with compromised immune systems are succumb to numerous infections which normal hosts fight off, a healthy host immune system is a prominent force protecting individuals from bacterial infections. The generation of extensive outermembrane bacterial diversity is thought to increase pathogen survival by avoidance of immune detection. Thus, the immune system is believed to be a strong selective force causing bacterial diversity (135, 163, 171, 235). (ii) Phages are specialized predators which recognize their bacterial prey through very specific contacts involving phage tail fibers and bacterial outermembrane proteins. Because of this specific interaction, a single amino acid change in an outermembrane structure

can prohibit phage binding and confer bacterial survival. Phages have been shown to play a strong role in bacterial mortality in the water column and other environments, so point mutations observed among bacterial genes encoding outer membranes may be the result of phage predation thus making phages an important selective pressure (18, 85, 105, 239, 260). (iii) Amoebae are general predators that prey upon any bacterium which it can identify as a food source. Because of this amoebae characteristic, more extensive bacterial diversity may be observed such as presenting different outermembrane structures (rather than single nucleotide changes as with the specific phage predators). The removal of amoebae and other protists like ciliates from an environment results in an increase in bacterial counts suggesting this interaction is important force effecting bacterial survival (71, 93, 101, 206, 277).

Predator-prey relations are important interactions which effect the survival of many organisms. One classic study performed in 1942 by Elton and Nickolson showed population cycles of the muskrat and lynx. Fluctuations in either prey or predator numbers was hypothesized to depend on the abundance or the other animal (54, 55); the presence of one effected the other. Likewise, when bacterial predators are removed from an environment, bacterial counts significantly increase suggesting that these interactions are strong forces effecting survival (71, 169, 190). Because bacterial survival depends on prey avoidance, bacteria have evolved different genetic mechanisms in order to survive against predators. I will discuss examples involving the immune system, phage, and amoebae and how bacteria may have evolved to escape these predators. With all examples, the phenotypic diversity and the mechanisms generating genotypic diversity will be discussed along with the spatial or temporal selective regime. By investigating the mechanisms generating diversity and how it is maintained, the predator forces should be revealed.

1.4 SELECTIVE PRESSURES FROM THE IMMUNE SYSTEM

Microbiologists have traditionally viewed bacterial phenotypic diversity as a result of immune selective pressures. Phase variation is a mechanism in which the reversible turning on or off of genes by various genetic mechanisms alters membrane structures [Figure 1, (12, 17, 41, 70, 86, 155, 156, 163, 254, 262)]. These different structures are believed to increase bacterial fitness against immune predation. For instance, if epitope recognition occurs against a pathogen's outer membrane structure such as during an immune encounter, this individual may be counter-selected against. However, there will be a few cells already having generated novel antigenic profiles through phase variation which escape recognition and survive due to phase-variable profiles. This frequent random change ensures that rare types already exist in the population even before environments change (120). It has been proposed that mechanisms such as strand slippage, homologous recombination, or site-specific recombination that mediate phase variation have evolved as a result of selective pressures from the immune system acting over either time or space. These mechanisms allow bacteria to survive the "arms-race" occurring between the pathogen and the predatorial immune system. Bacteria evade the host's immune system by adapting through phase variation and the predator is also constantly changing through hyper-variability of immunoglobulin structures allowing for diverse recognition of foreign epitopes (32, 33, 38). The battle of survival continues with each entity generating diversity to increase its fitness. I will discuss only bacterial diversity generated from this conflict.

1.4.1 Strand slippage in *Neisseria meningitidis* and *Haemophilus influenza*

H. influenza and *N. meningitidis* are commensals of the nasal pharyngeal tract that are most famous for causing bacterial meningitis (23, 235). Upon infection, few individuals invade the blood stream and then are recognized by the host's predatory immune system. These cells will certainly not survive if they retain clonality and express the same outermembrane proteins after each generation. However, bacteria such as *H. influenza* and *N. meningitidis* do adapt and persist during an infection through antigenic variation suggesting immune avoidance (41, 70, 162, 208, 209, 254).

A common mechanism by which *H. influenza* and *N. meningitidis* undergo random reversible on/off epitope switching is through strand slippage which confers a high fitness advantage in a rapidly changing environment (Figure 3). Loci under the control of strand slippage exhibit 1-8 nucleotide repeats in either the promoter or open reading frame (12, 39, 70, 83, 99, 156, 200, 262). During replication, random strand slippage of DNA polymerase leads to the addition or removal of one or more nucleotides in the promoter region (Figure 3a and b) or open reading frame (Figure 3a and c) resulting in either the phase-on or phase-off state depending on the nature of the switch. Promoter slips lead to a gradient response in transcription depending on the ability of polymerase binding efficiency to the changed promoter site. Nucleotide slips in the open reading frame effects translation through frame shifts resulting in a now-in-frame early stop codon. With both promoter and open reading frame slippage, gene expression may be re-established during the next division event when polymerase binding is restored or translation shifts back in frame. Random phase variation allows these organisms to survive within one host during temporal changing environmental conditions especially when the environment cannot be predicted (120).

Strand slippage is observed with many genes found in the genomes of *H. influenza* and *N. meningitidis*. In *H. influenza*, this mechanism controls the expression of *hifB* (252) and *lic1A* and *lic3A* (86, 88, 102, 147, 261) which encode the protein found in fimbriae and the enzymes responsible for modification of lippopolysaccharide (LPS), respectively (Table 1). With these genes, nucleotide repeats are located in the open reading frame causing frame shifts and truncation in translation as depicted in Figure 3c. *N. meningitidis* also has numerous genes controlled by this slippage mechanism that show antigenic variation. For instance, *porA* (9) and *opa* (145, 236) that encode outer membrane proteins, *hpuAB* and *hmbR* which produce hemoglobin receptors (131), and *lgtA*, *lgtC*, and *lgtD* that encode glycosyltransferases for LPS modification by the addition of different sugars (80, 221, 275) are all mediated by strand slippage (Table 1). Numerous other genes in *H. influenza* and *N. meningitidis* are also under this type of phase variation that promotes antigenic variation (12, 83, 131, 236). It is the frequency of this slippage that controls the expression of many genes in which these bacteria assemble different structures allowing the offspring to be different from its parent. Indeed, *in vivo* studies have shown populations of both pathogens – here a population consists of only bacterial cells within one infectious host – undergo phase variation and adapt during an infection (23, 155, 209, 235). Since variability is generated over a short amount of time within one host, phase variation is operating over time to evade the adaptive immune system.

Genetic diversity resulting in epitope switching is maintained because no one structure confers highest fitness during an infectious cycle. This genetic variability is not explained by the neutral theory because diversity is being selected for, it is not neutral. Here, fitness depends on novel structures that go undetected by immune surveillance - the control of *lgtA* with *N. meningitidis* may produce a rare O-antigen conferring a higher fitness compared to other

individuals in the host being attacked by the immune system. Because rare structures are advantageous, the mechanism maintaining diversity is frequency dependent selection (Figure 2a and Table 1). Mechanisms like this show that selection is responding to environmental variation on a temporal scale and organisms that have such mechanisms, be they slippage sites or any other rapid diversity generating mechanisms is similarly responding. The absence of such mechanisms indicates that environmental variability is encountered on a broader scale (which I define as spatial, since the cell need not be in the same location).

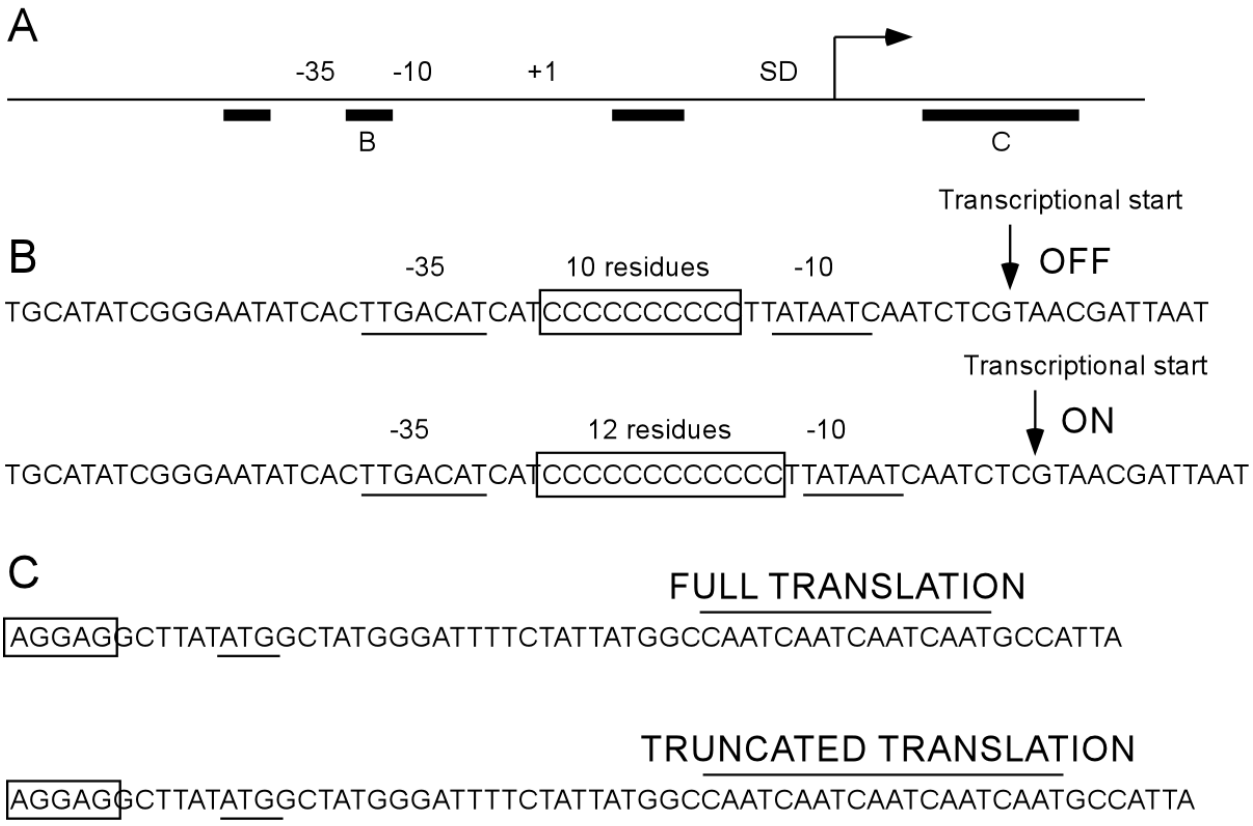

Figure 3. Phase variation controlled by strand slippage.

A. Regions in a gene where strand slippage occurs (thick black lines). B and C refer to respective figures below showing a close up of the specific region. **B.** Strand slippage located between -10 and -35 promoter region. The incorporation of two C residues allows promoter binding and results in the gene on state. Poly-C tracts involving slippage are boxed. **C.** Strand slippage in the open reading frame. An additional CAAT repeat in the bottom sequence leads to a frame shift and truncates translation at a now-in-frame stop codon. CAAT repeats have a line above them. The boxed region contains the Shine-Dalgaro sequence and the start codon is underlined.

Table 1. Polymorphic traits observed among bacteria and phage.

Organism	Gene	Function	Selective pressure	Temporal or Spatial Pressures	Diversity Generating Mechanism	Maintenance of Diversity
Neisseria meningitidis and *gonorrhoeae*	*lgtA*, *lgtC*, and *lgtD*	modify O-antigen by addition of sugars	immune system	temporal	contingency loci	FDS
	pilS and *pilE*	type IV pilin expression	niche expansion	spatial	cassette switching	DS
	opa	outer membrane expression	immune system	temporal	contingency loci	FDS
Neisseria meningitidis	*hpuAB and hmbR*	hemoglobin receptor expression	immune system	temporal	contingency loci	FDS
Haemophilus influenza	*lic1A*	modify LPS by addition of choline	immune system	temporal	contingency loci	FDS
	lic2A	modify LPS by addition of Neu5Ac	immune system	temporal	contingency loci	FDS
	hifA and *hifB*	fimbriae expression	immune system	temporal	contingency loci	FDS
Escherichia coli	*fim* operon	fimbriae expression	immune system	temporal	site specific recombination	FDS
	ompA and *ompC*	outer membrane proteins	phage	temporal	point mutations	FDS
Borrelia burgdorferi	*ompC*	modify outer membrane protein	niche expansion	spatial	point mutations	DS
Phage Mu	*gin*	tail fiber expression	niche expansion	temporal	site specific recombination	FDS
Phage P1	*cin*	tail fiber expression	niche expansion	temporal	site specific recombination	FDS
Bordetella Phage	*mtd*	tail fiber expression	niche expansion	temporal	reverse transcriptase	FDS
Salmonella enterica	*fliC* and *fljB*	flagella expression	immune system	temporal	site specific recombination	FDS

hsd	DNA restriction modification	phage protection	spatial	lateral transfer	FDS
gnd	glycolytic enzyme	-	-	-	-
rfb	O-antigen diversity	niche survival	spatial	lateral transfer	DS

1.4.2 Other phase variation mechanisms

Other phase variation mechanisms such as homologous or site specific recombination have also evolved in bacteria allowing adaptation of these organisms to changing predatorial environments. Since bacterial diversity is generated within one host during the course of an infection, the immune system is the selective pressure. Frequency dependent selection (FDS) maintains diversity because it is important to be rare and go unnoticed by the predator in order to survive. As described with strand slippage, the phase variation mechanisms described here in Chapter 1.4.2 generate diversity over short periods of time in which FDS is acting on a temporal scale. If these phase variable mechanisms are not present, then a broader scale of selection is present.

Genetic variability is observed with type I pili among *E. coli* serotypes. These pili, also referred to as fimbriae, are protein structures extending away from the surface of a bacterium which physically interact with the surrounding environment. These structures are usually recognized as foreign by the immune systems during an infection because they are on the outermembrane of the cell and interact with the environment. *E. coli* may survive an infectious cycle because they have the ability to change their fimbriae structures that assist in immune escape.

E. coli exhibits phase variation at its *fimA* locus controlling the expression of type I fimbriae. Here, the mechanism involving variation differs from that observed with loci using strand slippage (Figure 3). The *fimA* locus undergoes phase variation through site specific recombination (Figure 4). A DNA segment consisting of 296 base pairs and flanked by two 9 base pair inverted repeats includes the promoter for *fimA* (115). Random flipping of this segment turns “on” *fimA* when the promoter is in the correct orientation and “off” when in the other. DNA

inversion depends on factors including integrases such as FimB and FimE that act separately on inverted repeats (115, 151). In addition, other host factors being also affect the frequency of *fimA* inversion (45). Thus, the switching of *fimA* provides a means of diversity among other individuals trying to survive within one host during an infectious cycle.

S. enterica also have evolved site specific recombination but with flagella. Flagella are long proteinacious structures that protrude from the cell surface and play roles in attachment and motility. The flagella filament of *Salmonella* consists of one of two proteins FliC or FljB; phase variation controlling the expression of these genes dictates which one will compose the filament (20). Site specific recombination mediated by Hin recombinase acts on two 26 base pair inverted repeats resulting in flipping of a 966 base pair DNA segment (the H segment) that includes the promoter for *fljBA* and *fliC* transcription [Figure 5 (90)]. When the promoter controls expression of the *fljBA* operon, the FljB flagellar filament protein is expressed along with the FljA repressor that inhibits *fliC* transcription (20, 90, 121). Reverse switching ceases *fljB* transcription and lifts *fliC* repression which is then expressed. As with other structures undergoing phase variation, random switching confers a higher fitness in a rapidly changing environment.

FDS is the model used to explain the maintenance of genetic diversity when rare alleles provide a higher fitness to an individual. In the above examples, loci encoding outer membrane structures provide the ability of switching to rare epitopes during an infection thus giving a bacterium the ability to escape immune surveillance. FDS traditionally has been viewed as the model explaining diversity and only recently has diversifying selection (DS) been viewed as another plausible mechanism explaining diversity among bacteria.

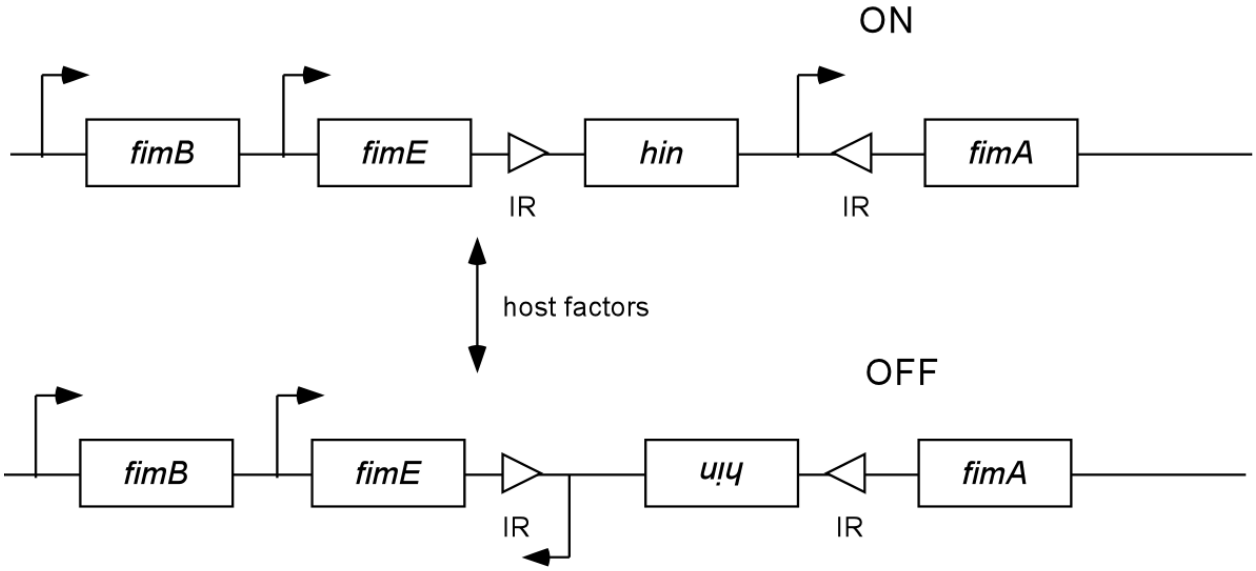

Figure 4. Site specific recombination of *fimA* in *E. coli*.

Site specific recombination occurs with the DNA segment containing the *hin* recombinase and the *fimA* promoter. FimA is turned on if the orientation of the *fimA* promoter drives FimA expression. Recombinases FimB, FimE, and the expression of host factors affect DNA inversion. Triangles represent inverted repeats (IR) on which the recombinases act.

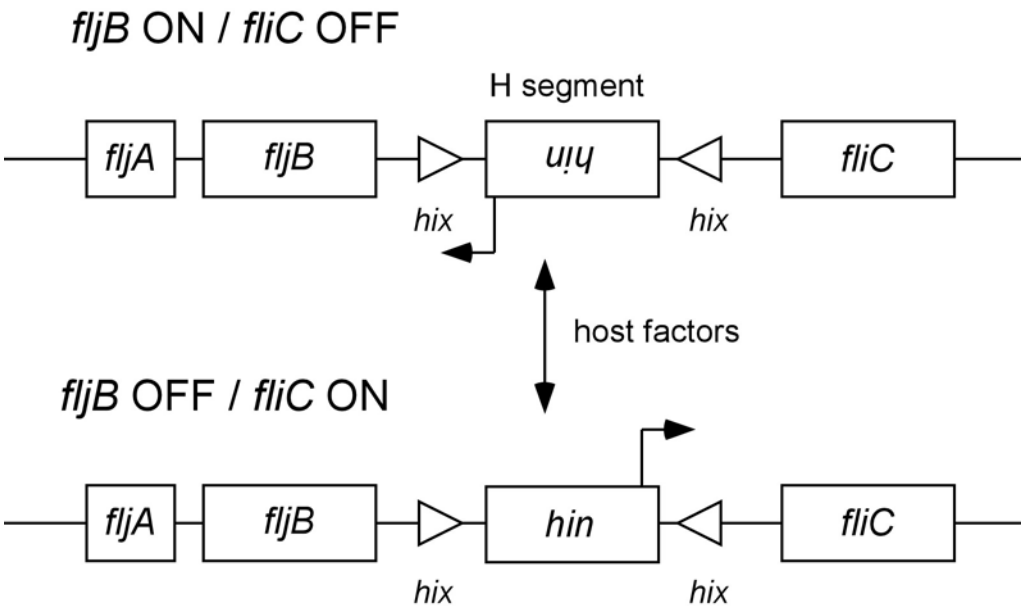

Figure 5. Site specific recombination of *fljB* and *fliC* in *S. enterica*.

The expression of the *fliC* or *fljB* gene depends on the orientation of the H segment which contains a promoter and *hin* recombinase. Orientation of the segment is controlled by host factors which act on *hix* sites. In one orientation, the promoter drives the expression of flagellar protein FljB along with FljA. FljA is the *fliC* repressor which prevents *fliC* transcription. In the other orientation, FljA is not expressed releasing *fliC* repression and the promoter drives FliC expression.

1.4.3 The maintenance of diversity by DS

Diversifying selection is another model that is used to explain the maintenance of genetic diversity. Although FDS is used to explain most bacterial diversity, it does not explain all cases. For instance, immune pressures can act over space as in the case of *ospC* with *B. burgdorferi* in which each individual host represents a different environment. Here, fitness depends on the host a bacterium resides in, and diversity is not generated to confer survival within one host but instead to allow survival within different host environments (Table 1).

Borrelia burgdorferi is a Gram negative bacterium that causes Lyme disease in humans. This pathogen exhibits extensive genetic diversity among genes that encode its outer membrane protein *ospC*. Here, diversity does not occur through phase variation where variability is generated on a short time scale like within one host during an infection. Instead, diversity is thought to have originated from many point mutations which accumulated over long periods of time (77, 240). As a result, the OspC protein is relatively stable compared to structures undergoing antigenic variation that switch often (Figures 3-5). Because this organism does persist in many hosts during its lifecycle including ticks, mice, deer, and humans, stable alleles are thought to confer a higher fitness advantage in different hosts through spatial selective regimes, DS may a plausible model used to explain the maintenance of diversity.

Another example of bacterial diversity being maintained by DS is the type IV pili with *Neisseria meningitidis* and *Neisseria gonorrheae*. Similar to fimbriae produced by *fimA* in *E. coli* (Figure 4 Table 1), this structure is used in attachment. However, diversity with type IV pilin are controlled through Rec-A dependent homologous recombination (72) and not site specific

recombination as with *fimA*. *N. gonorrhoeae* and *N. meningitidis* can express up to six different pilin whose genes encoding these products are located at silent sites throughout the genome denoted *pilS* (217). Expression of *pilS* takes place when recombination occurs at the *pilE* expression site (217, 219); this mechanism is referred to as cassette switching. Cassette switching with *pilE* can also be controlled by the on/off mechanism of phase variation; if recombination is faulty, a frame shift may occur effecting translation and gene expression will be terminated (103, 216, 226).

Homologous recombination involving *pilE* among *Neisseria* may be maintained due to selective pressures acting over spatial scales. Here, increased fitness occurs through re-infection of a non-naïve host via different PilE structures - ones which the immune system was not exposed to during previous infections. Diversity is not generated within one host to increase fitness of the bacterium during the time of an infection; instead, diversity is generated between infectious cycles to increase bacterial fitness through re-infection. Thus, immune pressures act over space. DS is the model used to explain variability because the ability to enter and survive an environment depends on a *pilE* allele and not the frequency it occurs in the population.

1.4.4 The immune system is a selective pressure causing bacterial diversity

Chapters 1.3-1.4 discuss how bacterial diversity is caused through predator-prey selective pressure which is mediated by the immune system. An "arms-race" describes the interaction between these two entities; both are changing to counter the other. Bacteria are the prey both during a short term infection and when re-infecting a non-naïve host in which immune selective pressures act over time or space, respectively. While responding to temporal immune pressure, bacteria rapidly change. In other instances as with *B. burgdorferi*, different hosts resemble

separate environments and selective pressures act spatially – diversity is generated allowing survival between hosts. Because outermembrane diversity is observed with most pathogens, the immune system is believed to be the main selective pressure influencing bacterial diversity. Although FDS has mostly been used to explain the maintenance of bacterial diversity among pathogens, DS has recently been used as an alternative model (Table 1). Besides the immune system, other forces have thought to play an important role in the evolutional of bacterial diversity. These other selective pressures are found in almost all environments where bacteria persist.

Phase variable mechanisms have evolved to allow bacteria to adapt to changing environments, but we can only speculate on the selective pressures causing bacterial diversity. If these rapidly changing environments are where a bacterium spends most of its lifecycle, then it is reasonable to assume selective pressures causing diversity originate from those habitats. However, the lifestyle of organisms like *E. coli* and *S. enterica* is not the life of a professional pathogen which is constantly attacked by the host immune system; instead, these organisms usually reside in the intestines of hosts (65, 67, 76, 173). It is probable that selective pressures which strongly influence diversity originate from a bacterium's ecological niche.

1.5 SELECTIVE PRESSURES FROM PHAGE PREDATION

Phages are another important force believed to influence bacterial diversity (47, 79, 108, 118, 161). These entities are found in most environments where bacteria usually reside (260). In the

water column where both organisms are in close contact, phages have been estimated to be 10^8/ml water (260), and studies have shown that removing bacteriophages results in increased bacterial counts (18, 178). Phage predators must infect a bacterium and use its cell machinery in order to replicate and produce more phage particles. To infect a bacterial cell, both lytic and lysogenic phages must first recognize its host through specific attachment sites involving the phage tail fibers and the bacterial outermembrane structure(s). After a series of events involving phage attachment, entry, DNA amplification, and assembly of new virus particles, the phages are released giving rise to numerous ones that infect the same bacterial species. Because encounters with phages are often lethal to bacteria and present a possible strong selective pressure, bacteria may have evolved mechanisms to resist phage infection. As with the immune system, interaction between the phage predator and its bacterial prey can be viewed as a constant "arms-race" in which both entities are battling each other; the phage tries to infect while the bacterium tries to escape.

1.5.1 Preventing phage attachment

Phages identify bacterium through highly specific contacts involving phage tail fibers and bacterial outer membrane structures. These binding interactions ensure the phage is infecting the proper bacterium whose cell machinery will be used to propagate more virus particles. If phage relied on non-specific binding, infectivity would probably occur in bacteria not supporting phage growth resulting in a dead end to the phage's life cycle. Thus, specific attachment is one way to identify the proper prey to infect.

Phage attachment has been investigated using the well characterized bacterium *E. coli*. Because attachment involves recognition through outermembrane structures, extensive genetic

diversity in genes encoding outer membrane proteins is thought to be due to phage predation. Studies have shown that point mutations in outer cytoplasmic loops of membrane or transporter proteins prevent specific phage binding suggesting increased fitness of prey through diversity (110, 118, 161, 214, 248). For instance, the outermembrane proteins OmpA and OmpC of *E. coli* have been shown to be attachment sites for phage [Table 1, (118, 161, 248)]. Studies introducing singles amino acid changes inhibit phage infectivity. This suggests selective pressures act to avoid phage attachment which results extensive genetic diversity among genes encoding outer membrane loops of these structures - bacteria increase their fitness by preventing phage attachment. While the outer cytoplasmic loops exhibit diversity due to selective pressures from phage predation, the inner protein loops found within the cell are much less diverse presumably due to lack predation (19).

The maintenance of this diversity is explainable by FDS. Mutations in genes encoding outer membrane proteins give rise to different amino acid sequences at phage binding sites preventing attachment and entry. This results in increased bacterial fitness. Selective pressures act over time because bacteria and phage are found in the same environments thus selecting for frequent bacterial change due to their interactions. Other outermembrane structures which phage may contact such as the O-antigen also show diversity, but this diversity may not be caused from phage predation. Here, single amino acid changes do not contribute to the majority of phenotypic diversity observed in O-antigens. Instead, phase variation gives rise to very different structures as observed with *H. influenza* and *N. meningitidis* due to selective pressure form the immune system (Chapter 1.4.1 and Table 1). Other O-antigen diversity like that observed at the *rfb* locus among *Salmonella* is also not proposed to be caused by phages. Diversity at the *rfb* region is the result of horizontal gene transfer which occurs very infrequently and between different

environments (Chapters 2 and 3). This extensive variability is caused by selective pressure acting over space (which I define as diversity that occurs between environments which provides a higher fitness advantage in the new environment). Thus, phages are not the selective force driving O-antigen diversity among bacteria because this diversity is not due to frequent change by temporal selective pressures.

1.5.2 Preventing phage infection

Because phages are a main bacterial predator and found in all environments, bacteria have devised other ways to combat phage in addition point mutations. If a phage can attach and enter a bacterium, the phage can be attacked internally after bacterial invasion. The genes that encode products of DNA restriction modification (RM) activities are intracellular proteins that protect the cell from phage invasion and also exhibit extensive diversity (8, 35, 134). RM systems protect its own DNA by recognizing unmethylated foreign DNA and cleaving it before being expressed and translated or incorporated into their DNA. Among *Salmonella*, type I RM systems are encoded by the *hsd* locus which consists of three genes, *hsdR*, *hsdM*, and *hsdS*. The S subunit alone is responsible for sequence specificity (61), the S and M subunits methylate certain residues (164-166), and R subunit is responsible for restriction endonuclease activity (164, 165). Because the S subunit dictates sequence specificity for both restriction and modification activities it can evolve and retain activity of the other two subunits (164, 165). However, all three genes show a high degree of genetic diversity, and studies have shown that allelic diversity mediated by lateral transfer (not phase variation) has contributed to this polymorphic site (35, 134, 136, 164).

Investing mechanisms that generate diversity and the structures involved should reveal the selective regime. It is interesting to note that the mechanisms generating diversity from phage pressure are very different from the mechanisms that evolved from immune pressure (point mutations compared to phase variation, respectively). Phage selective pressures acting on the *hsd* locus and other loci encoding outer membrane proteins which show point mutations are believed to work over temporal scales from constantly being prey upon by phages (164). Diversity is maintained at both the *hsd* locus and with genes encoding outermembrane proteins because different structures confer a higher fitness when they are rare (Figure 2a) - fitness depends on the frequency of an allele. Novel point mutations and restriction modification systems would be the best arsenal in preventing phage attachment and infection. Thus, bacteria have adapted to different selective pressures which is evident by the mechanisms used to increase their fitness.

1.5.3 Phage are active predators

Thousands of phage progeny produced during a lytic cycle expresses identical tail fibers. Because phage binding is specific, bacterial susceptibility will be problematic when phage infect all hosts or when the bacteria undergo point mutations. As a result, phage fitness will decrease due to superinfection and host availability. In response to this, phages have evolved mechanisms to increase their survival. As described above in Chapter 1.4 with bacteria and the immune system, bacteria and phage interactions can be depicted as an “arms race” in which both are evolving against each other. Phages have evolved strategies to infect bacteria by switching their tail fibers so they could bind to another receptor. Thus, both entities are changing in response to one another.

1.5.3.1 Phages Mu and P1

Phages P1 and Mu both have evolved similar mechanisms to presumably increase host binding range (94, 97). The G and C segments in phage Mu and P1 are recombinant regions of DNA that are mediated by either *min* or *cin* recombinase, respectively, and host factors (94, 96, 109, 187). Both recombinases invert a 3-4 kilobase DNA segment which contains two sets of tail fibers conferring different receptor specificities (79, 97, 212). With P1, one orientation expresses *s* and *u* genes encoding tail fibers for *E. coli* K12 attachment. The other orientation which expresses *s'* and *u'* genes and different tail fibers allows specificity to a different *E. coli* strain [Figure 6, (94-96, 108, 213)]. This same general mechanism is observed with phage Mu.

Random inversion of either the G or C segment is thought to have originated from temporal selective pressures (Table 1). Phage and bacteria inhabit the same environments and are in constant contact with one another, so predator-prey selective pressure is ongoing. If phage infect a small colony of a bacterial strain, thousands of phage particles are produced and chances are that these phage will deplete the host. The ability to change tail fibers through random inversion events increases phage fitness by broadening its host bacterial range. If infected bacteria are surrounded by other similar strains (possibly with slightly different outer membrane loop structures), then phage with different tail fibers can infect these hosts and continue to reproduce. This phage tail fiber diversity is likely maintained through FDS because fitness depends on the ability to be rare and bind to a different host.

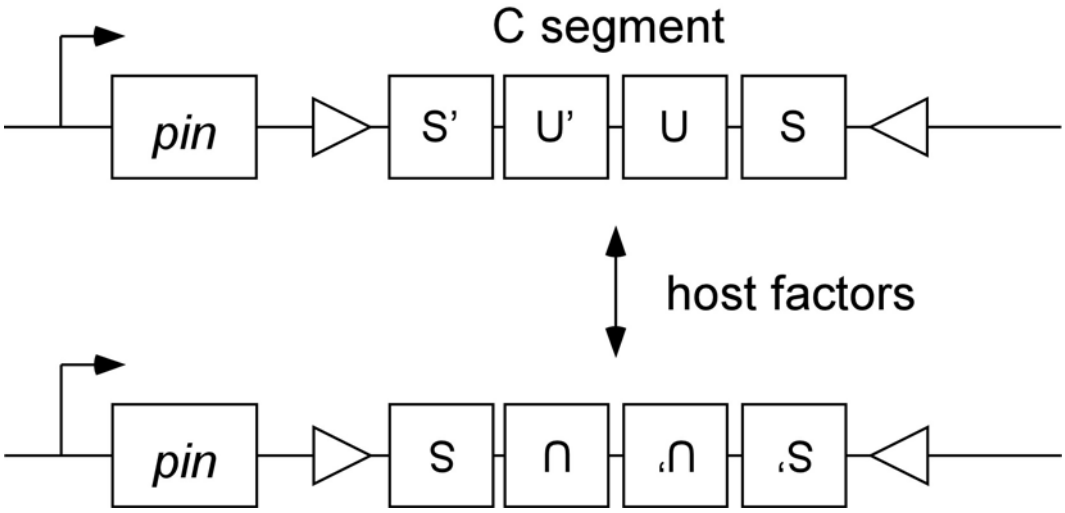

Figure 6. The C segment of bacteriophage P1 generates tail fiber diversity.

Tail fiber diversity allows for phage P1 binding to different strains of *E. coli*. The C segment contains tail fiber genes S and U or S' and U' which encode dissimilar tail fibers. Expression of either set of genes depends on C segment orientation. C segment inversion events are controlled by *pin* recombinase which act on the inverted repeat elements (triangles) flanking the C segment.

1.5.3.2 *Bordetella* phage, a more active predator

The *Bordetella* phage is a more active predator compared to other phages like Mu and P1. *Bordetella* phage hunts its prey (the *Bordetella* bacterium) by generating extensive genetic diversity with a locus that encodes its tail fibers enabling it to recognize diverse outermembrane bacterial structures (89, 152, 187, 212). With every cell burst, *Bordetella* phage produce phage that have different tail fibers. The variability generated by this predator can produce 10^{13} different sequences which allow the phage to attach to its bacterial host through totally different receptors (152).

The *Bordetella* phage infects its host even when the bacterium changes outermembrane structures. *Bordetella* bacterial species do express different outer membrane structures depending on environmental signals relayed by its two component system (246, 247). However, *Bordetella* phage seems to have adapted to this membrane switching by generating and maintaining extensive genetic diversity of their own that allows for the production of different tail fibers, thus the attachment to its host no matter what receptors are present (46, 47, 140). The mechanism creating this extensive genetic and phenotypic diversity within this phage is much different from mechanisms generating bacterial phase variation or phage DNA inversion events (Figure 6). Here, different tail fibers are generated through a template-dependent, reverse transcriptase-mediated process that introduces nucleotide substitutions at defined adenine locations within a particular gene called the major tropism determinant (140). This cassette-based mechanism is capable of providing a vast repertoire of potential tail fiber interactions (152).

Selective pressures causing *Bordetella* and other phage diversity may have originated do to predator-prey interactions. Phage diversity may have evolved so phage progeny infect cells

which are not super infected, have changed outer membrane structures through mutation (as with point mutations with *E. coli*), or due to differential gene expressions (as with *Bordetella*). With these bacterial-phage interactions, the selective pressure acts over time because these interactions are likely always occurring. Since it is beneficial to have a rare tail fiber so a phage can infect a different hosts, diversity is maintained by FDS (Table 1). These interactions are another example of the continuing "arms-race" where predator and prey try to survive against each others fitness strategy.

1.6 SELECTIVE PRESSURES FROM PROTOZOAN PREDATION

Protozoa are single-celled eukaryotic organisms which are found in almost every environment and like phage pose predatorial threats to bacteria (73, 205, 206). Protozoa are most abundant in the water column, soil, and intestinal environments, all places where bacteria survive (205, 266). Studies have shown that protozoan predators significantly affect the size of a bacterial population; if the predators are removed, bacterial population counts increase (36, 73, 85, 93). Some of predators have specific diets and eat certain organisms while others are general scavengers and eat any bacteria. Because protozoan predation is a threat to bacterial survival, selective pressures from these eukaryotic organisms may impact bacterial evolution resulting in the phenotypic and genotypic diversity.

Interactions between protozoa and bacteria exemplify the typical predator-prey interaction involving contact, prey capture, and ingestion. For prey to survive, it must escape either capture or ingestion of the predator. Indeed, studies suggest that bacteria have adapted strategies to survive protozoan predation by both pre- and post-ingestional events. For instance,

bacteria such as *Comamonas acidovorans* and *Ochromonas* sp. form long filaments when subject to increased predation. This increase in size prevents prey ingestion simply because the bacteria are too big (74, 75). Other mechanisms like motility (149) and biofilm formation (73) have also been suggested to have evolved in response to escape protozoan predation while toxin production has been proposed to increase prey fitness after ingestion (148). Thus, bacteria have the ability to escape protozoan predation through different mechanisms. These mechanisms are different from phase variation or point mutations generated from other non-protozoan selective pressures (Chapter 1.4 and 1.5). I have investigated the genetic variability at the *rfb* locus among *Salmonella*; this extensive diversity may be maintained because it allows for predator escape.

1.6.1 The lifestyle of *Salmonella*

Salmonella are considered pathogens because they infect humans and animals, both invertebrate and invertebrate, and cause disease (135). Usually infections occur after ingesting contaminated food or water. Infectious cases result in Salmonellosis characterized by diarrhea, fever, and abdominal cramps 6 to 72 hours after ingestion. If untreated, most individuals will recover between 4 to 7 days. Although most cases of Salmonellosis go unreported, it has been estimated that *S. enterica* is one of the major causes of food borne illnesses in the U.S., causing more than 1.4 million illnesses per year (153). *Salmonella* is viewed as a pathogen because of its ability to cause disease, and it is sometimes difficult to perceive *Salmonella* any other way.

Salmonella is an Gram-negative enteric bacterium which spends most its lifecycle in the intestines of hosts (28, 65, 76). Because *Salmonella* is an intestinal bacterium, it is likely that selective pressures influencing it may originate from ecological forces acting within the gut instead of the pressures from the immune system or from phage (Table 1). However, since

Salmonella is a leading cause for food borne illnesses, the immune system has historically been viewed as the selective pressure causing phenotypic and genotypic diversity. This may be true for loci such as such as *fljB*/*fliC* which are controlled by phase variation enabling random switching during changing environmental states. Interestingly, the *Salmonella rfb* locus (whose products assemble the O-antigen) and the *gnd* gene (which encodes 6-phosphogluconate dehydrogenase) also show extensive genetic diversity but are not under the control of strand slippage or any other phase variable mechanism (Table 1). Thus, it is conceivable that *Salmonella* O-antigen diversity has evolved other selective pressures.

1.6.2 Extensive genetic diversity at the *gnd* locus

Extensive genetic diversity observed with *gnd* and its adjacent locus *rfb* are both different from other examples of diversity discussed so far. The *gnd* locus encodes 6-phosphogluconate dehydrogenase of the pentose phosphate pathway which is a central route for carbohydrate metabolism among enteric bacteria (231). Because *gnd* is an important metabolic enzyme and does not contribute to outer membrane structures, it is expected that periodic selective sweeps would purge diversity at this locus. Other metabolic genes like malate dehydrogenase (*mdh*) (21) and the glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase (*gapA*) (168) reflect this and show little diversity (21). Most metabolic enzymes do not exhibit extensive genetic diversity because different alleles do not contribute to a higher fitness.

The *Salmonella gnd* locus is different from other metabolic enzymes because it exhibits extensive genetic diversity (14, 167, 218). Diversity observed at the *gnd* locus may be due to its tight linkage to the *rfb* locus (158). First, *gnd* diversity is thought to be maintained due to

hitchhiking during a *rfb* lateral transfer events (48, 167, 168). Increased lateral transfer selecting for diversity at the *rfb* operon has resulted in higher frequencies of recombination within the *gnd* locus adding diversity to this gene. Second, selective sweeps are precluded at the *gnd* locus because of the proximity to *rfb* allowing for natural variants to persist. Thus, diversity at the *gnd* locus is thought to be maintained as a result of its tight linkage to *rfb* and not because different alleles are maintained because certain ones provide a higher fitness advantage in particular situations.

Recombination within the *gnd* gene could result in extensive genetic diversity. Dykhuizen and Green compared gene phylogenies of the tryptophan (*trp*) operon, the gene encoding alkaline phosphatase (*phoA*), and the *gnd* locus [Figure 7 and (48)]. If bacterial strains encoding these genes were truly clonal then no differences would be observed among gene phylogenies. However, gene phylogenies and branch lengths differ when using these genes from different strains suggesting intragenic recombination has occurred (Figure 7). Moreover, diversity resulting from recombination is much greater at *gnd* than *trp* or *phoA*. While gene trees between *phoA* and *trp* showed that most strains had a recent common ancestor at a distance of 0.011 nucleotide changes per site, the *gnd* locus does not (Figure 7). Instead, the *gnd* locus shows greater diversity.

The variability at *gnd* is likely due to its linkage to *rfb*. However, there is a possibility that *gnd* diversity is due to balancing selection in which different alleles provide a higher fitness advantage. Since diversity also occurs at genes close to *gnd*, all enzymes would be under balancing selection because either they are all in the same place or because of their linkage to *rfb*. It is more probable that *gnd* extensive genetic diversity is maintained as stable alleles from the result of intragenic recombination due to close proximity to the *rfb* region. The *galF* gene

which flanks the opposite side of *rfb* also exhibits diversity presumably due to its tight linkage to the *rfb* locus. Counter selection of sweeps at the *rfb* locus results in high genetic diversity at adjacent loci such as the *gnd* gene. It is the selective pressure acting on the *rfb* locus and maintaining diversity at this locus which is influences *gnd* diversity.

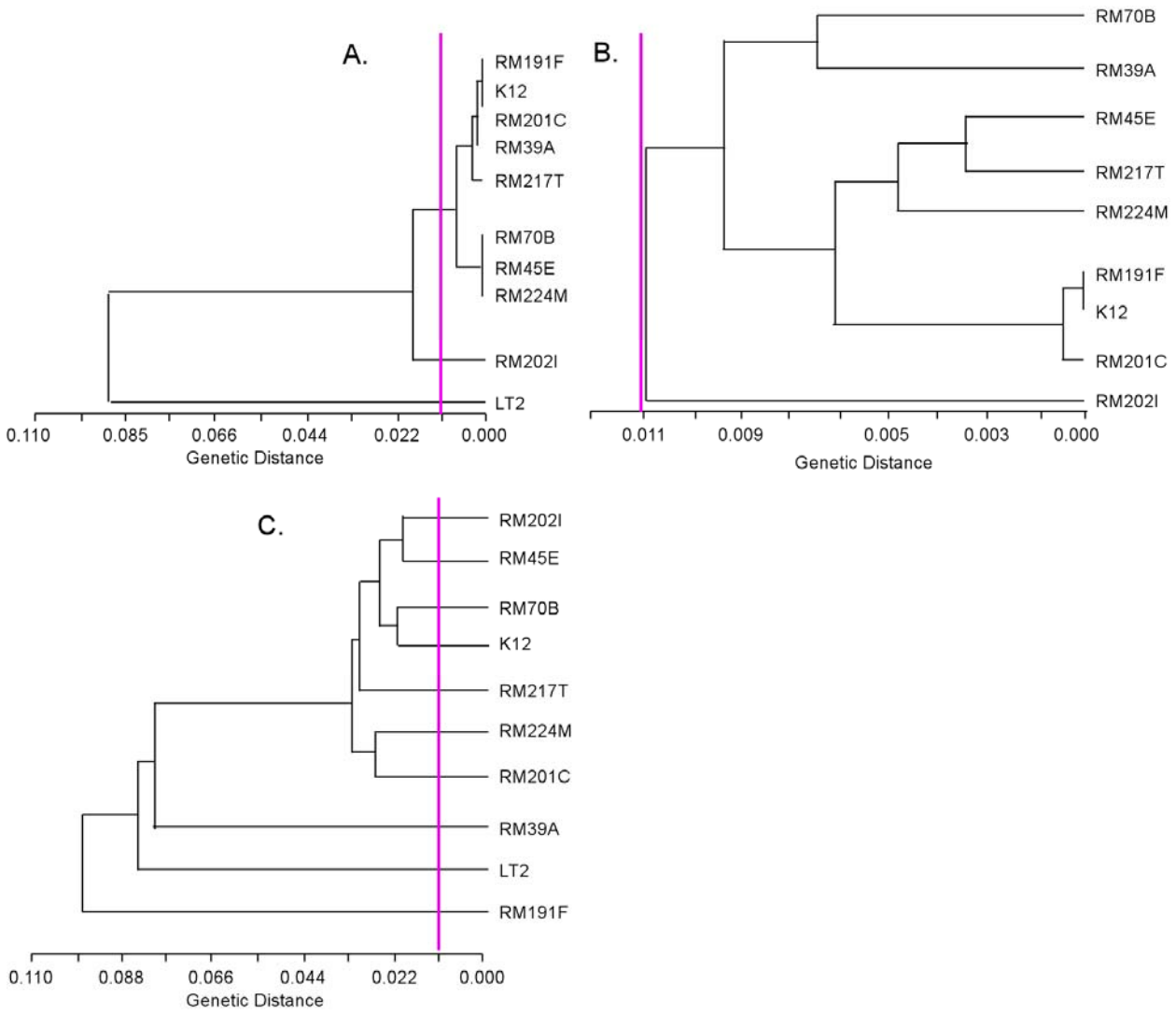

Figure 7. Gene trees of *trp*, *phoA*, and *gnd* loci of *E. coli* and *Salmonella* strains.

Gene trees showing relationships of the *trp* gene among *E. coli* strains and *Salmonella* (A.), the *phoA* gene among *E. coli* (B.), and the *gnd* gene (C.) among *E. coli* strains and *Salmonella* (LT2). Different gene trees suggest recombination has occurred in these organisms. The pink line represents the distance of 0.011 nucleotide changes per site. This figure was modified from reference 48 with permission.

1.6.3 Extensive genetic diversity at the *rfb* diversity

The *rfb* locus, which is proposed to influence the variability at its flanking *gnd* and *galF* genes, also shows extensive genetic diversity among *Salmonella* [Figures 8, 9, (158, 270)]. However, the diversity observed at the *rfb* locus is different from the neighboring genes. Diversity at *rfb* is defined by different genes at this locus instead of alleles. This diversity observed at the *rfb* region gives rise to the more than 70 different O-antigens among *Salmonella* [Figure 1 and (188)]. Because numerous H- and O-antigens (Figures 5 and 9) can be expressed, there are >2000 different *Salmonella* serotypes. A different serovar name is assigned to each type depending on the H- and O-antigen combination. Figure 8 shows the *rfb* region of *S. enterica* serovar Typhimurium along with the individual O-antigen its products construct. The *rfb* locus is 20 kilobases in length and contains 16 genes encoding either sugar synthetase or transferase enzymes (137, 138).

The *rfb* locus among *Salmonella* serovars encodes numerous genes. Some genes encode very different sugar synthesis or transferase products which account for the genetic diversity and the dissimilar O-antigens that are produced among serovars (Figures 8 and 9). This genetic diversity is the result of lateral transfer from other bacteria (26, 139, 270, 271), and not the result of phase variation mechanisms turning these genes on or off. Strand slippage, site specific recombination, or DNA methylation have never been discovered in the *rfb* region thus abrogating any antigenic variation due to these mechanisms. Different genes at the *rfb* locus produce the extensive genetic diversity and not phase variable mechanisms. As a result, different O-antigens are expressed which are stable and do not frequently change.

The hypothesis that selective pressure from the immune system influencing O-antigen (*rfb*) variability among serovars and its maintenance by FDS fails with *Salmonella*. First,

Salmonella serovars do not randomly undergo phase variation to produce different O-antigen structures. Second, serovar-host specificity is observed among *Salmonella* meaning that a host is usually infected a particular serovar that expresses a certain O-antigen: humans are infected by serovar Typhimurium expressing the epitope 1,4,[5],12, chickens are infected by serovar Pollarum expressing the epitope 1,9,12, pigs are infected by the serovar Cholereasuis expressing the epitope 6,7, and horses are infected by the serovar Abortusequis expressing the epitope 4,12 [Figure 10 and (188)]. This is inconsistent with FDS which states that being rare is beneficial (both of these points will be discussed in further detail in Chapters 2 and 3). Because *Salmonella* is a native resident of the gut, it is possible that selective pressures may originate here, in an environment possibly more stable than when infecting a host and facing the immune system. Stable environments would not require phase variation or the necessity to differ from your parent.

In this thesis, I propose that *rfb* diversity is ecologically influenced by protozoan predators whereby a certain O-antigen confers higher fitness in a particular intestinal environment, and because no one allele does well in all environments, diversity is maintained by DS. The experiments performed in Chapters 2 and 3 test my hypothesis. Results may shed light upon the evolution *rfb* diversity among *Salmonella* and offer an alternative hypothesis explaining outermembrane diversity with bacteria.

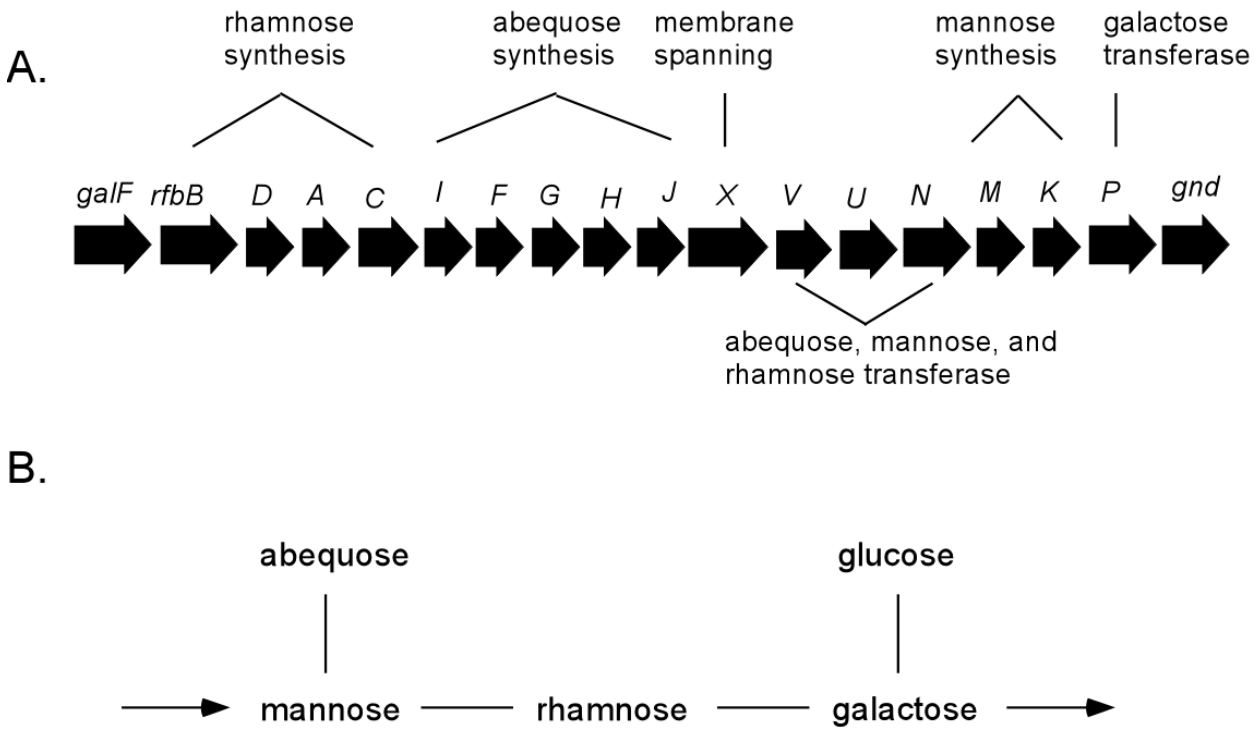

Figure 8. The *rfb* region of *S. enterica* serovar Typhimurium.

A. Genes spanning the *rfb* locus encode sugar synthesis or transfer enzymes responsible for making up the O-antigen. The flanking *galF* and *gnd* (Chapter 1.6.2) genes also exhibit diversity due to tight linkage to *rfb*. **B.** The structure an individual O-antigen sugar unit that makes up the O-antigen of serovar Typhimurium. The left arrow represents the connection to the LPS core and the right arrow represents the placement of addition of repeating sugar units.

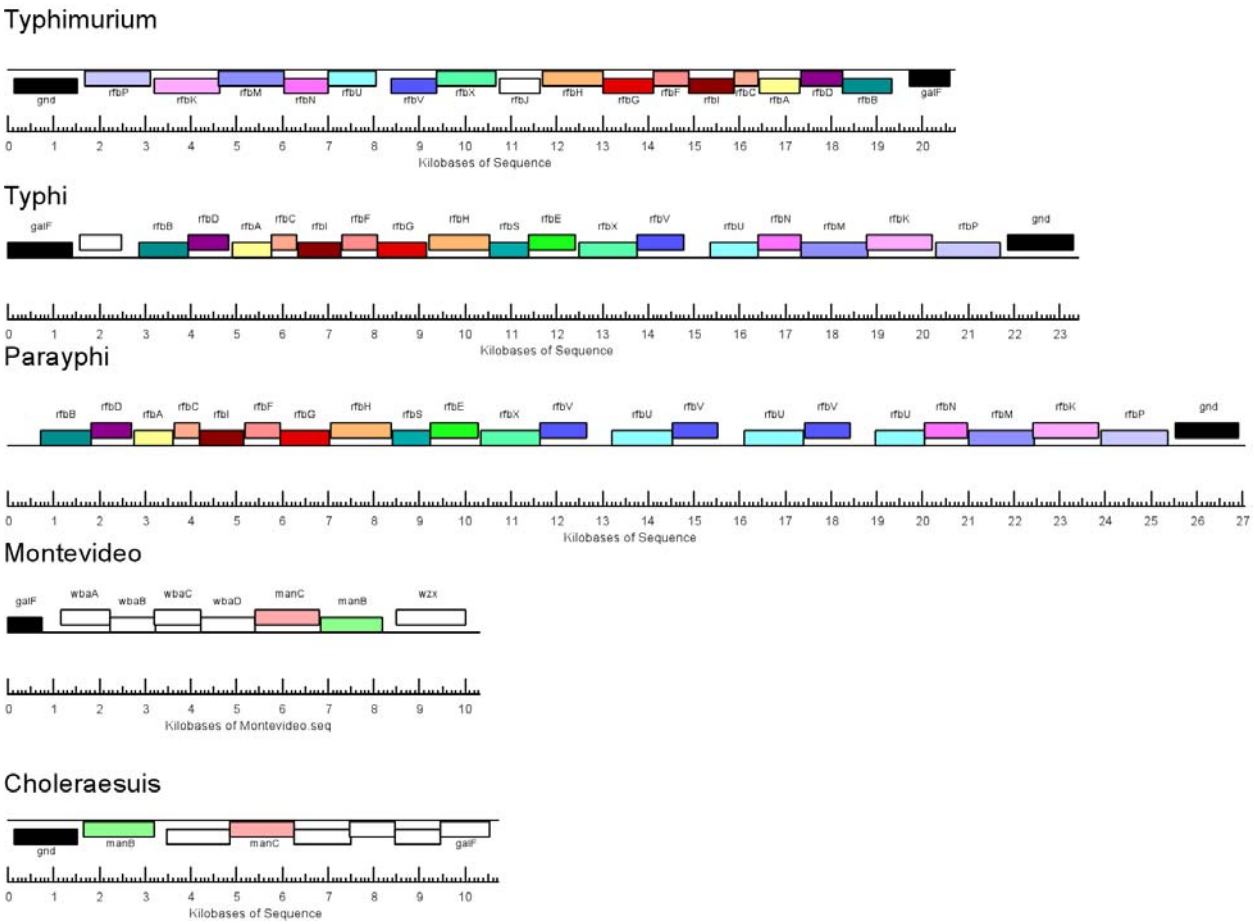

Figure 9. Different *rfb* regions of *S. enterica* serovars.

The *rfb* regions observed among *Salmonella* serovars dramatically differ due to the number and type of genes encoded in this region. This extensive diversity arose by lateral transfer and *rfb* genes are stability maintained at the population level among serovars. Colored boxes represent genes encoding sugar synthetase or transferase genes. Different colors represent different genes involved in O-antigen construction and assembly. Black boxes represents *galF* and *gnd* that flank the *rfb* locus.

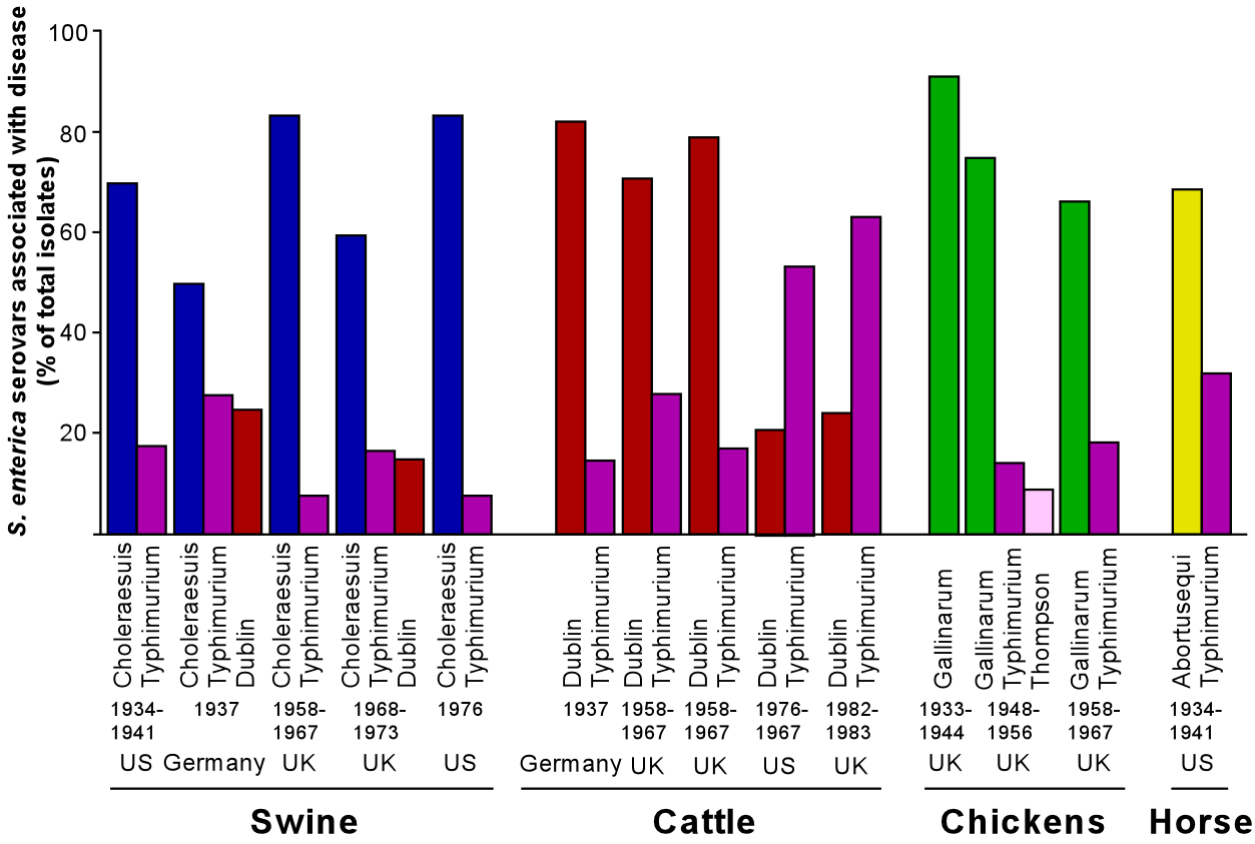

Figure 10. Serovar-host specificity among *Salmonella*.

Salmonella serovars expressing a certain O-antigen usually infect and cause disease among particular hosts (191). The percent of total *Salmonella* serovars isolated from swine, cattle, chickens, and horses through 1934-1983 from different region of the world. Nonrandom distribution of serovar found in certain hosts falsifies the hypothesis that *rfb* diversity is maintained through FDS. This figure was modified from reference 191 with permission.

2.0 PREDATION AND *RFB* DIVERITY AMONG *SALMONELLA*

I propose *rfb* diversity among *Salmonella* originated from protozoan predation and is maintained by DS. Experiments presented here, in Chapter 2, were aimed at falsifying the hypothesis that selective pressures may be due to protozoan predation. For this to be true, amoebae predators must recognize separate O-antigens with different efficiencies.

2.1 O-ANTIGEN DIVERSITY AMONG *SALMONELLA*

Among all *S. enterica* there are at least 70 different O-antigens that are expressed and presented to immune systems (188). The O-antigen is the outermost leaflet of LPS (Figure 1) and decorates the entire outer surface of the cell; it is anchored to the outer membrane-bound lipid A of LPS via the highly conserved core oligosaccharide (5, 63, 78, 82). An O-antigen consists of repeats of two to six linked sugars, and the identity of the monosaccharides in these repeating units varies between different O-antigens. The enzymes responsible for O-antigen synthesis in *Salmonella* and other enteric bacteria are encoded by the *rfb* operon, which exhibits extensive genetic diversity (26, 27, 126, 127, 137-139, 211, 233, 237, 255, 256, 270, 271). Different alleles of *rfb* arose via lateral transfer, whereby genes encoding diverse sugar synthetases and transferases were introduced into the *rfb* operon and directed the synthesis of novel sugars and linkages (137, 138, 211, 256). High genetic diversity at the *rfb* locus is maintained because no one allele, or O-

antigen, confers the highest fitness among serovars; therefore, no allele initiates a selective sweep (7, 129).

Historically, extensive genetic diversity at the *rfb* locus has been attributed to frequency dependent selection (7, 129) imposed by the host immune system (114, 157, 195). Novel *rfb* loci would have an advantage since their cognate O-antigens would be unrecognized by immune systems (Figure 2a); strains carrying rare loci would have higher fitness and would avoid rapid stochastic loss, rising to higher frequency. Yet selective advantages decrease with abundance; as a result, strains with common *rfb* loci cannot dominate the population, or elicit a selective sweep (129), since their fitnesses becomes lower as they become more abundant. In concert, frequency-dependent selection prevents the loss of rare alleles, or the dominance of common alleles, thus maintaining diversity (7).

According to the frequency-dependent selection model, expression of different LPS molecules through gene regulation allows invading bacteria to escape host immunity, survive, and proceed throughout its life cycle; this hypothesis explains O-antigen variation very well for some bacteria. As discussed above, *H. influenza* and *N. meningitidis* are commensal bacteria of the upper respiratory tract that can cause life threatening diseases once they invade their host (23, 25, 162, 176, 250, 274). Upon host entry, *H. influenza* and *N. meningitidis* replicate in the blood stream, resulting in a steadily increasing bacteremia within hours after infection (208, 209, 235, 250). Bacterial survival within the host's blood stream is dependent upon the ability to escape the innate and adaptive immune systems, and *H. influenza* and *N. meningitidis* both have multiple genes under control of phase variation which result in antigenic variants arising every generation, allowing for immune evasion (12, 17, 56, 86, 99, 197, 200, 204, 224). One important example is LPS phase variation via contingency loci which allows these bacteria to express different LPS

molecules after every generation enhancing their ability to survive and escape host immune cell recognition (12, 86, 99). Thus, strong selective pressure from the immune system during bacterial invasion is believed to be the driving force of LPS variation among these bacteria.

Unlike *H. influenza* and *N. meningitidis* (chapter 1.4.1), *Salmonella* is a commensal bacterium of the intestine and does not invade the blood stream when entering the host. Instead, *Salmonella* resides within intestinal epithelial cells or resident macrophages – typically not in the bloodstream – and escapes immense attack from the immune system (87, 133, 171). Although *Salmonella* is exposed to the host immune system while in the intestine through mucosal surveillance, including potential sampling by dendritic cells resulting in IgA release into the intestinal lumen (144, 170), it is not bombarded by the strong host immune pressure experienced by *H. influenza* and *N. meningitidis* during an infection. One could infer that selective pressure driving O-antigen variation among *Salmonella* serovars is not mediated by exposure to the immune system to the degree it may be with *Haemophilus* or *Neisseria*. Thus, it is not surprising that LPS phase variation is absent in *Salmonella* since commensurately strong selective pressure from the immune system is also absent during invasion. However, population-level LPS variability is still observed among *Salmonella* serovars, which display more than 70 different O-antigens (188).

To this end, we believe local frequency-dependent selection fails to explain extensive polymorphism at the *Salmonella rfb* locus, wherein strains present the same O-antigen during infection. Other observations add to our skepticism. First, non-pathogenic bacteria show extensive diversity at their *rfb* loci (8, 81), abrogating extensive, direct exposure to the immune system as a necessity for extreme variability; in addition, loci not encoding surface antigens show high diversity [*e.g.,* the *hsd* locus; see Chapter 1.6.2 and (8, 157)]. Indeed, pathogenic

strains of *E. coli* are limited to few antigenic types – like the enterohaemorrhagic serovar O157:H7 – rather than sharing the breadth of variability at the *rfb* locus seen among natural isolates of *E. coli*. Second, and perhaps more salient, *Salmonella* exhibits host-serovar specificity – that is, certain serotypes infect and cause disease in specific hosts (11, 191) – which is entirely incompatible with, and contradictory to, the FDS model.

Alternatively, excess polymorphism can be maintained by DS, whereby organismal fitness depends upon the environment (Figure 2b). When different alleles confer varying fitness values in dissimilar contexts, selective sweeps are also precluded, resulting in high genetic diversity; this model has been invoked to explain diversity in *Plasmodium* antigens (10) and *E. coli* flagellar antigens (257). Before *Salmonella* can invade their host, they must pass through the harsh environments of stomach acid and bile salts and colonize the intestinal epithelium in competition with more abundant bacterial species. In addition, they must evade generalist predators such as protozoa, which also inhabit intestinal environments (40, 59, 220). Bacterial populations are constrained by the action of protozoan predation, including *Yersinia* in river water (31), *Rhizobium* in groundwater and soil (36, 101), *Xanthomonas* in soil (71), Archaea in the rumen (169), and numerous bacterial species resident in the water column (85, 107, 189, 205, 222, 267, 277) or in soil (6, 71, 101, 205, 206). Since amoebae are abundant predators in vertebrate intestinal tracts (40, 59, 117, 185, 215, 220, 258), they likely act in similar manners to control populations of enteric species.

If protozoan predators from separate environments recognize O-antigens with different efficiencies that is, their receptors have different affinities for the different O-antigen epitopes – they may provide a mechanism by which diversifying selection maintains diversity at the *rfb* locus. Hence, O-antigen variability among *Salmonella* may allow differential serovar persistence

in different host intestinal environments by abating predation in a niche-specific manner. If serovar-host specificity began as an ability to evade host-specific protozoa, the diversifying selection model would provide a new and testable explanation for this pattern of *Salmonella* pathogenicity, and provide a framework for niche differentiation and potential lineage diversification.

2.2 MATERIALS AND METHODS

Line Tests and prey fitness calculations. Strains were streaked on NM solid media (15.5 mM K_xPO_4 pH 7.5, 0.2% peptone, 0.2% glucose, 2.0% agar) from the center of the plate outward then incubated overnight at 37˚C; four replicates of two strains were streaked on each plate, interleaved as depicted in Figure 11a. All 36 pairwise comparisons between nine SARB strains [strains 1, 2, 3, 8, 20, 30, 36, 52, and 59] were performed. A total of 10^4 protozoan cysts (numbers determined via direct counting on a hemocytometer) in 10 μl 0.9 NaCl was added in the middle of the plate on a sterile paper disk; plates were incubated at 34˚C. Plates were photographed every six hours; predation rates were determined from the distance of predation feeding front relative to the line's starting position. Regressions were calculated for distance consumed *vs*. time (R^2 typically > 0.95). The significance of the difference between the two sets of four slopes was determined using a *t*-test. Overall consumption rates were calculated as mean slopes for each plate, which were then averaged across the eight independent pair-wise competition plates bearing that strain. Cell density in lines was estimated by counting cells eluted from 6 cores samples from 6 replicate plates spread with lawns of each strain grown to stationary phase. Cell densities were calculated both by final OD_{600} in liquid NM media with soluble agar

components, or by eluting cells from solid media; comparable results were obtains for both methods. Overall prey fitness values were calculated by multiplying the overall rate of consumption (mm^2/hr) by the normalized cell density (cells/mm^2), normalizing fitness (cells/hr) to the value of the least-preferred strain.

Near-isogenic strain construction. A strain of *Salmonella enterica* serovar Typhimurium was constructed (LD869) containing *hisD*9953::Mu*dJ* and *rfbI*::Tn*10d*Cm mutations. This strain was transduced to histidine prototrophy using either P22 or ES18 bacteriophage lysates (depending on host sensitivity). Transductants were screened for chloramphenicol sensitivity to isolate a strain that mobilized the *rfb* operon. Agglutination tests using antibodies against the 4 and 7 O-antigen epitopes verified that transductants had altered their O-antigen epitope profile. We estimate that ~30 kb of DNA was introduced into the LT2 strain background to create the near-isogenic strains; the *his-rfb* intergenic region does not contain genes for flagella, fimbrae, outer membrane proteins or other potential epitopes, but does include the *wzz* genes for O-antigen chain length determination.

Isolation of protozoa. Amphibians and insects (see legend for Figure 14) were collected from a pond and their intestinal contents were removed via sterile dissection into 0.9% NaCl. Protozoa were separated from bacteriophage and carnivorous bacteria by five rounds of low-speed centrifugation. Cells in pellet were diluted and plated on NM media spread with 10^8 *Salmonella* cells; protozoan cysts were collected from cleared plaques, diluted and reisolated to ensure purity.

Identification of protozoa. The 18S rDNA locus was amplified by the PCR using universal primers; a 1.4 kb band was routinely produced and the sequences of both strands were determined using an ABI-310 sequencer. Strains were given genus designations (14) by virtue of

their close relationship with previously identified protozoa. Sequences have been deposited in GenBank, accession numbers AY576362 – AY576367.

Bacterial survival in the presence of amoebae. An aliquot of 100 µl containing 10^4 cfu's of each of two strains of bacteria were plated on at least 16 NM plates; half of the plates were inoculated with 10^4 cysts of an amoeboid predator at one end of the plate. At the start of the experiment, half of the plates (an equal number with or without predator addition) were immediately eluted with 2 ml of 0.9 % NaCl, diluted and plated on appropriate media, either MacConkey agar with 1.0% xylose (where SARB2 appeared white; other strains appeared red) or Kligler Iron Agar (strains bearing a *phs*-208::Tn*10d*Gn mutation appeared white; other strains appeared black). Indicator plates were incubated overnight at 37˚C, and the numbers of each strain were determined via their colorimetric differences. Experimental plates were incubated at 34˚C until the front of moving amoeboid predators had transversed the plate, ensuring a uniform feeding efficiency for each replicate; the remaining bacterial cells were eluted and counted as above. Significant differences between the proportions of each cell type were compared directly using a *t*-test if $0.3 < p < 0.7$, otherwise data were normalized by standard arcsine($\sqrt{p}$) transformation before analysis. The addition of Tween-80 (used in ciliate competition experiments) to the plates did not alter the results of competition experiments (data not shown).

Bacterial survival in the presence of ciliates. *Tetrahymena pyriformis* was propagated axenically in 2 % Peptone, 0.1% Yeast Extract, 0.2% glucose, and 20 µg/ml kanamycin (to prevent bacterial growth). An aliquot of 100 µl containing 10^4 cfu's of each of two bacterial strains was added to 5 ml TH liquid media (0.5% Peptone, 0.5% Tryptone, 8 mM K_2HPO_4) with 0.02% Tween 80 (to prevent cell clumping) in a 25 ml flask at 30˚C and grown with agitation; four of eight replicates received 10^3 *T. pyriformis* predators. Aliquots were removed after 0, 6,

and 24 hours, diluted in 0.9% NaCl; the numbers of each strain and significant differences in their proportions were determined as above.

2.3 PREDATORS DISCRIMINATE PREY BASED ON THE O-ANTIGEN

2.3.1 *N. gruberi* can distinguish among natural isolates of *Salmonella.*

I examined the abilities of nine *Salmonella* SARB strains (22) to avoid six different amoeboid predators, including 1 laboratory isolate (*Naegleria gruberi*) and five amoebas isolated from intestinal environments. The rate of predation was measured on solid media via sets of pairwise comparisons (Figure 11a). Results showed that a single predator consumes *Salmonella* serovars at different rates (Figure 11b). All pair-wise tests were performed (Table 2) and the data were consistent with a single hierarchical ranking of prey preference for each amoeboid predator (Figure 12 and Table 2).

Rates were corrected for bacterial density (strains with lower growth yields resulted in lines exhibiting faster rates of consumption, R^2=0.32, Figure 13a), although the difference in cell density was small relative to the difference in the rate of line disappearance, suggesting its impact would be low. Neither the width of the bacterial streaks nor the efficiency of prey consumption was found to differ between strains; relative fitnesses were assigned by normalizing corrected consumption rates to that of the least-preferred strain (Figures 12 and 13); correction for variation in cell density only slightly influenced relative fitness values (r^2_s = 0.96; Figure 13).

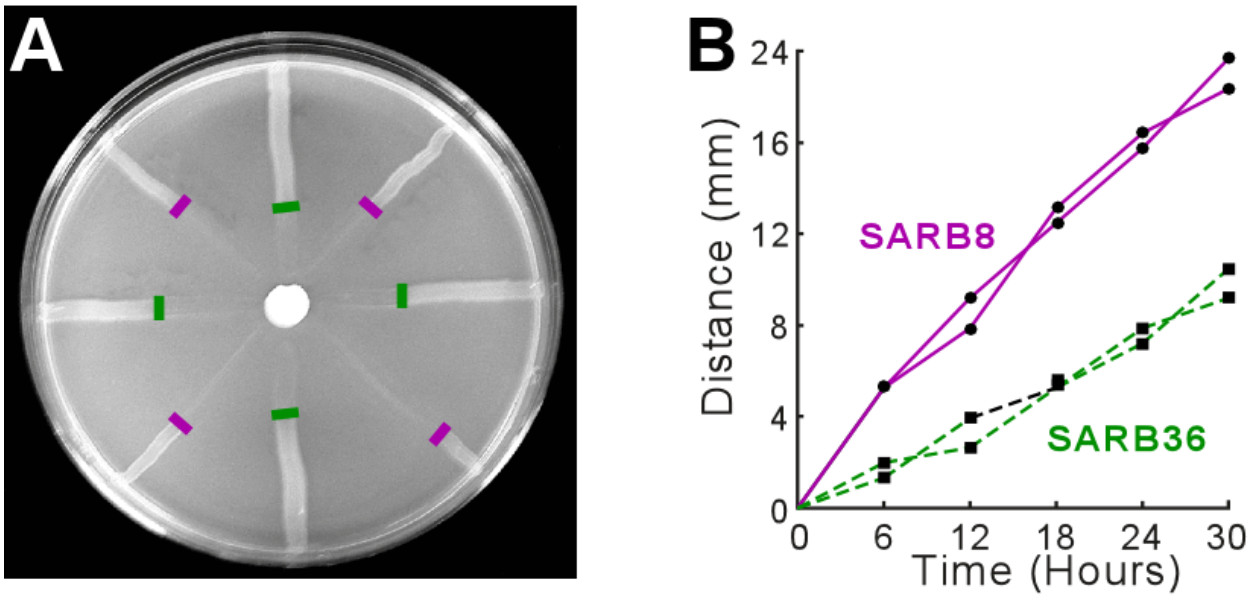

Figure 11. Determining the rates of amoeboid predation by line tests.

A. Pairwise line test between SARB8 (O-antigen 6,7) and SARB36 (O-antigen 6,8) against *Naegleria gruberi*; green and purple lines delineate the feeding fronts for SARB 36 and 8, respectively. **B.** Results of test in panel A; two of four lines are shown for simplicity.

Table 2. Pairwise results of SARB hierarchies.

*Identities of the more slowly consumed strains for each experiment are indicated above the diagonal, with P values noted below; *ns* : not significant.

SARB	01	02	03	08	20	30	36	52	59
01	* -	01	01	08	20	30	*ns*	52	59
02	<0.001	-	02	08	20	30	*ns*	52	59
03	0.027	<0.001	-	08	20	30	*ns*	52	59
08	0.006	<0.001	<0.001	-	20	*ns*	08	52	*ns*
20	0.031	<0.001	<0.001	0.070	-	20	20	20	*ns*
30	0.010	<0.001	<0.001	*0.175*	0.011	-	30	52	59
36	*0.339*	*0.269*	*0.248*	<0.001	<0.001	<0.001	-	52	59
52	0.014	<0.001	<0.001	<0.001	0.001	<0.001	<0.001	-	52
59	<0.001	<0.001	<0.001	*0.450*	*0.142*	0.012	<0.001	0.011	-

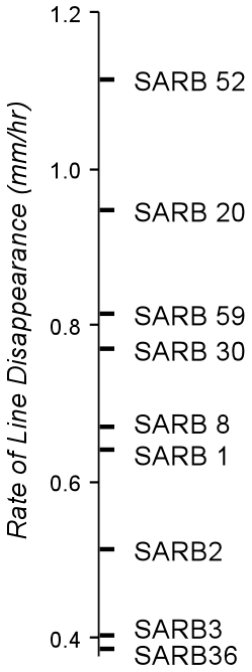

Figure 12. *Naegleria gruberi* rates of predation determined by line tests.

Average rate of consumption of nine SARB strains averaged across all 8-pair-wise comparisons. Figure 12 depicts line tests showing rate of predation with SARB8 and SARB36.

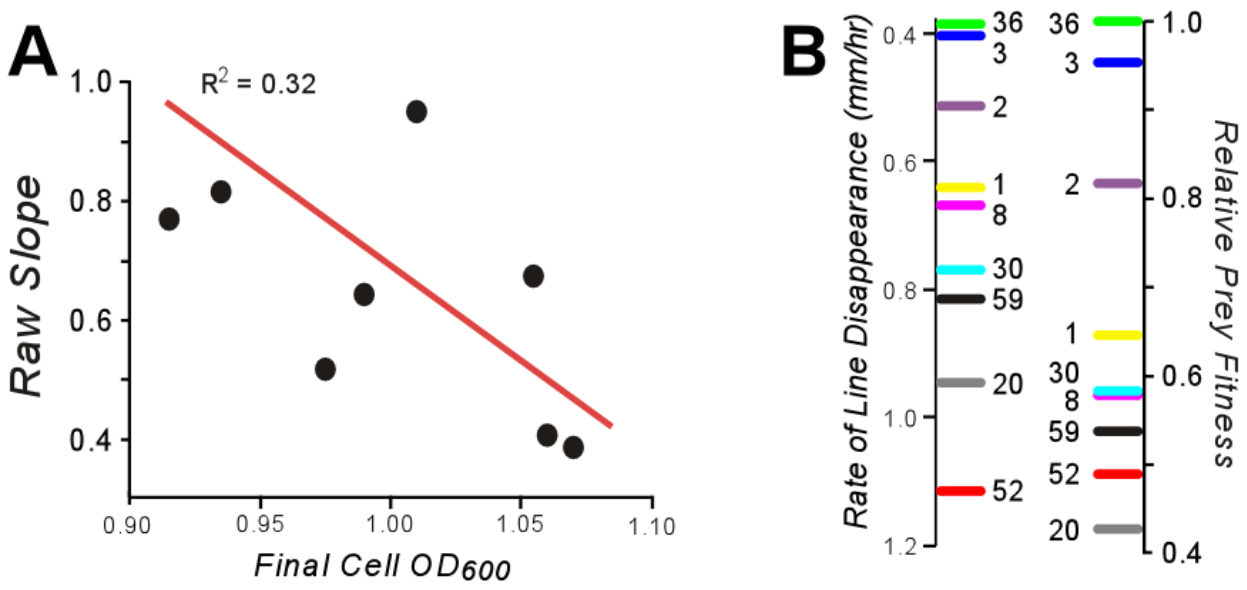

Figure 13. Determining SARB fitness values.

A. Rate of consumption (see Figure 12 and Table 2) is correlated with cell density (calculated as final growth density; see methods). **B**. Comparison of rates of consumption (Figure 14 and Table 2) and fitness for SARB strains facing *N. gruberi* as a predator.

2.3.2 Different protozoa show different feeding preferences.

Serovar fitness depends dramatically upon the predator they face [Figure 14 and (45)]; e.*g.*, SARB52 exhibits a low fitness when faced with *Naegleria gruberi* but higher fitness against *Acanthamoeba*, consistent with the diversifying selection model (Figure 2b). Here, different predators represent the different environments. Since predators have different tolerances to temperature, salinity and pH (data not shown), we expect to find them (and isolated five of them) from disparate intestinal environments; as expected, preliminary results suggest that amoeboid predators isolated from the same host are more uniform than what would expect at random ($P < 0.023$).

Not only do prey have different fitnesses when faced with different predators, the magnitude of the selection coefficients (s) are quite large, on the order of 10^{-1} (fitness = 1-s). Section coefficients on the order of 10^{-4} are readily detected in small-scale chemostat experiments with *Escherichia coli* over the course of 20 generations (142); here, mutation during the course of the chemostat experiment limits the level of detection. Extrapolation of those results to the effective population size of enteric bacterial species suggests that selective coefficients of significantly lower values would have dramatic impacts on the fate of bacterial strains. From this perspective, the two-fold differences in predator susceptibility are enormous by comparison.

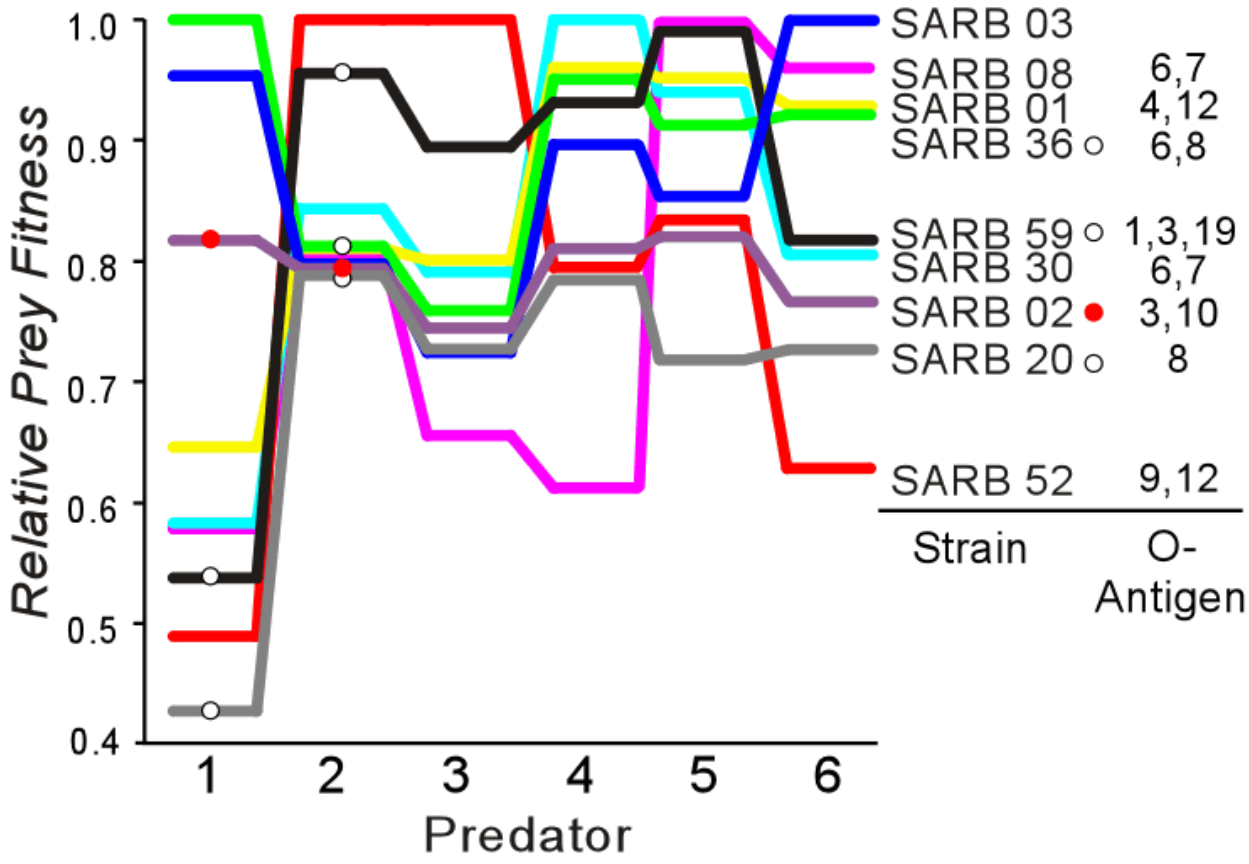

Figure 14. Fitness values determined from predation rates after correction for cell density.

SARB fitness values against different amoeboid predators. Predators (source) are 1, *Naegleria gruberi* (laboratory strain); 2, *Acanthamoeba* sp. (*Hyla crucifer*); 3, *Hartmanella* sp. (*Rana catasbiena)*; 4, *Hartmanella* sp. (*Notophthalmus viridicens)*; 5, *Naegleria* sp. (*Belastoma)*; and 6, *Naegleria* sp. (pond water). Circles indicate strains used in Figure 11; red circles denote SARB2.

2.3.3 Feeding differences reflect predator choice.

Serovars may excrete substances which affect protozoan predators differentially. If so, the putative differential selection coefficients would vanish if a predator were presented with two prey simultaneously, since the excreted substance from either strain would impair the predator. We performed such experiments using subsets of the strains shown in Figure 12. In each case, SARB2 was one of the strains, since it contained a natural inability to consume xylose, allowing discrimination of prey on MacConkey indicator plates. Proportions of prey strains were measured at the onset of the experiment and following predation (which is not 100% efficient; about 1% of cells remain before amoebas encyst due to paucity of prey), and results were compared with predator-free controls (Figure 15).

Differences in the relative abundance of each strain will change over the course of the experiment due to differential growth rates of the two strains. For this reason, the impact of predator must be assessed by comparing the relative abundances of strains in the presence *vs*. absence of predator, not merely from the onset *vs*. the conclusion of the experiment. While no differences in serovar abundance were observed at the start of the experiment (Figure 15, Start), significant differences were observed after predation (Figure 15, Finish), demonstrating that predators can discriminate between prey. More importantly, these results reflected the same protozoan feeding preferences shown by the line tests (strains marked with circles in Figure 14). Taken together, these experiments demonstrate that *Salmonella* fitness values we measured are a function of the feeding preferences of the protozoa in their environment, and not a function of excreted toxins.

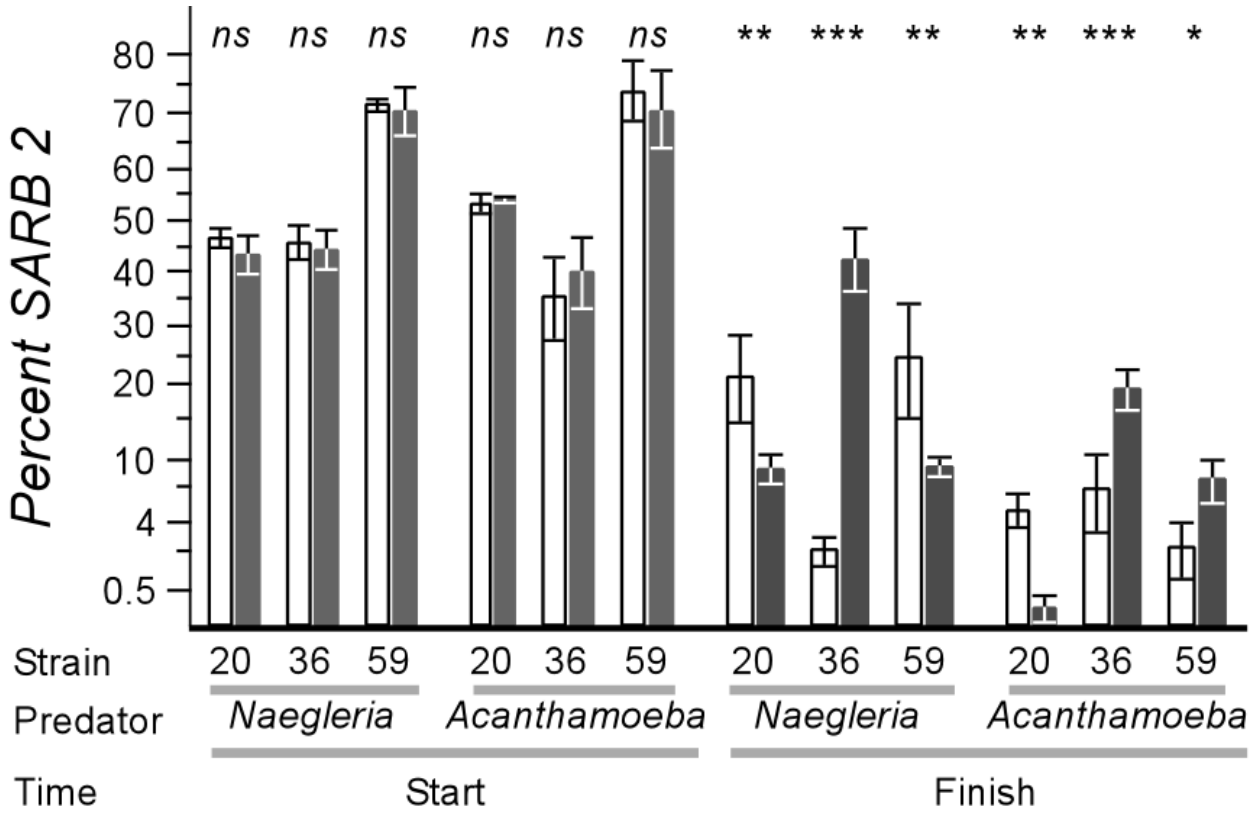

Figure 15. Predator choice among natural isolates of *S. enterica*.

The strain noted was grown with strain SARB2, which fails to utilize xylose; at least 4 replicates were examined. Bars represent the percent of SARB2 present in the population; error bars represent one standard deviation. Open bars report experiments in the presence of predator, whereas filled bars report experiments in the absence of the predator noted. P-values compare the mean percentage of SARB2 between sets of plates with and without predators. Data are plotted along a transformed arcsine (□p) axis, as was used for statistical tests (see METHODS). *ns* denotes $P > 0.05$; * denotes $P < 0.001$; ** denotes $P < 0.0005$; *** denotes $P < 0.0001$. O-antigen designations are as follows: SARB2=3,10; SARB20=8,20; SARB36=6,8; SARB59=1,3,19.

2.3.4 Predators can distinguish prey differing solely at the O-antigen.

For differential predation to occur, predators must recognize some bacterial structure to identify their prey (and to avoid self-predation, engulfment of inorganic matter, *etc*). The abundance of the O-antigen makes it a good candidate for a broad-spectrum ligand recognized by a predator's cognate receptor. There are outer membrane proteins that could act as ligands, such as the flagellum or pili, or outer membrane proteins like BtuB, LamB, OmpA, OmpC, OmpF or PhoE (16, 142, 143). Yet these potential ligands are not all constitutively expressed, present only small loops as binding epitopes, and most would be hidden by the lengthy O-antigen polysaccharide. For these reasons, we postulate that the O-antigen is likely a major ligand for predator recognition; this discrimination would mediate diversifying selection at the cognate *rfb* locus. Other antigens are also likely used, since SARB strains with identical O-antigens did not evade predation equally well (Figure 14).

To test if predation is influenced by the O-antigen, we created strains *of Salmonella enterica* LT2 that vary only in the *rfb* region; strains were tested by antibody agglutination to verify their O-antigen structures. Near isogenic strains were created that encode the *rfb* regions from SARB3, SARB4, or SARB44, designated r03, r04, or r44, respectively. Strains r03 and r44 have similar O-antigens (epitopes 1,4,12 and 1,4,[5],12, respectively) while strain r04 bears a substantially different O-antigen (epitopes 6,7). Experiments with wildtype strains or their respective near isogenic derivatives show that one strain is strongly preferred by the predator when O-antigens differ, but no preference is seen when O-antigens are identical (Figure 16). Moreover, protozoan discrimination of near isogenic strains mirrors the discrimination of cognate wild-type parents. These data indicate not only that the O-antigen influences protozoan predation, but that it may be a primary recognition epitope.

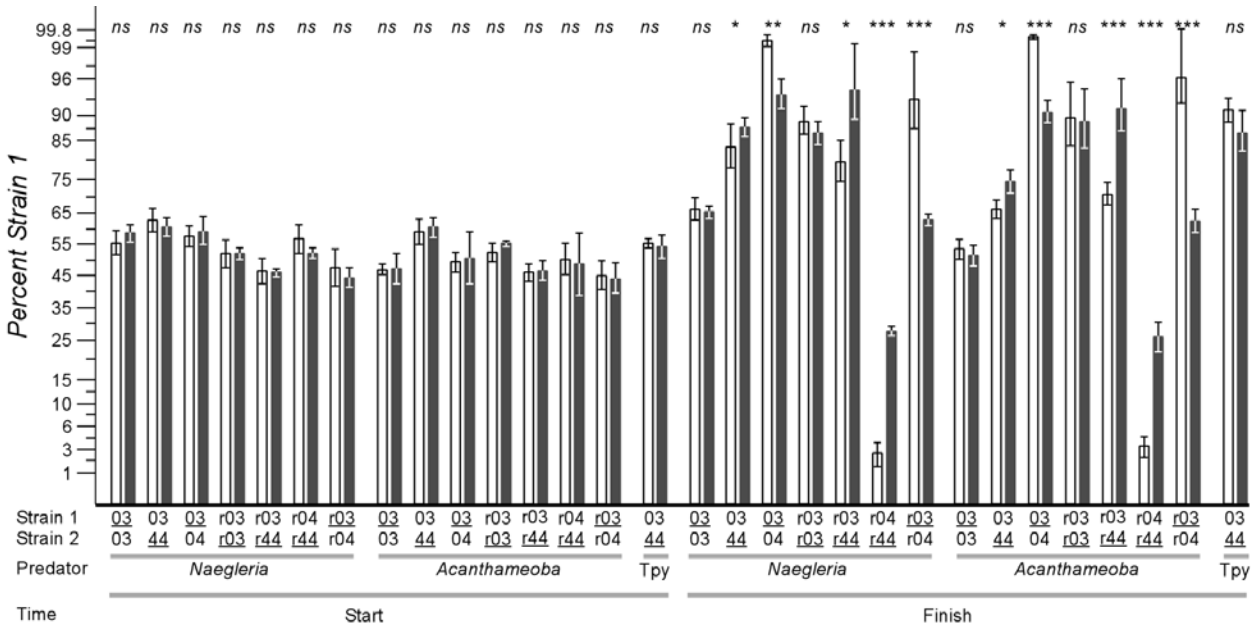

Figure 16. Predator choice among natural isolates of *S. enterica* and isogenic strains.

The strain noted was grown with strain SARB2, which fails to utilize xylose; at least 4 replicates were examined. Bars represent the percent of SARB2 present in the population; error bars represent one standard deviation. Open bars report experiments in the presence of predator, whereas filled bars report experiments in the absence of the predator noted. P-values compare the mean percentage of SARB2 between sets of plates with and without predators. Data are plotted along a transformed arcsine (√p) axis, as was used for statistical tests (see METHODS). *ns* denotes $P > 0.05$; * denotes $P < 0.001$; ** denotes $P < 0.0005$; *** denotes $P < 0.0001$. O-antigen designations are as follows: SARB2=3,10; SARB20=8,20; SARB36=6,8; SARB59=1,3,19.

2.3.5 Predators that do not utilize cell-cell interaction cannot discriminate among prey.

Unlike amoeboid predators of the viscous enteric environment, ciliates filter prey by size; there is no prey recognition through cell-cell contact. We believe serovar recognition by cell-cell contact drives diversifying selection since predators do not demonstrate differential feeding efficiencies (*i.e.,* digestive differences), which would have been detected by the line tests. As expected, we could not detect any feeding preferences in the ciliate *Tetrahymena pyriformis* using strains that have different O-antigen structures which could be distinguished by both amoeboid predators tested (sample data shown in Figure 16). Here, the ciliates grazed upon mixed cultures of *Salmonella* with differing O-antigens, and while the numbers of cells decreased 100-fold over the course of the experiment (the same decrease following amoeboid predator grazing), no measurable preference for one strain over another could be detected.

2.4 DIVERSIFYING SELECTION AND *RFB* DIVERSITY

2.4.1 DS is a viable model for maintaining diversity at the *Salmonella rfb* locus.

Like models used to explain extensive diversity at the *hsd* and *ospC* loci, we believe diversifying selection provides the best framework in which natural selection can act to maintain numerous variant alleles of a gene within a population without rapid alternation among haplotypes. Our data provide an explanation for how and why extensive genetic diversity arose at the *rfb* locus,

and clarify previous unexplained observations: that is, non-pathogenic enteric bacteria have different O-antigens to evade protozoan predators, not the immune system; as a result, we would expect host-serovar specificity.

While avoiding predation may be critical for survival of *Salmonella*, recognition of *Salmonella* is likely of little importance to the predator, since *Salmonella* is not a major constituent of the intestinal flora (~0.1% of cells). Rather, strict anaerobes comprises 95-99% of the microbial intestinal flora (50, 84, 132, 174, 276). Abundant *Bacteroides* expresses numerous different polysaccharides through phase variation (29, 119, 122, 177, 259), perhaps preventing predators from adapting to its O-antigen. In addition, sampling by dendritic cells [resulting in IgA excretion into the intestinal lumen (144, 170, 243)], would have little impact on *Salmonella* population since, as a minor constituent, it would not be sampled as often as strains of *Bacteroides*. Also, preliminary experiments suggest that neither *Naegleria* nor *Acanthamoeba* change feeding preferences, even after 100 generations of consuming non-preferred strains (data not shown). Collectively, these data imply that predators would not change preferences in response to *Salmonella* availability, which could prevent diversifying selection from maintaining *rfb* variability; however, further experimentation is required to determine if predator preferences are indeed stable (Chapter 3).

I believe protozoan predators mediate this selection rather than bacteriophages predators since phages are highly specific in the strains they can infect; moreover, most bacteria acquire resistance to additional phages via their co-immune prophages. Therefore, phages are unlikely to represent a class of niche-specific predators. In contrast, perhaps all of the protozoan predators encountered by a bacterial cell are capable of ingesting it; we have not isolated any amoeba that will not eat any strain of *Salmonella* or *E. coli* as prey. As a result, protozoan predators are likely

to represent the more imminent, niche-specific threats to bacterial survival in intestinal environments than do bacteriophages, although differential distribution of bacteriophages in intestinal environments is largely unexplored.

2.4.2 Differential distributions.

For diversifying selection to provide an explanation for serovar-host specificity, both bacterial prey and their protozoan predators must be stably and differentially distributed between host species. Differential distribution of bacteria among hosts – that is, the nonuniform abundance of different genotypes among different environments – has been convincingly demonstrated for *Salmonella* (30, 191, 251), *E. coli* (30, 64-66, 68), *Enterococcus* (34, 264) and *Bacteroides* (34). Our preliminary data suggest similar results for protozoa (that is, *Naegleria polyphaga* was found preferentially in carnivorous metamorphosing *Rana catasbiena*, whereas *Hartmanella* was found in herbivorous tadpoles ($P < 0.023$; randomization test). We would predict that predators isolated from the same hosts would show similar feeding preferences (see chapter 3).

In addition, differential distribution of protozoa has been described for pathogenic *Entamoeba*, where *E. invadens* causes disease in reptiles (44, 117), including ball pythons (117), whereas *E. histolytica* causes disease in humans (13, 125, 193). *E. suis* and *E. chattoni* infect non-human mammals, yet a related but distinct species preferentially infects ostriches (146). The amoeba *Vannella platypodia* was found to infect multiple fishes (49), while members of the genus *Neoparamoeba* preferentially colonize gills (58). The Microsporidian *Encephalitozoon cuniculi* is a pathogen of domesticated rabbits and dogs, whereas *E. intestinalis*, *E. hellem,* and *E. bieneusi* are opportunistic pathogens of humans (258). Commensal protozoa also show differential distribution among hosts (Chapter 3 and Table 2). For example, the non-pathogenic

amoeba *Paravahlkampfia ustiana* was isolated multiple times from the intestines of skinks (215). These and other studies suggest that protozoa, like bacteria, are not distributed uniformly across all environments.

2.4.3 A road to host-serovar specificity.

Ultimately, this paradigm offers insight into the origins of serovar-host specificity. If amoebae predators that reside in different intestinal environments collectively prefer one serovar over another (Chapter 3), the fitness of *Salmonella* serovars would be host-specific. As a result, serovar Dublin may cause disease in cattle because, perhaps historically, it could better escape predators within cattle, increasing the likelihood of invasion. If this serovar is transmitted to swine, fitness diminishes because its O-antigen would be easily recognized by swine-borne predators, whereas native serovar Choleraesuis avoids these predators. For this to occur, We would predict amoebae from an environment collectively prefer one serovar over another - a serovar is found in an intestinal environment because it can escape native predators (Chapter 3). Thus diversifying selection could lay the groundwork for the acquisition of additional loci that would confer host-specific pathogenicity traits. In this manner, diversifying selection at the *rfb* locus could act as a reproductive isolation mechanism, precluding admixture of these diverging population and allowing for niche specialization to occur. In summary, this work refutes the conventional wisdom regarding how and why diversity at the *rfb* virulence locus evolved, provides a selective mechanism for the maintenance of genetic diversity that may lead to niche differentiation of bacterial populations and subsequent speciation, and offers a sound ecological basis for the origin and maintenance of extensive genetic variation at an important pathogenicity locus.

3.0 AMOEBOID PREDATION MAY MAINTAIN *SALMONELLA RFB* DIVERSITY

In Chapter 2, I showed that amoeboid predation may be a strong selective pressure influencing O-antigen diversity. In Chapter 3, I try to falsify the hypothesis that *rfb* diversity is maintained among *Salmonella* through DS. For DS to be upheld, predators in an environment must collectively prefer one serovar over another. If this is true, then selective pressure could act over a spatial scale between intestinal environments.

3.1 PREDATION IS A STRONG SELECTIVE FORCE

Given the enormous numbers of rapidly reproducing bacteria in nearly every ecosystem, the control of their population sizes and growth rates ranks as one of the most powerful regulators of biomass on the planet. The numbers of viable bacteria – residing in the water column, soil, and subsurface sediments – have been estimated to be $4\text{-}6\text{x}10^{30}$ cells, with turnover times measured in days (266); this constant replacement of bacteria by cell division suggests that bacterial mortality occurs at a high rate. The $\sim 1.2\text{x}10^{29}$ bacteria in all aquatic habitats (266) coexist with $\sim 10^{8}$ bacteriophage/ml (239); these specialized predators recognize their prey primarily through specific motifs on outer membrane proteins or carbohydrates, thus ensuring that phage replication occurs in a compatible host. Similar amounts of bacteria – $2.6\text{x}10^{29}$ (266) – are found in soils, where protozoa are believed to be dominant predators. Unlike phage, which infect

subsets of strains within particular species of bacteria, amoeboid predators are generalist scavengers which prey upon bacteria relatively indiscriminately.

Both amoebae and phage are effective in controlling bacterial populations, whose cell counts increase significantly in the absence of either predator (6, 36, 42, 60, 69, 93, 100, 101, 178, 181, 190, 206, 260, 263, 272, 277). Similar "top-down" predation models govern the dynamics of many eukaryotic populations, *e.g.,* shrimp increase in abundance in the absence of predatory cod (269), barnacles will out-compete mussels only in the absence of starfish predators (154), and populations of terrestrial plants decrease with increasing abundance of arthropods (160, 207). Alternative, "bottom-up" ecological models are appropriate when the abundance of prey controls predator abundance – for example, abundance of the wood mining insect *Phytobia* is dependent upon its host trees (273), and marine phytoplankton blooms are triggered by nutrient load and not low abundance of zooplankton predators (92) – but these models do not appear to apply to most bacteria. Because bacteria increase in numbers in the absence of predators, we infer that predation plays a role in controlling bacterial populations. Therefore, understanding predator-prey dynamics will provide insight into the distribution and abundance of bacteria. Here, we focus on the impact of predator choice on distribution of potential prey among enteric environments.

Bacterial abundance in intestinal environments has been estimated at 10^{11}-10^{12} cells/ml in humans, cattle, sheep and pigs (266). Dominant species of bacteria include members of the Gram-positive taxa *Bacteriodes*, *Clostridium*, and *Lactobacillus*, whereas minor constituents are Gram-negative enteric bacteria such as *Escherichia coli*, *Salmonella enterica*, and *Citrobacter freundii* (50, 132, 174, 276). As minor constituents of the intestinal flora, changes in the Gram-negative members of this community would be unlikely to affect the behavior of the predators,

which feed primarily on Gram-positive bacteria. As such, we can more directly assess how changes in predator behavior may alter the distribution and abundance of Gram-negative prey, decoupling changes in prey populations with changes in predator populations.

Being the most abundant molecule on the surface of bacterial cell, LPS and its outermost structure the O-antigen is a likely structure used by predators to recognize their prey (150, 268); for example, it is the recognition of LPS that elicits an immune response upon infection. As mentioned in Chapter 2, *Salmonella* have more than 70 different O-antigens each which express O-antigens with specific epitopes; for example, serovar Typhimurium expresses the 1,4,[5],12 O-antigen (188). *Escherichia coli* and *Citrobacter freundii* also show great diversity in O-antigen types, and display at least 170 (265) and 45 (116) different O-antigens, respectively.

I hypothesized in Chapter 2 that predation from intestinal amoebae provides selective pressure for maintaining *rfb* genetic diversity among *Salmonella*; that is, a serovar may better escape predators in a particular environment by virtue of the O-antigen it possesses, and different serovars flourish in separate environments with different predators. Supporting this DS model, we showed that intestinal amoebae consumed *Salmonella* serovars at different rates which expressed dissimilar O-antigens (Figure 14). Furthermore, the O-antigen itself is sufficient to elicit a predation feeding preference; predators could discriminate among *Salmonella* that differed solely at their O-antigen (Figure 16).

In chapter 2, I mentioned that serovar-host specificity may be the result of protozoan predation. Serovar-host specificity is the clinical observation in which a serovar expressing a certain O-antigen usually infects and causes disease in a host. For example, when cattle succumb to a *Salmonella* infection it usually carries serovar Dublin, while swine are typically infected by serovar Choleraesuis, horses by Abortusequis, sheep by Abortusovis, chickens by Gallinarum

and Pullorum, and humans by host-restricted serovar Typhi and serovars Typhimurium and Enteriditis [Figure 10 (191)]. Outside of the clinical domain, other studies have suggested that certain *Salmonella* serovars are natural inhabitants of particular hosts (30). In addition, different subgroups of *E. coli* are differentially distributed among the vertebrate hosts (65-68). In Chapter 2 I explained why FDS is not a plausible model used to explain the maintenance of O-antigen (and *rfb*) diversity. FDS is entirely incompatible with the well-established phenomenon of serovar-host specificity (268), whereas our diversifying selection model cleanly explains this otherwise puzzling aspect of bacterial natural history. Because O-antigens are recognized by predators, then *rfb* diversity may be maintained by DS because a single O-antigen does not allow a bacterium to escape predation (that is, confer high fitness) in different intestinal environments inhabited by diverse predators.

A prediction of the DS model is that predators in a particular environment will collectively prefer one serovar over another based on the identity of the prey's O-antigen. This may occur if (a) co-resident amoebae are related and simply share ancestral feeding preferences, or (b) unrelated predators share feeding preferences because the environment influences this phenotype. Alternatively, if amoebae in a particular environment have different feeding profiles, then no single serovar would have an advantage in escaping all predators and *rfb* genetic diversity could not be maintained by this DS model. To discriminate among these alternatives, I isolated numerous amoebae from different environments, identified them based on their 18S rDNA, and tested their feeding preferences to determine if bacterial serovars could escape communities of predators. If so, then *rfb* diversity in these bacteria may have evolved due to selective pressures of amoeboid predation which are maintained through diversifying selection.

3.2 MATERIALS AND METHODS

Media and growth conditions.

The bacterial isolates described below, or strains from the SARB collection (22) – SARB1, 2, 3, 8, 20, 30, 36, 52, 59 – were grown on solid LB medium overnight at 37˚C for routine propagation. NM medium was prepared as 15.5 mM K_xPO_4 pH 7.5, 0.2% peptone, 0.2% glucose, 2.0% agar for propagation of amoebae; amoebae and bacteria were incubated at 34˚C during predation experiments. NM-LG (low glucose) medium was prepared as NM, except with 0.02% glucose and 1.5% agar. SBG sulfa enrichment media, XLT4 media and MacConkey-lactose media were purchased from Difco. LB was prepared using 10 g of tryptone, 5 g yeast extract, 5 g of NaCl, and 12 g agar.

Isolation and identification of intestinal amoebae.

Bullfrog tadpoles (*Rana catesbeiana*) were collected from Geneva Pond #1 in Crawford county Pennsylvania; goldfish (*Carassius auratus auratus*) were purchased from a local pet store, and turtles (*Trachemys scripta*) were purchased from Ward's Scientific Supply House. Lower intestinal contents were removed via sterile dissection into sterile water. Amoeboid cysts were separated from bacteriophage by differential centrifugation. Aliquots of 10-100 µl of the intestinal sample were spread on NM media seeded with 10^8 *Salmonella enterica* serovar Typhimurium LT2 cells as food. Protozoan cysts were collected from cleared plaques, diluted and reisolated to ensure purity. Bearded dragons (*Pogona barbata*) were pets of a colleague; amoeba isolation was performed as above using freshly-collected fecal samples as starting material. Chromosomal DNA was isolated from amoebae using the DNeasy kit from Qiagen. An internal fragment of the 18S rDNA gene was amplified using primers U509F (5'ACTCGAGTGCCAGCAGCCGCGGTAA) and E1789R (5'TCCGCAGGTTCACCTACGGA), and the

nucleotide sequences of both strands of the resulting product were determined using ABI-310 and ABI-3100 sequencers. Strains obtained are listed in Table 1.

Isolation and identification of intestinal bacteria.

Aliquots of 10-200 µl of intestinal samples from turtles, bearded dragons, and goldfish were placed in SBG sulfa enrichment media and incubated overnight at 37˚C. Cultures were then diluted and plated on XLT4 or MacConkey-lactose media. Black or pink colonies, respectively, were picked and streaked for singles on LB. Bacteria from bearded dragons (*Pogona* vitticeps) of a house pet and a screeching owl (*Otus asio*), a resident of the National aviary in Pittsburgh, were isolated from freshly-collected fecal samples. An internal fragment of the bacterial 16S rDNA was amplified using primers rp1 (5'CCCGGGATCCAAGCTTACGGTTACCTTGTTACGACTT) and fd2 (5'CCGAATTCGTCGACAACAGAGTTTGATCATGGCTCAG). A 1.4 kb band was routinely produced and the sequences of both strands were determined using ABI-310 and ABI-3100 sequencers. The strains obtained are listed in Table 1.

Line Tests and fitness calculations.

Procedure for line tests were modified from the protocol of Wildschutte et al. (268). Eight strains were streaked on NM or NM-LG medium from the center of the plate outward; four replicates of two strains were struck on each plate. Plates were incubated overnight at 37˚C until lines were fully grown. A total of 10^4 protozoan cysts (numbers were determined via direct counting on a hemocytometer) in 10 µl of 0.9% NaCl was added in the middle of the plate on a sterile paper disk and plates were incubated at 34˚C. Plates were photographed every six hours; predation rates were determined from the distance of predation feeding front relative to the line's starting position. Regressions were calculated for distance consumed *vs*. time (typically, $R^2 > 0.95$). The significance of the difference between the two sets of four slopes was determined

using a *t*-test. To assay fitness differences, all 36 (9 strains) or 10 (5 strains) pair-wise comparison plates were examined. Overall consumption rates were calculated as mean slopes for the four replicates on each plate, which were then averaged across the independent pair-wise competition plates bearing that strain. Cell density in bacterial lines was estimated as described (268). Overall fitness values were calculated by multiplying the overall rate of consumption (mm^2/hr) by the normalized cell density (cells/mm^2), normalizing corrected consumption rates (cells/hr) to the value of the least-preferred strain to obtain fitness. In addition, the robustness of the fitness hierarchy was validated by consistency of the overall relationships with the results of individual pair-wise competition plates; that is, an overall hierarchy of A>B>C was validated by individual competition plates having yielded A>B, B>C and A>C.

Feeding preference comparisons.

The Pearson correlation coefficient (R) was determined for each pairwise comparison of feeding preferences. For a collection of more than 2 sets of feeding preferences, an average value for R ($R_{Average}$) was determined as simple arithmetic means of the individual pairwise values. To determine if $R_{Average}$ were significantly different from zero, rates of predation were randomly assigned to prey and $R_{Average}$ was computed for these randomized data. The significance of $R_{Average}$ for the observed data was computed as the number of randomized sets of comparisons with a value of $R_{Average}$ that met or exceeded this value. P-values were determined from 1,000,000 randomization trials.

3.3 AMOEBAE FROM AN ENVIRONMENT SHARE FEEDING PREFERENCES

3.3.1 Related amoebae have similar feeding preferences.

Feeding preferences of amoebae represent their "search image" for prey; by definition, amoebae eat bacteria matching this search image more quickly than bacteria which do not. These preferences may be a variable trait among amoebae populations if different members of the same species search for different prey. If feeding preferences change in this way, then otherwise genotypically and ecologically similar amoebae inhabiting the same intestinal environment may not recognize the same prey. As a result, no single O-antigen would allow a *Salmonella* serovar to escape all predators. Alternatively, generalist amoeboid predators could retain the same feeding preferences if altering them provided no benefit to amoeba, regardless of how these preferences affect their prey populations.

To discriminate between these alternatives, I obtained strains of six free living *Acanthamoeba* (FLA), kindly provided by Paul Fuerst. These amoebae were isolated from a marine environment and were >99% identical at their 18S rDNA loci. The feeding preferences of the FLA1 amoeba were determined using nine serotypically-diverse strains of *Salmonella* from the SARB collection. The predator's ability to consume prey was measured using line tests as described below in the methods section and as employed previously (material and methods Chapter 2). Among the nine strains tested, strain SARB52 (serovar Pullorum expressing the 1,9,12 antigen) was consumed the most slowly and was assigned a fitness value of 1.0 (Figure 17a). The remaining strains were reproducibly consumed more quickly, indicating that the FLA1 amoeba can discriminate among *Salmonella* serovars.

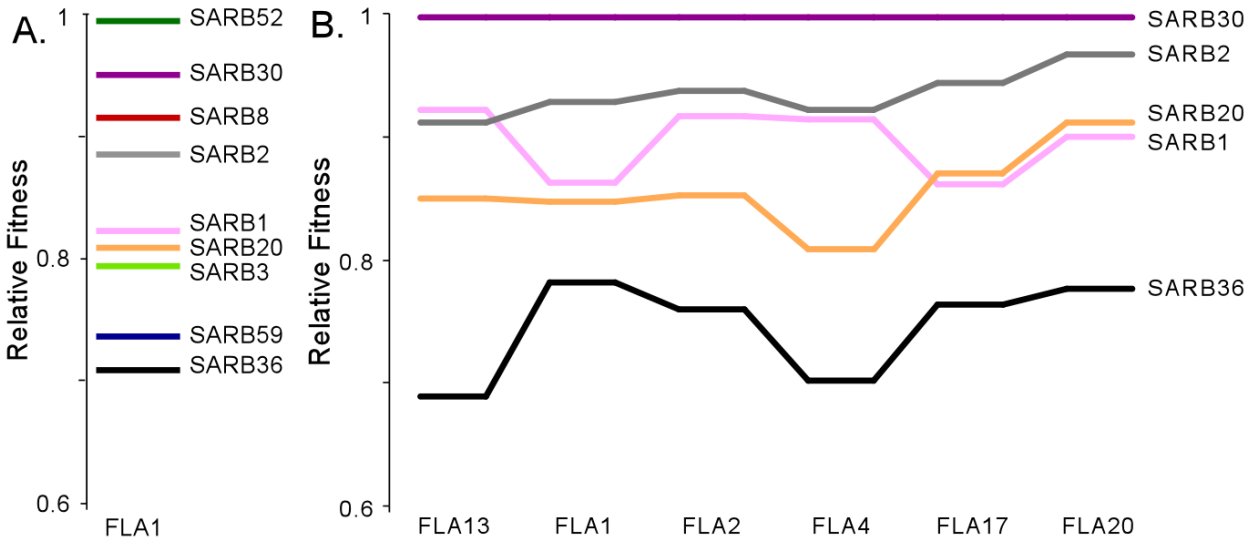

Figure 17. FLA feeding preferences.

A. The feeding preferences of amoeba FLA1 as determined by pairwise line tests. The least preferred SARB strain was assigned a fitness of 1.0. **B**. Feeding references of six FLA amoebae determined as above.

Five SARB strains were selected to encompass the range of fitness values observed with FLA1 and feeding preferences for five other FLA amoebae were determined using these strains (Figure 17b). The FLA1 preference was recapitulated, even for close fitness values, showing that the results of line tests are robust. As a group, all six amoebae show very similar feeding preferences, with strains SARB30 and SARB52 eaten most slowly, therefore being most fit, and strains SARB36 eaten most quickly and being least fit. The average Pearson correlation coefficient for pairwise comparisons among these profiles showed that the preferences of the six FLA *Acanthamoeba* were quite similar, $R_{Average} = 0.972$, $R^2=0.9447$. To determine if this average value of R was significantly greater than zero, I devised a randomization test whereby rates of predation were randomly assigned among bacterial strains, and fitness values calculated; significance was calculated as the number of average values for Pearson's R for random sets of feeding preferences that exceeded the value observed. For this set of 6 predators and 5 prey, the observed similarity of feeding preferences is greater than expected at random (Figure 17b; $P<0.000001$). These results establish that closely related, but not identical, amoebae share feeding preferences. Therefore, I predict that related amoebae residing in the same environment could also share feeding preferences.

3.3.2 Amoebae within tadpoles have similar feeding preferences.

A bacterium may escape a collection of predators in an environment only if those amoebae share feeding preferences. While the results above show that related amoebae may share preference, results in Chapter 2 show that feeding preferences of unrelated amoebae are not shared among all

predators (Figure 14). If prey could escape all predators in an environment, then predators in a single environment must share feeding preferences. To test this hypothesis, I isolated amoebae from the intestinal tracts of three bullfrog tadpoles, *Rana catesbeiana*, at various stages of development. The smallest tadpole was 2.9 cm in length, lacked limbs as was an herbivorous algavore; the largest tadpole was 3.4 cm in length and, given front leg development, this animal was likely transitioning to a carnivorous diet. A total of five strains of *Acanthamoeba* and 2 strains of *Hartmannella* were cultivated and identified based on the sequences of their 18S rDNA genes (Table 3).

To determine feeding profiles, I used the same strategy as discussed above in section 3.3.1. *Acanthamoeba* strain T2-10 was first tested against 9 SARB strains to determine its feeding preferences (Figure 18a); to assay additional predators, five SARB strains were chosen to represent the range of fitness values obtained. Strikingly, all seven amoebae isolated from these three tadpoles shared similar feeding preferences (Figure 18b; R^2=0.9119; $p<0.0002$). These results show that related *Acanthamoeba* from an intestinal environment – here up to 20% different at their rRNA loci –collectively prefer one *Salmonella* serovar over another. More importantly, the feeding preferences of the two *Hartmannella* strains were similar to those of the *Acanthamoeba* isolates, even though these amoebae are distantly related (Figure 19).

These data contrast strongly with the marked differences in feeding preferences for members of these genera isolated from different hosts (Chapter 2 and Figure 14). These data again show that related intestinal amoebae share feeding preferences, but also suggest that feeding profiles may be similar among all amoebae in a particular environment regardless of their relatedness.

Table 3. Amoebae and bacterial environmental isolates.

Environment	Amoebae Cultivated	Bacteria Cultivated	Notes	Figure
Marine	6 *Acanthamoeba*	nd	-	12
Tadpole				13
T2	4 *Acanthamoeba*	nd	Herbivore	
T3	2 *Hartmannella*	nd	Omnivore	
T5	1 *Acanthamoeba*	nd	Carnivore	
Fish				15
F1	6 *Naegleria*	4 *Citrobacter*	Reared at 30˚C	
F3	4 *Tetramitus*	6 *Citrobacter* and 1 *Aeromonas*	23˚C for 3 days	
F5	1 *Tetramitus*	4 *Citrobacter* and 4 *Aeromonas*	23˚C for 5 days	
F8	5 *Tetramitus*	12 *Aeromonas*	23˚C for 8 days	
Bearded Dragon				17
BD1	3 *Tetramitus*	7 *S. enterica* Typhimurium	Juvenile, carnivore	
BD2	1 *Acanthamoeba*	4 *S. enterica* Bahrenfeld and 8 *S. enterica* Typhimurium	Adult, herbivore	
Turtle				17
R1	1 *Acanthamoeba*	-	Herbivore	
R2	2 *Acanthamoeba*	-	Herbivore	
	-	6 *S. enterica* Typhimurium	Herbivore	
Screeching Owl	-	3 *S. enterica* Montevideo	Carnivore	-

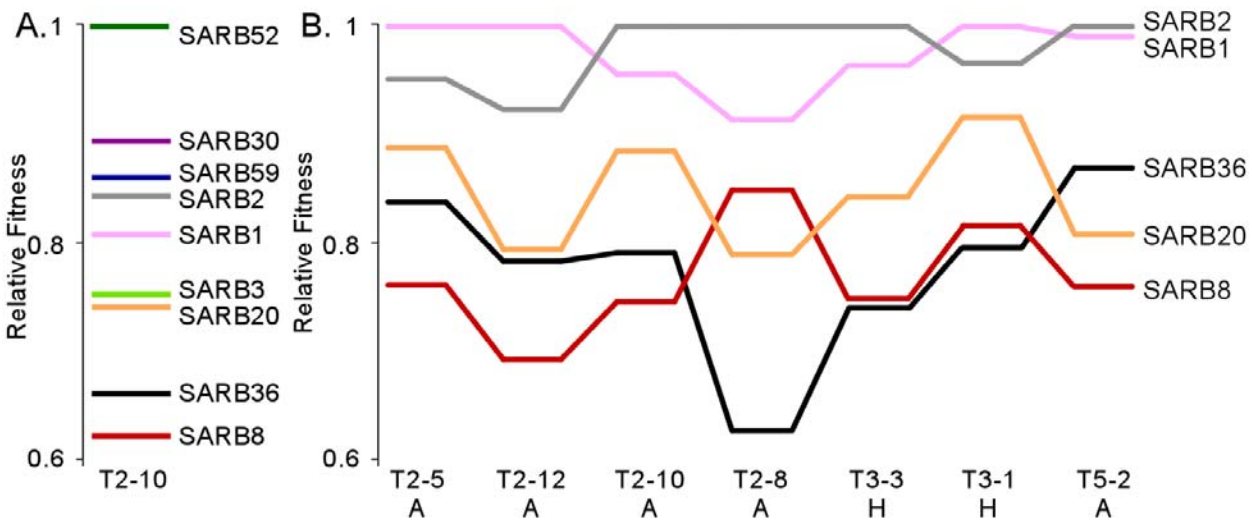

Figure 18. Feeding preferences of amoebae isolated from tadpoles.

Feeding preferences of intestinal amoebae isolated from bullfrog tadpoles. **A**. The feeding preferences of amoeba T2-10 as determined by pairwise line tests. The least preferred SARB strain was assigned a fitness of 1.0. **B**. Feeding references of 7 amoebae isolated from 3 tadpoles. Amoebae are labeled according to the tadpole from which it was isolated (T2, T3, or T5) followed by a strain number; A=*Acanthamoeba*, H=*Hartmannella*.

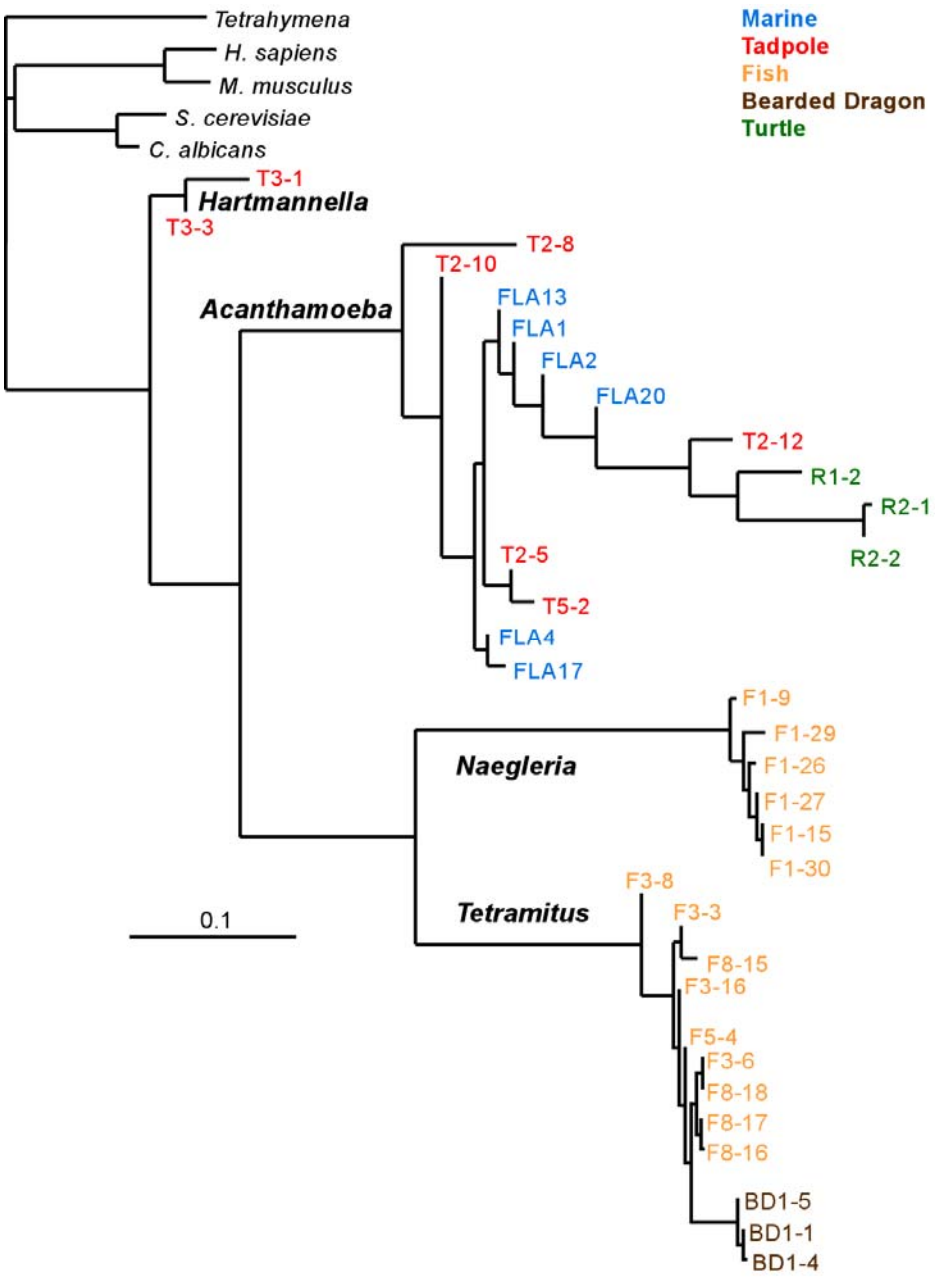

Figure 19. Neighbor-joining tree of amoebae isolated from environments.

Individuals of *Acanthamoeba*, *Naegleria*, *Hartmannella*, and *Tetramitus* isolated from different environments all grouped to their respective clades. Amoebae isolates with the same color were isolated from the same environment.

3.3.3 Amoebae within fish have similar feeding preferences.

The similarity in feeding preferences among the *Acanthamoeba* and *Hartmannella* isolated from tadpole intestines may reflect either the influence of factors within host from which they were isolated or more general, non-host-specific environmental conditions experienced by pond-dwelling creatures, such as the pH, the temperature of the water or the composition of the gut flora. To distinguish between these alternatives, I isolated amoebae from the intestines of goldfish (*Carassius auratus auratus*) which were reared at two temperatures and tested their feeding preferences. Fish F1 was housed at 30˚C; from it I isolated 5 strains of *Naegleria* that were >99% similar at their 18S rDNA loci (Table 3). Three other fish were housed at 23˚C; from them I isolated 11 strains of *Tetramitus* that were >98% similar at their 18S rDNA loci (Table 3). *Naegleria* and *Tetramitus* are only distantly-related, being only 57% identical at their 18S rDNA loci and represent members of different classes (Figure 19). As above, subsets of SARB strains were used to assay the feeding preferences of the sixteen amoebae.

The 5 *Naegleria* from fish F1 all had similar feeding preferences (Figure 20; R^2=0.8770, P<0.000001). Likewise, the 4 *Tetramitus* from fish F3 had similar feeding preferences (R^2=0.8732, P<0.000001) as did the 5 *Tetramitus* from fish F8 (R^2=0.9049, P<0.000001). These data reinforce the results obtained with tadpole T2: related amoebae from a single host animal collectively prefer one serovar over another. More importantly, all amoebae isolated from fish – either *Naegleria* from fish reared at 30˚C or *Tetramitus* from fish reared at 23˚C – shared common feeding preferences that are highly significantly similar (R^2=0.8596, P<0.000001). These data suggest that feeding preferences of intestinal amoebae reflect properties of the host, not common environmental conditions.

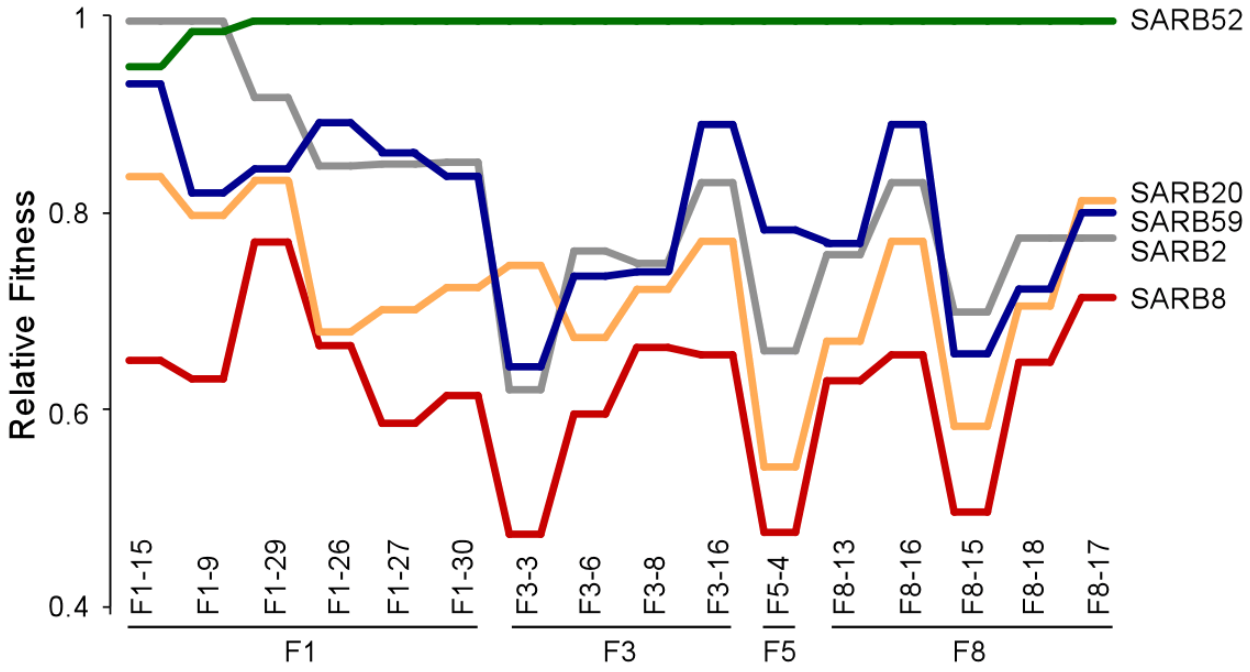

Figure 20. Feeding preferences of amoebae isolated from fish.

Feeding preferences of 16 amoebae isolated from 4 goldfish. Amoebae are labeled according to the fish from which it was isolated (F1, F3, F5 or F8) followed by a strain number.

3.3.4 Fish reared at different temperatures have different microbial flora.

To determine if goldfish reared at different temperatures had dissimilar microbial flora, Isolated gram-negative bacteria from their intestinal contents. Four strains of *Citrobacter* were isolated from fish F1 which was housed at 30˚C (Table 3); *Citrobacter* are antigenically diverse enteric bacteria (116) related to *Salmonella* and *E. coli*. In contrast, 12 strains of *Aeromonas* were isolated from fish F8, which was house at 30˚C (Table 3). A total of 10 strains of *Citrobacter* and 5 strains of *Aeromonas* were isolated from fish F2 and F5, which were transitioned from 30˚C to 23˚C and had been housed at 23˚C for a shorter period of time than had fish F8. These results suggested that the bacterial flora in the goldfish were dissimilar among the fish housed at different temperatures, just as the amoeboid predators were. In addition, I examined the O-antigens of the *Citrobacter* isolates by gel electrophoresis (Figure 21).

Although the identity of the O-antigen cannot be determined in this way, one may distinguish between different carbohydrates based the patterns seen in silver-stained gels. While *Citrobacter* has more than 45 different O-antigens, all strains isolated from goldfish were serologically identical or nearly identical (Figure 21). Together, these data show that amoebae from an intestinal environment have similar feeding preferences, regardless of their relatedness, and suggest that feeding preferences are a function of the host, not the identity of prey bacteria, temperature or other environmental factors.

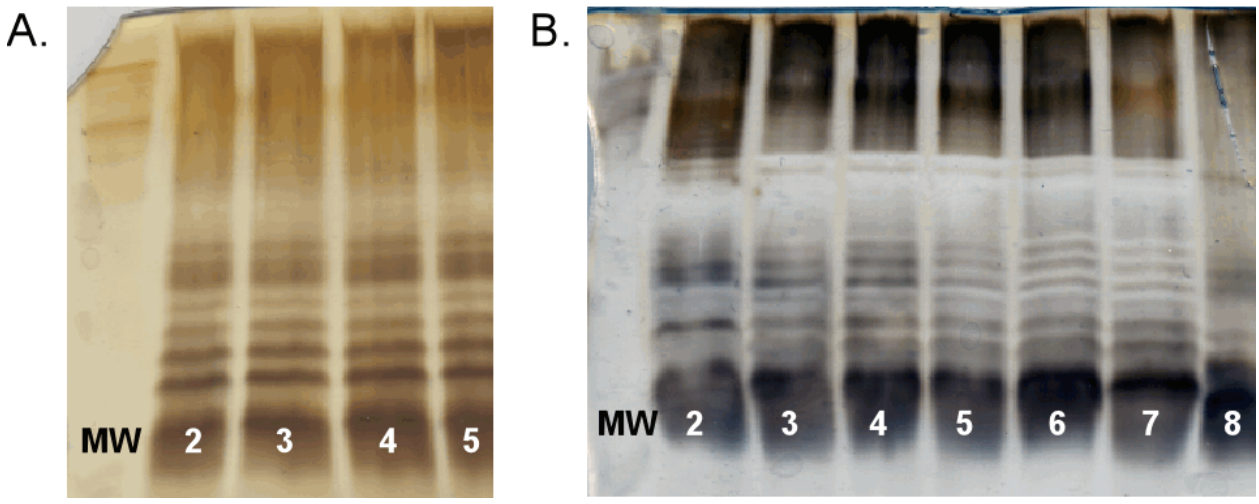

Figure 21. Silver stained O-antigens of bacteria isolated from fish.

A. Gel of 4 *Citrobacter* isolated from F1. B. Gel of 6 *Citrobacter freundii* strains isolated from the intestines of F3 (lanes 2, 3, 4, 5, 6, and 7) and one *Aeromonas hydrophila* strain isolated from F5 (lane 8). O-antigens are similar, if not identical, in appearance.

3.3.5 Amoebae from reptiles have similar feeding preferences.

Tadpoles and fish have relatively undifferentiated intestinal tracts. Because enteric bacteria also reside within hosts with more complex intestinal environments, I tested the hypothesis that amoebae from more differentiated intestines also share feeding preferences by isolating amoebae from reptiles. Three strains of *Acanthamoeba* (R1-2, R2-1, and R2-2), 88% similar based at 18S rDNA genes, were isolated from 2 red-eared sliders, *Trachemys scripta* (Table 3). Feeding preferences were determined as described above and were again more similar than expected at random (Figure 22; R^2=0.5096, P<0.038). Three strains of *Tetramitus* (BD1-1, BD1-4, BD1-5; >99% similar at their 18S rDNA loci) were isolated from the feces of a carnivorous, juvenile bearded dragon, *Pogona barbata* (Table 3), and 1 strain of *Acanthamoeba* (BD2-1) was isolated from an herbivorous adult. Line tests were performed as described previously and results show these amoebae share feeding preferences (Figure 23; R^2=0.4763, P<0.026). Thus, amoebae isolated from hosts with differentiated intestines also collectively prefer one *Salmonella* serovar over another; the larger variation in feeding profiles may reflect the greater diversity of habitats within more differentiated intestines.

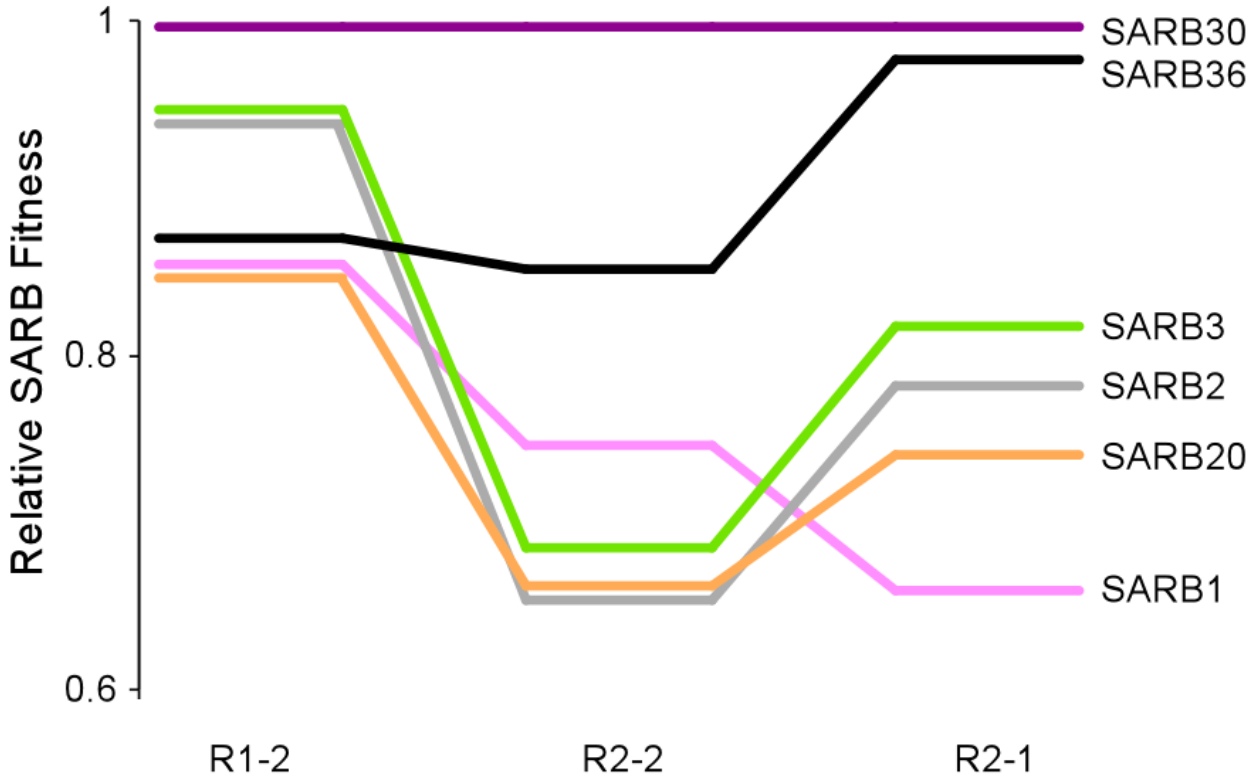

Figure 22. Feeding preferences of amoebae isolate from turtles.

Feeding preferences of 3 amoebae isolated from 2 turtles as determined by pairwise line tests. Amoebae are labeled according to the turtle from which it was isolated (R1, R2) followed by a strain number.

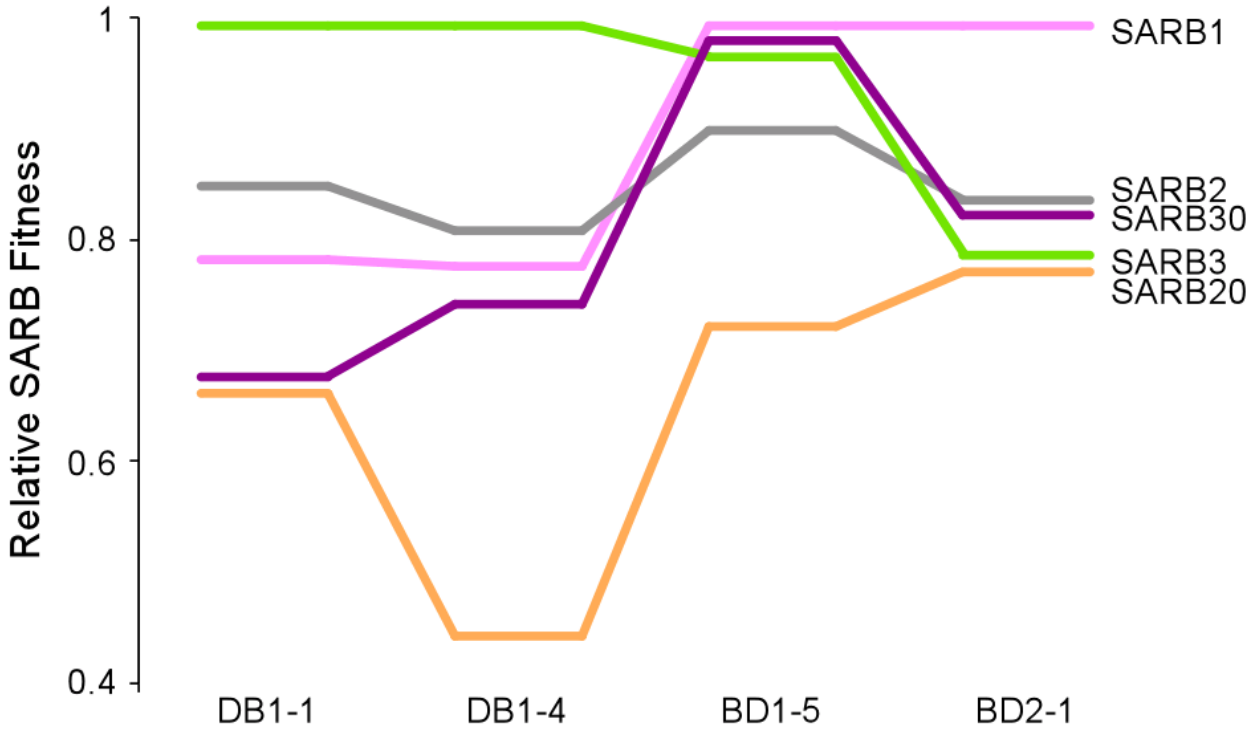

Figure 23. Feeding preferences of amoebae isolated from bearded dragons.

Feeding preferences of 4 amoebae isolated from the feces of 2 bearded dragons as determined by pairwise line tests. Amoebae are labeled according to the lizard from which it was isolated (BD1, BD2) followed by a strain number.

3.4 AMOEBOID PREDATION MAY EFFECT DIFFERENTIAL DISTRIBUTION OF *SALMONELLA*

3.4.1 Predators form environment-specific threats to particular prey.

Traditionally, microbiologists have viewed antigens – the highly variable proteins and carbohydrates that decorate the surface of bacterial cells – as being intimately associated with the adaptive immune system. These outer-most cellular structures are the most accessible to antibodies or white blood cells, so it is not surprising that they are quite variable among bacteria that interact with eukaryotes in this way. This need for pathogens to escape immune recognition continually selects for new, rare antigenic types; here, frequency-dependent selection (FDS) favors the newly-different cells by allowing for prolonged infections or the re-infection of non-naïve hosts (41, 86, 98, 195, 254, 262). Yet many non-pathogens are antigenically diverse, and many pathogens – like *Salmonella enterica* – have antigenic diversity that is inconsistent with FDS (268). For these bacteria, I proposed that intestinal protozoa – rather than bacteriophagic white blood cells – were the predators being avoided. Here, a *Salmonella* serotype may have an advantage in escaping all of the predators in a particular environment; because predators are not uniformly distributed, different prey serotypes flourish in different environments, a process germs diversifying selection (DS). A critical prediction of this model is that predators within

particular intestinal environments should share a common set of feeding preferences, thus allowing for some prey to escape predation better than others.

The data collected here uniformly support this position in five different systems: marine-isolated *Acanthamoeba* shared similar feeding preferences (Figure 17); tadpole-born *Acanthamoeba* and *Hartmannella* also showed significant similarity in prey choice (Figure 18); fish-born *Naegleria* and *Tetramitus* showed striking similarity in feeding preferences even though their hosts were reared at different temperatures (Figure 20); and both *Acanthamoeba* from turtles and *Tetramitus* and *Acanthamoeba* from bearded dragons showed similar feeding preferences after isolation from the differentiated colons of the reptilian hosts (Figures 22 and 23, respectively). It is perhaps not surprising that related amoeba – *e.g.,* those from the same species – would share similar feeding preferences, since the proteins mediating prey recognition would be highly similar. Being closely related, one may expect to find them in the same environment. Therefore, one could say that the similar feeding profiles of *Acanthamoeba* within tadpoles (Figure 18) and *Tetramitus* or *Naegleria* within fish (Figure 20) are not terribly surprising.

But what is unexpected is that unrelated amoebae share feeding preferences if and only if they are found in the same environment (Figure 24). Unrelated amoebae may be found in a single host due to fluctuating conditions. For example, only *Hartmannella* were isolated from the smallest tadpole, which was likely herbivorous, whereas only *Acanthamoeba* was isolated from the largest tadpole whose carnivorous diet would have fostered growth an entirely different microflora. Similarly, only *Naegleria* were found in the fish housed at 30˚C, whereas only *Tetramitus* were isolated from fish housed at lower temperatures. Yet in both these cases, the unrelated amoebae from each host shared a common set of feeding preferences, despite other

major environmental variables between individuals. I can conclude that there must be other, host-specific environmental factors within the intestinal lumen that influence the residents. That is, amoebae which do not share these feeding preferences do not persist, selecting for sets of unrelated predators with common proclivities in prey choice.

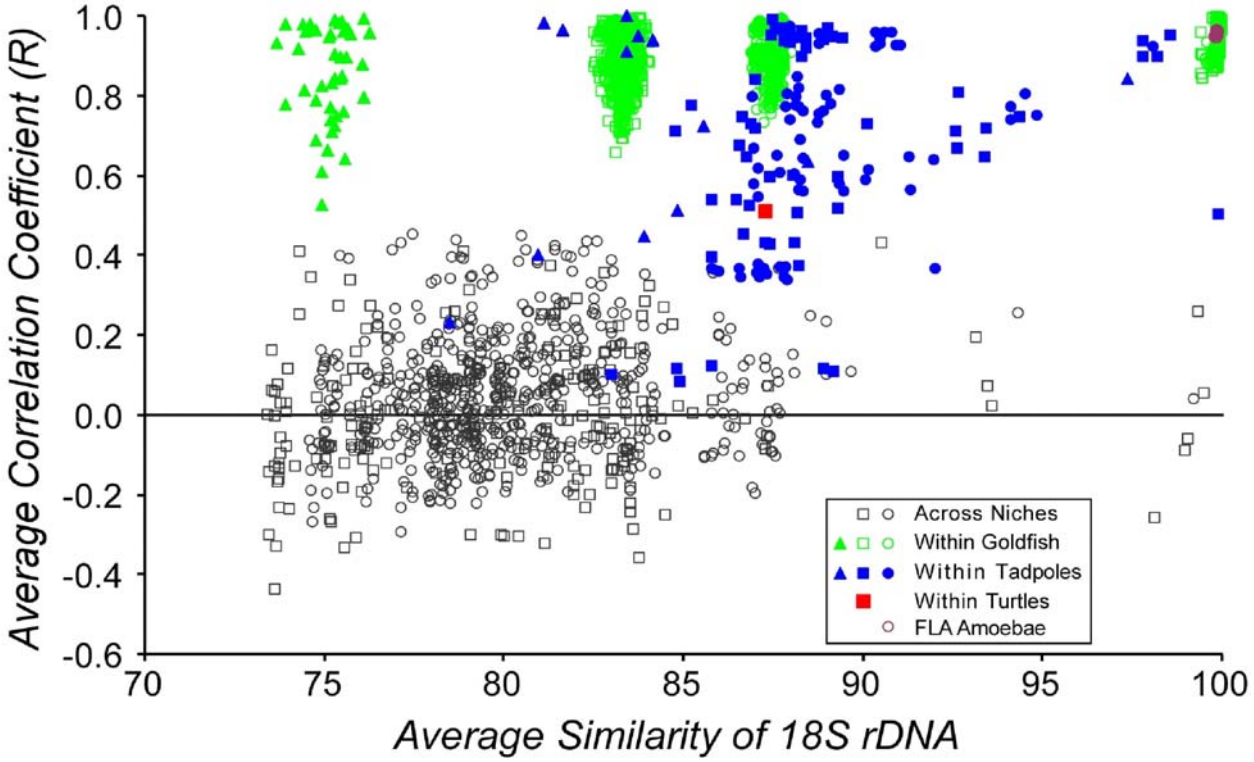

Figure 24. Similarity of amoebae feeding preferences as a function of genetic relatedness.

Data are gathered from Figures 17, 18, 20, 22, 23, and from reference (268). Average Pearson's correlation coefficients are reported for groups of 2, 3 and 4 line tests, plotted against the average similarity at the 18S rDNA locus. Comparison of feeding preferences between amoebae isolated from different environments are shown in gray, among tadpole-isolated amoebae in blue, among goldfish-isolated amoebae in green, among turtle-isolated amoebae in red and among FLA amoebae in violet. Triangles represent comparisons with 2 amoebae's feeding preferences; comparisons using 3 sets of preferences are shown as squares and those using 4 sets of preferences are shown as circles. Open (transparent) symbols are used when a large density would obfuscate the number of data points.

3.4.2 Predators discriminating among environmental carbohydrates.

As the most abundant molecular on the outside of the cell, the O-antigen of bacterial lipopolysaccharide is a likely target for amoeboid predators; consistent with this hypothesis, I have shown that predators do discriminate among prey based on the identity of the O-antigen (268). Yet amoebae encounter other abundant polysaccharides in their intestinal environments, notably mucins. Mucins are proteins that are heavily substituted with oligosaccharides that are O-linked to serine and threonine sites. Generally, the mucin protein consists of a core made up of variable tandem repeats rich in serine, threonine, and proline (201, 202, 253). Attached to the end of the oligosaccharides is usually a sialic acid or sulphate group which contributes to the negativity of the protein or a blood group epitope (201-203, 253). In humans, six core main mucin structures appear throughout the intestinal tract, and modification of these occurs through different glycosylations and differential formation of oligosaccharide side chains which results in tens of dissimilar structures (123, 201-203). Mucins are either secreted or attached to intestinal cells and the primary functions of mucins are thought to be intestinal protection and aiding in gut flora binding (2, 13, 43, 51, 223, 234). While O-antigens appear on bacterial prey, mucins decorate the intestinal wall; and while bacteria are viable food sources, the intestinal wall is not.

I propose that amoebae will differentiate between structures that should and should not be eaten. Recognition of host-specific mucin polysaccharides would allow the amoebae to use these sugars for simple attachment, while avoiding attempts at consuming the intestinal wall. Consistent with this hypothesis, the commensal human amoeba *Entamoeba histolytica* has been shown to strongly bind to the abundant mucin sugar N-acetylgalactosamine (GalNAc) (1, 91, 128, 186, 193). This sugar may be a receptor attachment site of *E. histolytica* allowing it to reside in its adapted niche and avoid rapid expulsion from the colon. It is unlikely that *E.*

histolytica uniformly binds to all mucins since they differ in structure. If amoebae differentially bind to mucins, then O-antigens more similar to native intestinal mucins may provide a higher fitness to that bacterium via host mimicry. That is, bacteria whose O-antigens resemble the local mucins may escape predation more readily because they are recognized as housing, not as food. Mucins do not vary between members of a vertebrate host species with different diets or those reared at different temperatures, so I would predict that even the unrelated amoeboid predators that take up residence under these different environmental conditions would avoid eating bacteria that resemble local mucins, thus providing them with similar feeding preferences. And since mucins do vary between host species, I would predict that amoebae isolated from separate environments would avoid different serovars of *Salmonella*. As a whole, if feeding preferences of intestinal amoebae are influenced by native mucins, then these mucins are important influences on feeding phenotypes of amoebae through niche adaptation and (indirectly) shape the composition of the bacterial flora.

3.4.3 Differential distribution of *Salmonella* may result from predation.

Salmonella serovar-host specificity has historically been viewed as a product of bacterial interaction with host immune systems, whereby a serovar, expressing a specific O-antigen, could infect a certain host after immune evasion and then cause disease. Previously, I showed that amoeba are a possible selective pressure influencing O-antigen variability (268), and here, I show that groups of amoebae within an environment collectively prefer one serovar over another (Figures 17, 18, 20, 22, and 23). Furthermore, amoebae between dissimilar environments have different feeding preference (Figures 14, 24, and 25). As a result, amoeboid predation may influence bacterial survival in environments resulting in the differential distribution of bacteria

among hosts: that is, bacteria may be found in an environment because they can survive better against native predators. Under this model, *Salmonella* serovar-host specificity may have originated after a serovar had established the ability to escape native predators in certain environments. Following adaptation to its specific niche, a serovar can acquire genes allowing it to infect that host and cause disease. That is, the specificity for *Salmonella* in causing disease more readily in particular hosts may be intimately associated with that serovar's ability to avoid the predators within that host; *Salmonella* must avoid predation before it invades intestinal epithelium. Consistent with this hypothesis, *Salmonella* and *E. coli* have been found to be differentially distributed among the intestinal environments of their hosts (65-67, 191). I found similar result here where, for example, the serovars of *Salmonella* within turtles and bearded dragons were significantly different from those I isolated form birds (Table 1). Because a single O-antigen would not confer high fitness in all environments, O-antigen (and *rfb*) variability would be maintained among *Salmonella*. Thus protozoan predation may be the selective pressure maintaining O-antigen diversity among *Salmonella*.

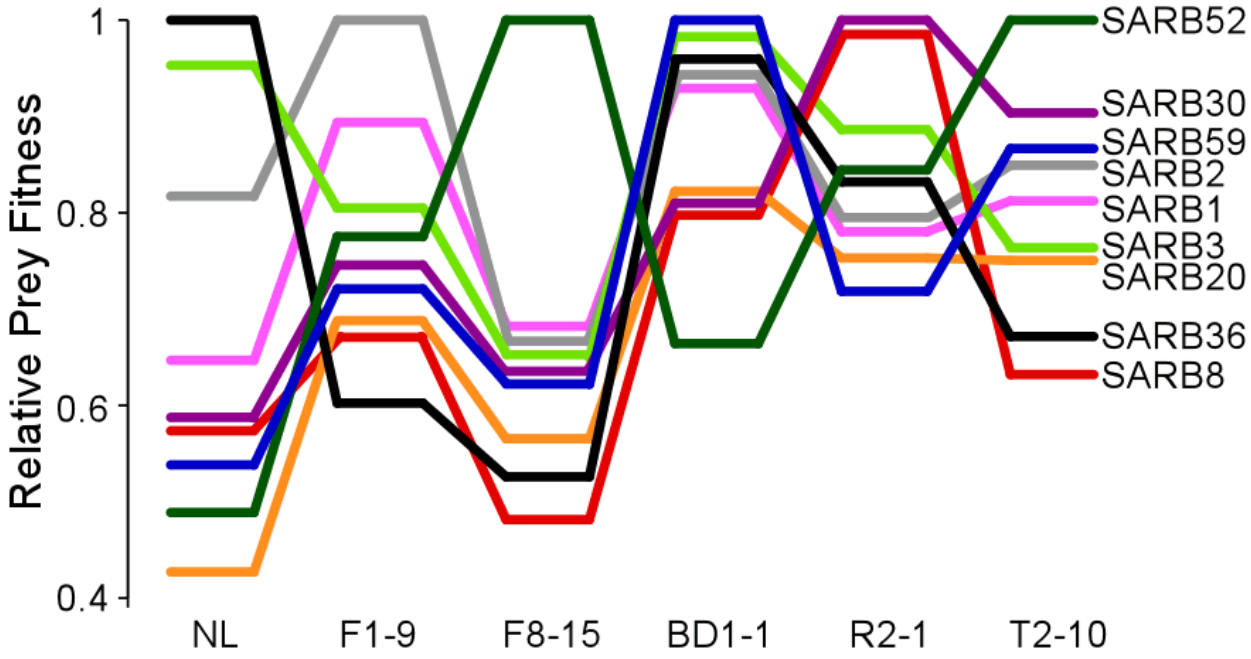

Figure 25. Amoebae from separate environments have different feeding preferences.

Data are gathered from Figures 17, 18, 20, 22, 23, and from reference (268). Each amoeba was tested against 9 SARB strains. The feeding preferences of amoeba were determined by pairwise line tests using 9 SARB strains. The least preferred SARB strain was assigned a fitness of 1.0. Amoebae are labeled as previously described in this chapter; NL is a lab strain of *Naegleria*.

4.0 CONCLUDING REMARKS ON *RFB* DIVERSITY

O-antigen diversity among *Salmonella* was originally thought to be maintained by FDS mediated by exposure to the immune system (196). This hypothesis was believed to adequately explain O-antigen diversity among *Salmonella*. In Chapters 1.6 and 2, I discussed why this hypothesis does not work with *rfb* diversity observed among *Salmonella*, and I proposed that selective pressures from protozoan predation may be influencing *rfb* diversity among *Salmonella* which is maintained by DS. In Chapter 2, experiments were performed to falsify my proposal that protozoan predation is the selective pressure influencing *rfb* diversity, and in Chapter 3 I tested if *rfb* diversity could be retained through DS. Now, I will revisit my results and show that my hypothesis provides the framework for understanding *rfb* diversity and it gives insight to the previously unexplained observations.

4.1 THE ORIGIN OF *RFB* DIVERSITY

The O-antigen evolved in a way which provides protection to Gram-negative bacterium and possibly a fitness advantage under particular ecological conditions. This polysaccharide, which is the most abundant outer membrane structure and covers the entire surface of the cell, extends out from the bacterial cell surface through its attachment of core and the lipoprotein Lipid A (Figure 1). With the exception of flagellae and fimbrae, which are not constitutively expressed, the O-

antigen is the most distal structure from the cell, and therefore will likely be the first – if not the only – molecule on the bacterial cell that contacts its environment. The numerous different O-antigens within a bacterial species suggest that this diversity has evolved in a way by which dissimilar structures provide a fitness advantage under particular circumstances. Population level variability observed at the *rfb* locus is certainly not the result of the accumulation of neutral variants. While some organisms have maintained mechanisms for active switching of O-antigen phenotypes others have evolved novel O-antigens through the process of lateral gene transfer (LGT), not the process of random mutation. Because the origin of *rfb* genetic diversity is different between bacteria this suggests that not all of these sugar structures evolved for the same reason.

As discussed in Chapter 1, the numerous serovars of *Salmonella enterica* are not created by genetic reassortment of intragenomic information on temporal scales; rather, stable alleles were created through LGT and maintained by selection in different serovars [Figure 9, (139, 194, 256)]. LGT works through homologous or illegitimate recombination in which a gene or set of genes are incorporated into the recipient genome from some source other than its parent. The incorporation of a gene or genes into the *rfb* region through lateral transfer is not a process that occurs during the course of an infection such as the changing of an O-antigen through phase variation which occurs in *Neisseria* or *Haemophilus* (Table 1). Compared to phase variation, LGT rarely occurs (124).

Some infrequent events must occur for lateral transfer to be successfully incorporated into a bacterial genome that results in expression of a different O-antigen. First, the DNA must be incorporate into the recipient's cell. DNA encoding genes must be transferred through either phage transduction, bacterial conjugation, or by taking up naked DNA. This DNA can encode

either a single gene or a large cluster of genes. Second, the foreign DNA must escape degradation by the restriction/modification system and then recombine into the genome. Finally, the region must be expressed and provide a fitness advantage so the genes can be selected for and retained. All these events happening together rarely occur. Consequently, LGT does not contribute to *rfb* diversity among *Salmonella* through a process that rapidly generates diversity over a short period of time, so selective pressures influencing O-antigen diversity are not temporal. Instead, selective pressures likely act spatially and LGT is the mechanism that generates O-antigen diversity.

Extensive genetic diversity is observed at the *Salmonella rfb* locus because serovars have acquired different genes through LGT. These genes, encoding sugar synthesis and transferase products, are located in the same chromosomal region among *Salmonella* but can be completely different between serovars thus making dissimilar O-antigens (Figures 8 and 9). Many bacteria have these genes grouped in together in operons which are expressed as either one or several transcripts. Because most genes that make the O-antigen are located in one region, LGT is a mechanism that can produce many different O-antigen by the transfer of either the entire functional genetic region which could result in the expression of a new O-antigen or the transfer of a few genes which may result in the addition of a sugar to an existing O-antigen. The gain of an entirely new O-antigen or the modification of an existing one may give a fitness advantage to a serovar in its environment. Although LGT is a rare event, it is a good mechanism to generate diversity and give rise to very different O-antigens that may increase fitness.

The origin of *rfb* diversity among *Salmonella* originated from LGT and is maintained by selective pressures. Because LGT rarely occurs and results in genetic diversity which produces stable O-antigens, compared to other mechanisms like phase variation, selective pressures

influencing this locus may be constant in an environment in which rapid generation of diversity is not needed for a higher fitness. LGT is the mechanism that generates extensive genetic diversity among *Salmonella* and selective pressures may act spatially thus maintaining O-antigen diversity at a population level.

4.2 THE MAINTENANCE OF RFB DIVERSITY

Amoebae are excellent candidates which may provide the selective pressure maintaining O-antigen diversity. These single celled eukaryotic organisms are residents of the intestines and generally 60 times larger than average length of a bacterium (~1ul) giving this predator the ability to engulf tens of bacterial prey through every phagocytic event. In addition, amoebae are general scavengers that consume bacterial prey they likely identify through cell-to-cell contact. Being general predators, amoebae do not actively search for *Salmonella* serovars which make up less that 1% of the bacterial flora in warm blooded animals; rather, amoebae prey upon all intestinal bacteria. These qualities make amoebae predator candidates for maintaining O-antigen diversity among *Salmonella*. Other protozoa such as ciliates or flagellates are less probable sources of selective pressure. Ciliates have an oral groove and may not discriminate between *Salmonella* serovars as my results suggest (Figure 16) while flagellates are usually thrive in the water column and not in the intestines with undigested food debris. Thus, amoebae predators may be important players that strongly influence O-antigen diversity among *Salmonella*.

Amoebae presumably bind to and recognize *Salmonella* and other bacteria through cell-to-cell contact. Amoebae isolated from human intestines have been shown to differentially bind

to intestinal sugars through a receptor-epitope binding interaction. This method may represent a similar mechanism by which amoebae recognize its prey. Since the O-antigen is the most abundant molecule on the bacterial cell surface and is made up of sugar molecules, it is an easy target for amoebae predators and may be involved in binding. If true, then amoebae may differentially recognize different *Salmonella* serovars based on the O-antigen and prefer one serovar over another by way of receptor-epitope interactions. My results suggest that the amoebae do differentially prey upon serovars (Figures 14 and 25) and that the O-antigen is sufficient to elicit an amoeboid feeding preference (Figure 16).

Interactions between an amoeba and its ecological environment may play a critical role in shaping prey recognition. My results show that unrelated amoebae from the same environments have similar feeding preferences while related amoebae from separate environments have different feeding preferences (Figures 18, 20, 22, 23, and 24). This suggests that interactions within the environment are important in shaping amoebae feeding hierarchies; if amoebae are adapted to a specific intestinal environment, then that habitat may shape its lifestyle. For instance, amoebae found in the intestinal tracts of freshwater goldfish may be more adapted to a colder fluidic environment with low concentrations of oxygen compared to amoebae found in the human intestinal environment which is a differentiated intestine that usually removes water, is warmer, and anoxic in regions. Besides these factors, the intestines also are packed with mucins which amoebae seemingly are in constant contact. Amoebae may be adapted to these intestinal sugars and recognize them as non-prey objects. If this is true, then amoebae receptors could have evolved that recognize certain mucins, and particular amoebae reside in certain environments because they are adapted to the mucins as well as other biotic and abiotic factors.

Amoebae that are adapted to an environment may persist there because they strongly attach to the intestinal mucin. These intestinal predators may recognize the mucins as their environment (not as prey) and strongly bind to these sugar structures through receptor interactions preventing expulsion from the gut. This binding may shape amoeboid feeding preference if amoebae receptors use these mucin adapted receptors to identify an O-antigen. O-antigens resembling mucins may camouflage the serovar and provide a fitness advantage through predator avoidance. Because different animals express dissimilar mucins, amoebae may be found in separate intestines that have different feeding preferences - this is what my results show (Figure 14, 24, and 25). Thus, a *Salmonella* serovar which expresses an O-antigen that resembles the intestinal mucin may have a higher fitness in an intestinal niche (as discussed in Chapter 3.4.2).

Mucins may also affect the stability of amoeboid feeding preferences. Although mucins have been shown to change between intestinal regions of a host such as between the small and large intestine, mucin structure is constant within a particular gut niche. If amoebae reside in a particular region within the gut and are adapted to the mucins there, then a prediction would be that amoebae from that intestinal region would have stable and similar feeding preferences. My results suggest that both these concepts are true. First, independent amoebae isolates from a particle intestinal environment share preferences (Figures 18 and 20). If feeding preferences rapidly changed, then amoebae isolated from the same environment would show different preferences. Instead, amoebae preferences reflect stability which may be the result of constant binding to a mucin in a specific intestinal region. Second, amoebae from the same environment share feeding preferences. All amoebae isolated from either tadpoles (Figure18) or goldfish (Figure20) share preferences which may reflect the undifferentiated intestine of these hosts and

little mucin diversity. If amoebae reside in animals with a differentiated intestine and are adapted to the mucins in the separate intestinal regions, then a prediction would be that amoebae from that gut but from separate regions would have different feeding preferences. My results showing the preferences of amoebae isolated from reptiles may hint at this concept (Figures 22 and 23). Although feeding preferences are significantly more similar than expected at random, the P-value was not as strong compared to the feeding preferences of other amoebae isolated from nondifferentiated intestines (Figures 17 and 20). This may be due to the small sample size or possibly these amoebae are adapted to different intestinal habitats and have slightly different preferences. Taken together, intestinal mucins may play important roles in shaping both the feeding preferences and the stability of these preferences among amoebae.

If amoeboid feeding preferences are indeed shaped by mucins, then a serovar that resides in a particular intestinal niche may have a higher fitness advantage due to the O-antigen it expresses. Serovars expressing an O-antigen similar to the mucin an amoebae is adapted to would have higher fitnesses than serovars with dissimilar O-antigens. As a result, a range of amoebae feeding hierarchies would occur which my data suggests (Figures 12, 13 and Table 2). Because there are different mucins in separate intestinal environments, amoebae from particular habitats would exhibit different feeding preferences as shown in Figures 14 and 25 while amoebae from the same habitats would show similar feeding preferences (Figures 17, 18, 19, 22, and 23). Moreover, amoebae only exhibit similar feeding preferences if they were isolated from the same environment (Figure 24). For the O-antigen diversity to result from amoeboid predation then the O-antigen itself would have to influence feeding preferences which my results confirm (Figure 16). In concert, O-antigen diversity may be maintained by amoeboid predation.

Salmonella serovars may be found in a particular intestinal habitat as a result of amoeboid predation. A serovar may spend most of its lifecycle in a particular gut environment and become adapted to its niche. Selective pressures from within this environment may be the most influential pressures shaping this bacterium resulting in a serovar that is the most fit in its niche. This is what may be observed among *Salmonella* serovars; serovar host specificity suggests that certain serovars usually infect particular hosts. For instance, when a pig is infected by *Salmonella*, it is usually serovar Cholereasuis which expresses the O-antigen epitope 6,7. If a horse gets Salmonellosis, it is usually serovar Abortusequis expressing the O-antigen epitope 4,12. Amoeboid selective pressures from spatially separated intestinal environments of different hosts may be influencing O-antigen diversity. If a serovar retains a *rfb* locus because it provides a selective advantage in a particular host environment then O-antigen diversity could be maintained and certain serovars may have higher fitness values in some environments compared to others as observed with serovar host specificity.

Diversity maintained through spatial selective pressures, such as predation between different intestinal environments, can be explained by diversifying selection which denotes that fitness of an individual depends on the environment it resides. This is what my results show - *Salmonella* serovars do well when they are tested against amoebae from separate intestinal environments (Figures 14 and 25). This suggests that selective pressures influencing O-antigen diversity are acting over space and different *rfb* regions obtained through LGT provide fitness advantages to serovars in other ecological habitats. By knowing the mechanism of diversity and the process by which the selective regime acts, the component maintaining variability should be revealed.

In conclusion, I tested my hypothesis that protozoan predation may be a selective pressure maintaining *rfb* diversity among *Salmonella*. The results I presented in Chapter 2 showed (i) a predator prefers one serovar over another, (ii) not all predators have the same feeding preference, and (iii) predation is influenced by the O-antigen itself. All which suggest that predation may be an important force influencing the O-antigen of *Salmonella*. In Chapter 3, I showed (i) amoeba feeding preference are stable (ii) groups of amoebae in an environment share feeding preferences, and (iii) groups of amoebae between environments have different preferences. These results suggest that diversifying selection is a good model used to explain the extensive genetic diversity observed at the *rfb* locus. Thus, protozoan predation may be a strong selective pressure maintaining *rfb* diversity among *Salmonella* which is explained through diversifying selection.

4.3 FUTURE DIRECTIONS

My results established a foundation that O-antigen diversity among *Salmonella* may be due to protozoan predation. I showed that amoebae from intestinal environments prefer certain *Salmonella* serovars over others and that this is mediate by the O-antigen. Amoebae feeding preferences must be influenced by binding and recognition of an amoeba to a serovar. This will be investigated to help determine why amoebae exhibit feeding preferences and how they effect *Salmonella* survival. To determine if differential distribution of *Salmonella* among hosts may be a result of predation preference, a more comprehensive study of amoebae in differentiated

intestines needs to be performed. This will address if different types of amoebae are present in particular intestinal segments which *Salmonella* may encounter. Other *in vivo* experiments will determine if amoeboid predation occurs in the intestines. These studies will shed light upon predation in the intestinal environment and help determine if serovar host specificity can be a result of predation. Below are three aims extending from my work presented in Chapters 2 and 3 which further investigate our hypothesis that protozoan predation is the selective pressure causing O-antigen diversity which is maintained by DS.

Aim 1. Are amoebae feeding preferences mediated by differential binding to serovars?

Amoeboid feeding preference may be mediated by differential binding to the O-antigen. We have begun testing this hypothesis using a binding assay and preliminary results suggest that amoebae bind more strongly to the serovar they prefer to eat. Other strategies like microscopy will also be employed to determine if an amoeba binds one *Salmonella* serovar more than another. Together, these studies will shed light upon the feeding preferences of amoebae and further characterize the interactions between these organisms.

Aim 2. What is the diversity of amoebae in differentiated intestinal environments?

Amoebae isolated from a differentiated intestinal environment may be diverse. *Salmonella* usually resides in complex differentiated intestinal environments such as reptiles. I isolated seven amoebae from the intestines of four reptiles (Table 3). My results show that these amoebae do prefer one serovar over another. To get a more comprehensive understanding of amoebae and if different ones have separate feeding preferences based on their intestinal location, more amoebae need to be isolated from specific intestinal locations and tested for their feeing preferences. This would help in characterizing amoebae from a particular environment.

Aim 3. Does predation influence differential distribution of Salmonella among hosts?

Differential distribution of *Salmonella* among hosts may occur if amoebae exhibit feeding preferences in the intestinal environment. My *in vitro* predation results show that amoebae from the same environment share feeding preferences but between environments preferences differ (Chapter 3). To determine if predation effects *Salmonella* differential distribution among hosts, both amoebae and *Salmonella* need to be isolated from the same hosts. These isolates would then be tested through *in vitro* and *in vivo* experiments. *in vitro* experiments would show if predation occurs with natural *Salmonella* isolates. *in vivo* experiments would show if predation occurs in the intestinal environment. These tests would help address if *Salmonella* is differentially distributed among hosts due to amoeboid predation.

My studies are the first ones which test a hypothesis involving selective pressures that may act within an ecological habitat and influence bacterial diversity. I have discussed extensive genetic diversity among bacteria and have speculated on the selective regimes causing diversity and its maintenance (Chapter 1). In Chapters 2 and 3, I focused on *rfb* diversity among *Salmonella* and tested if protozoan predation may be a selective pressure causing O-antigen diversity among *Salmonella* which is subsequently maintained by DS. I performed experiments that have shed light on these predator-prey interactions. This research answered basic questions about selective pressures influencing O-antigen diversity among *Salmonella*, but more importantly, it opened the door for other investigations and new questions to be asked. Intestinal environments are very complex habitats consisting of numerous bacteria, protozoa, and elements of the hosts itself (84, 132, 201, 205). Selective pressures that most strongly influence a bacterium's genotypic and phenotypic structure almost certainly originate from where the bacterium spends most of its

lifecycle. *Salmonella* spends most of its lifecycle in the intestinal environment so it is here where selective pressures act to shape this organism. Over 150 years ago, selection was first proposed by Charles Darwin who laid down the foundation for understanding species and the diversity between them. Now, we try to identify these pressures and how they shape individuals.

BIBLIOGRAPHY

1. **Adler, P., S. J. Wood, Y. C. Lee, R. T. Lee, W. A. Petri, Jr., and R. L. Schnaar.** 1995. High Affinity Binding of the *Entamoeba histolytica* Lectin to Polyvalent N-Acetylgalactosaminides. J Biol Chem **270:**5164-5171.
2. **Aksoy, N., D. Thornton, A. Corfield, C. Paraskeva, and J. Sheehan.** 1999. A study of the intracellular and secreted forms of the MUC2 mucin from the PC/AA intestinal cell line. Glycobiology **9:**739-746.
3. **Alvarez-Castro, J. M., and G. Alvarez.** 2005. Models of general frequency-dependent selection and mating-interaction effects and the analysis of selection patterns in *Drosophila* inversion polymorphisms. Genetics **170:**1167-1179.
4. **Anxolabehere, D.** 1980. The influence of sexual and larval selection on the maintenance of polymorphism at the sepia in *Drosophila* melanogarter. Genetics **95:**743-755.
5. **Appelmelk, B. J., Y.-Q. An, T. A. M. Hekker, L. G. Thijs, D. M. MacLaren, and J. de Graaf.** 1994. Frequencies of lipopolysaccharide core types in *Escherichia coli* strains from bacteraemic patients. Microbiology **140:**1119-1124.
6. **Avery, S. V., J. L. Harwood, and D. Lloyd.** 1995. Quantification and characterization of phagocytosis in the soil amoeba *Acanthamoeba castellanii* by flow cytometry. Appl Environ Microbiol **61:**1124-1132.
7. **Ayala, F., and C. Campbell.** 1974. Frequency dependent selection. Annu Rev Ecol Syst **5:**115-138.
8. **Barcus, V. A., A. J. Titheradge, and N. E. Murray.** 1995. The diversity of alleles at the hsd locus in natural populations of *Escherichia coli*. Genetics **140:**1187-1197.
9. **Bart, A., J. Dankert, and A. van der Ende.** 1999. Antigenic variation of the class I outer membrane protein in hyperendemic *Neisseria meningitidis* strains in the netherlands. Infect Immun **67:**3842-6.
10. **Baum, J., A. W. Thomas, and D. J. Conway.** 2003. Evidence for diversifying selection on erythrocyte-binding antigens of *Plasmodium falciparum* and *P. vivax*. Genetics **163:**1327-1336.
11. **Baumler, A. J., B. M. Hargis, and R. M. Tsolis.** 2000. Tracing the origins of *Salmonella* outbreaks. Science **287:**50-52.
12. **Bayliss, C. D., D. Field, and E. R. Moxon.** 2001. The simple sequence contingency loci of *Haemophilus influenza* and *Neisseria meningitidis*. J Clin Invest **107:**657-662.

13. **Belley, A., K. Keller, M. Gottke, and K. Chadee.** 1999. Intestinal mucins in colonization and host defense against pathogens. Am J Trop Med Hyg **60:**10-5.
14. **Beltran, P., J. M. Musser, R. Helmuth, J. J. r. Farmer, W. M. Frerichs, I. K. Wachsmuth, K. Ferris, A. C. McWhorter, J. G. Wells, A. Cravioto, and e. al.** 1988. Toward a population genetic analysis of *Salmonella*: genetic diversity and relationships among strains of serotypes S. choleraesuis, S. derby, S. dublin, S. enteritidis, S. heidelberg, S. infantis, S. newport, and S. typhimurium. PNAS **85:**7753-7757.
15. **Bentley, S., D,, D. M. Aanensen, A. Mavroidi, D. Saunders, E. Rabbinowitsch, M. Collins, K. Donohoe, D. Harris, L. D. Murphy, M. A. Quail, G. Samuel, I. C. Skovsted, M. S. Kaltoft, B. Barrell, P. R. Reeves, J. Parkhill, and B. G. Spratt.** 2006. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. PLoS Genet **2**.
16. **Benz, R.** 1988. Structure and function of porins from gram-negative bacteria. Annu Rev Microbiol **42:**359-393.
17. **Berrington, A. W., Y. C. Tan, Y. Srikhanta, B. Kuipers, P. van der Ley, I. R. Peak, and M. P. Jennings.** 2002. Phase variation in meningococcal lipooligosaccharide biosynthesis genes. FEMS Immunol Med Microbiol **34:**267-275.
18. **Bettarel, Y., T. Sime-Ngando, C. Amblard, and J. Dolan.** 2004. Viral activity in two contrasting lake ecosystems. Appl Environ Microbiol **70:**2941-51.
19. **Bloch, M. A., and C. Desaymard.** 1985. Antigenic polymorphism of the LamB protein among members of the family Enterobacteriaceae. J Bacteriol **163:**106-110.
20. **Bonifield, H. R., and K. T. Hughes.** 2003. Flagellar phase variation in *Salmonella enterica* is mediated by a posttranscriptional control mechanism. J Bacteriol **185:**3567-3574.
21. **Boyd, E. F., K. Nelson, F. S. Wang, T. S. Whittam, and R. K. Selander.** 1994. Molecular genetic basis of allelic polymorphism in malate dehydrogenase (*mdh*) in natural populations of *Escherichia coli* and *Salmonella enterica*. PNAS **91:**1280-1284.
22. **Boyd, E. F., F.-S. Wang, P. Baltran, S. A. Plock, K. Nelson, and R. K. Selander.** 1993. *Salmonella* reference collection B (SARB): strains of 37 serovars of subspecies I. J Gen Micriobiol **139:**1125-1132.
23. **Branca, G., and H. Dym.** 2003. *Haemophilus influenza* cellulitis. A review and case report. N Y State Dent J **69:**34-36.
24. **Brisson, D., and D. E. Dykhuizen.** 2004. *ospC* diversity in *Borrelia burgdorferi*: different hosts are different niches. Genetics **168:**713-722.
25. **Brook, I.** 2003. Microbiology and management of acute suppurative thyroiditis in children. Int J Pediatr Otorhinolaryngol **67:**447-451.
26. **Brown, P. K., L. K. Romana, and P. R. Reeves.** 1991. Cloning of the *rfb* gene cluster of a group C2 *Salmonella* strain: comparison with the *rfb* regions of groups B and D. Mol Microbiol **5:**1873-81.
27. **Brown, P. K., L. K. Romana, and P. R. Reeves.** 1992. Molecular analysis of the *rfb* gene cluster of *Salmonella* serovar muenchen (strain M67): the genetic basis of the polymorphism between groups C2 and B. Mol Microbiol **6:**1385-94.
28. **Caugant, D. A., B. R. Levin, and R. K. Selander.** 1981. Genetic diversity and temporal variation in the *E. coli* population of a human host. Genetics **98:**467-490.
29. **Cerdeno-Tarraga, A. M., S. Patrick, L. C. Crossman, G. Blakely, V. Abratt, N. Lennard, I. Poxton, B. Duerden, B. Harris, M. A. Quail, A. Barron, L. Clark, C.**

Corton, J. Doggett, M. T. G. Holden, N. Larke, A. Line, A. Lord, H. Norbertczak, D. Ormond, C. Price, E. Rabbinowitsch, J. Woodward, B. Barrell, and J. Parkhill. 2005. Extensive DNA inversions in the *B. fragilis* genome control variable gene expression. Science **307:**1463-1465.

30. **Chambers, C. L., and A. C. Hulse.** 2006. *Salmonella* serovars in the Herpetofauna of Indiana County, Pennsylvania. Appl Environ Microbiol **72:**3771-3773.
31. **Chao, W. L., R. J. Ding, and R. S. Chen.** 1988. Survival of *Yersinia enterocolitica* in the environment. Can J Microbiol **34:**753-756.
32. **Chothia, C., and A. M. Lesk.** 1987. Canonical structures for the hypervariable regions of immunoglobulins. J Mol Biol **196:**901-917.
33. **Chothia, C., A. M. Lesk, A. Tramontano, M. Levitt, S. J. Smith-Gill, G. Air, S. Sheriff, E. A. Padlan, D. Davies, W. R. Tulip, and e. al.** 1989. Conformations of immunoglobulin hypervariable regions. Nature **342:**877-883.
34. **Cotta, M. A., T. R. Whitehead, and R. L. Zeltwanger.** 2003. Isolation, characterization and comparison of bacteria from swine faeces and manure storage pits. Environ Microbiol **5:**737-745.
35. **Daniel, A. S., F. V. Fuller-Pace, D. M. Legge, and N. E. Murray.** 1988. Distribution and diversity of hsd genes in *Escherichia coli* and other enteric bacteria. J Bacteriol **170:**1775-82.
36. **Danso, S. K., and M. Alexander.** 1975. Regulation of predation by prey density: the protozoan-*Rhizobium* relationship. Appl Microbiol **29:**515-21.
37. **Darwin, C.** 1859. On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struglled for Life. John Murray, London.
38. **Davis, M. M., and P. J. Bjorkman.** 1988. T-cell antigen receptor genes and T-cell recognition. Nature **334:**395-402.
39. **De Bolle, X., C. D. Bayliss, D. Field, T. van de Ven, N. J. Saunders, D. W. Hood, and E. R. Moxon.** 2000. The length of a tetranucleotide repeat tract in Haemophilus influenzae determines the phase variation rate of a gene with homology to type III DNA methyltransferases. Mol Microbiol **35:**211-22.
40. **de Moura, H., H. C. Salazar, O. Fernandes, D. C. Lisboa, and F. G. de Carvalho.** 1985. Free-living amoeba in the human intestine. Evidences of parasitism. Rev Inst Med Trop Sao Paulo **27:**150-156.
41. **de Vries, F., A. van Der Ende, J. van Putten, and J. Dankert.** 1996. Invasion of primary nasopharyngeal epithelial cells by *Neisseria meningitidis* is controlled by phase variation of multiple surface antigens. Infect Immun **64:**2998-3006.
42. **Dent, V. E., M. J. Bazin, and P. T. Saunders.** 1976. Behaviour of *Dictyostelium discoideum* amoebae and *Escherichia coli* grown together in chemostat culture. Arch Microbiol **109:**187-94.
43. **Deplancke, B., and H. R. Gaskins.** 2001. Microbial modulation of innate defense: goblet cells and the intestinal mucus layer. Am J Clin Nutr **73:**1131S-1141S.
44. **Donaldson, M., D. Heyneman, R. Dempster, and L. Garcia.** 1975. Epizootic of fatal amebiasis among exhibited snakes: epidemiologic, pathologic, and chemotherapeutic considerations. Am J Vet Res **36:**807-817.
45. **Dorman, C. J., and C. F. Higgins.** 1987. Fimbrial phase variation in *Escherichia coli*: dependence on integration host factor and homologies with other site-specific recombinases. J Bacteriol **169:**3840-3843.

46. **Doulatov, S., A. Hodes, L. Dai, N. Mandhana, M. Liu, R. Deora, R. W. Simons, S. Zimmerly, and J. F. Miller.** 2004. Tropism switching in *Bordetella* bacteriophage defines a family of diversity-generating retroelements. Nature **431:**476-81.
47. **Drickamer, K., and M. E. Taylor.** 2005. Targeting diversity. Nat Struct Mol Biol **12:**830-1.
48. **Dykhuizen, D. E., and L. Green.** 1991. Recombination in *Escherichia coli* and the definition of biological species. J Bacteriol **173:**7257-7268.
49. **Dykova, I., J. Lom, B. Machackova, and H. Peckova.** 1998. *Vexillifera expectata* sp. n. and other non-encysting amoebae isolated from organs of freshwater fish. Folia Parasitol **45:**17-26.
50. **Eckburg, P. B., E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill, K. E. Nelson, and D. A. Relman.** 2005. Diversity of the human intestinal microbial flora. Science **308:**1635-8.
51. **Ehre, C., A. H. Rossi, L. H. Abdullah, K. De Pestel, S. Hill, J. C. Olsen, and C. W. Davis.** 2005. Barrier role of actin filaments in regulated mucin secretion from airway goblet cells. Am J Physiol Cell Physiol **288:**C46-56.
52. **Eichinger, L., and A. A. Noegel.** 2005. Comparative genomics of *Dictyostelium discoideum* and *Entamoeba histolytica*. Curr Opin Microbiol **8:**606-611.
53. **Eichinger, l., J. Pachebat, G. Glöckner, and M.-A. Rajandream.** 2005. The genome of the social amoeba *Dictyostelium discoideum*. Nature **435:**43-57.
54. **Elton, C., and M. Nicholson.** 1942. Fluctuations in numbers of muskrat (*Ondatra zibethica*) in Canada. J Anim Ecol **11:**96-126.
55. **Elton, C., and M. Nicholson.** 1942. The ten-year cycle in the numbers of the lynx in Canada. J Anim Ecol **11:**215-244.
56. **Emonts, M., J. A. Hazelzet, R. de Groot, and P. W. Hermans.** 2003. Host genetic determinants of *Neisseria meningitidis* infections. Lancet Infect Dis **3:**565-577.
57. **Felsenstein, J.** 1975. Genetic drift in clines which are maintained by migration and natural selection. Genetics **81:**191-207.
58. **Fiala, I., and I. Dykova.** 2003. Molecular characterisation of *Neoparamoeba* strains isolated from gills of *Scophthalmus maximus*. Dis Aquat Organ **55:**11-16.
59. **Franke, E. D., and J. S. Mackiewicz.** 1982. Isolation of *Acanthamoeba* and *Naegleria* from the intestinal contents of freshwater fishes and their potential pathogenicity. J Parasitol **68:**164-166.
60. **Fuhrman, J. A., and R. T. Nobel.** 1996. Viruses and protists cause similar bacterial mortality in coastal seawater. Luminol Oceanogr **40:**1236-1242.
61. **Fuller-Pace, F. V., and N. E. Murray.** 1986. Two DNA recognition domains of the specificity polypeptides of a family of type I restriction enzymes. PNAS **83:**9368-9372.
62. **Garcia-Paris, M., D. A. Good, G. Parra-Olea, and D. B. Wake.** 2000. Inaugural article: biodiversity of Costa Rican salamanders: implications of high levels of genetic differentiation and phylogeographic structure for species formation. PNAS **97:**1640-1647.
63. **Gibb, A. P., G. R. Barclay, I. R. Poxton, and F. Di Padova.** 1992. Frequencies of lipopolysaccharide core types among clinical isolates of *Escherichia coli* defined with monoclonal antibodies. J Infect Dis **166:**1051-1057.

64. **Gordon, D. M.** 2001. Geographical structure and host specificity in bacteria and the implications for tracing the source of coliform contamination. Microbiology **147:**1079-1085.
65. **Gordon, D. M., S. Bauer, and J. R. Johnson.** 2002. The genetic structure of *Escherichia coli* populations in primary and secondary habitats. Microbiology **148:**1513-22.
66. **Gordon, D. M., and A. Cowling.** 2003. The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. Microbiology **149:**3575-86.
67. **Gordon, D. M., and F. FitzGibbon.** 1999. The distribution of enteric bacteria from Australian mammals: host and geographical effects. Microbiology **145 (Pt 10):**2663-71.
68. **Gordon, D. M., S. E. Stern, and P. J. Collignon.** 2005. Influence of the age and sex of human hosts on the distribution of *Escherichia coli* ECOR groups and virulence traits. Microbiology **151:**15-23.
69. **Guixa-Boixereu, N., K. Lysnes, and C. Pedros-Alio.** 1999. Viral Lysis and Bacterivory during a Phytoplankton Bloom in a Coastal Water Microcosm. Appl Environ Microbiol **65:**1949-1958.
70. **Haas, R., and T. F. Meyer.** 1987. Molecular principles of antigenic variation in Neisseria gonorrhoeae. Antonie Van Leeuwenhoek **53:**431-4.
71. **Habte, M., and M. Alexander.** 1975. Protozoa as agents responsible for the decline of *Xanthomonas campestris* in soil. Appl Microbiol **29:**159-164.
72. **Hagblom, P., E. Segal, E. Billyard, and M. So.** 1985. Intragenic recombination leads to pilus antigenic variation in Neisseria gonorrhoeae. Nature **315:**156-8.
73. **Hahn, M. W., and M. G. Hofle.** 2001. Grazing of protozoa and its effect on populations of aquatic bacteria. FEMS Microbiol Ecol **35:**113-121.
74. **Hahn, M. W., and M. G. Hofle.** 1998. Grazing Pressure by a Bacterivorous Flagellate Reverses the Relative Abundance of *Comamonas acidovorans* PX54 and *Vibrio* Strain CB5 in Chemostat Cocultures. Appl Environ Microbiol **64:**1910-1918.
75. **Hahn, M. W., E. R. Moore, and M. G. Hofle.** 1999. Bacterial filament formation, a defense mechanism against flagellate grazing, is growth rate controlled in bacteria of different phyla. Appl Environ Microbiol **65:**25-35.
76. **Hartl, D. L., and D. E. Dykhuizen.** 1984. The population genetics of *Escherichia coli*. Annu Rev Genet **18:**31-36.
77. **Hefty, S. P., S. E. Jolliff, M. J. Caimano, S. K. Wikel, and D. R. Akins.** 2002. Changes in temporal and spatial patterns of outer surface lipoprotein expression generate population heterogeneity and antigenic diversity in the lyme disease spirochete, *Borrelia burgdorferi*. Infect Immun **70:**3468-3478.
78. **Heinrichs, D. E., J. A. Yethon, and C. Whitfield.** 1998. Molecular basis for structural diversity in the core regions of the lipopolysaccharides of *Escherichia coli* and *Salmonella enterica*. Mol Microbiol **30:**221-232.
79. **Hiestand-Nauer, R., and S. Iida.** 1983. Sequence of the site-specific recombinase gene cin and of its substrates serving in the inversion of the C segment of bacteriophage P1. EMBO J **2:**1733-1740.
80. **High, N. J., M. P. Jennings, and E. R. Moxon.** 1996. Tandem repeats of the tetramer 5'-CAAT-3' present in lic2A are required for phase variation but not lipopolysaccharide biosynthesis in *Haemophilus influenzae*. Mol Microbiol **20:**165-74.

81. **Hohwy, J., and M. Kilian.** 1995. Clonal diversity of the *Streptococcus mitis* biovar 1 population in the human oral cavity and pharynx. Oral Microbiol Immunol **10:**19-25.
82. **Holst, O., and H. Brade.** 1992. Chemical structure of the core region of lipopolysaccharides., p. 134–170. *In* D. C. a. R. Morrison, J.L. (ed.), In bacterial endotoxic lipopolysaccharides. CRC Press, Boca Raton, FL.
83. **Hood, D. W., M. E. Deadman, M. P. Jennings, M. Bisercic, R. D. Fleischmann, J. C. Venter, and E. R. Moxon.** 1996. DNA repeats identify novel virulence genes in *Haemophilus influenzae*. Proc Natl Acad Sci U S A **93:**11121-5.
84. **Hoogkamp-Korstanje, J. A., J. G. Lindner, J. H. Marcelis, H. den Daas-Slagt, and N. M. de Vos.** 1979. Composition and ecology of the human intestinal flora. Antonie Van Leeuwenhoek **45:**35-40.
85. **Hornak, K., M. Masin, J. Jezbera, Y. Bettarel, J. Nedoma, T. Sime-Ngando, and K. Simek.** 2005. Effects of decreased resource availability, protozoan grazing and viral impact on a structure of bacterioplankton assemblage in a canyon-shaped reservoir. FEMS Microbiol Ecol **52:**315-27.
86. **Hosking, S. L., J. E. Craig, and N. J. High.** 1999. Phase variation of *lic*1A, *lic*2A and *lic*3A in colonization of the nasopharynx, bloodstream and cerebrospinal fluid by *Haemophilus influenzae* type b. Microbiology **145:**3005-3011.
87. **House, D., A. Bishop, C. Parry, G. Dougan, and J. Wain.** 2001. Typhoid fever: pathogenesis and disease. Curr Opin Infect Dis **15:**573-578.
88. **Howard, M. D., A. D. Cox, J. N. Weiser, G. G. Schurig, and T. J. Inzana.** 2000. Antigenic diversity of *Haemophilus somnus* lipooligosaccharide: phase-variable accessibility of the phosphorylcholine epitope. J Clin Microbiol **38:**4412-9.
89. **Huber, H. E., S. Iida, W. Arber, and T. A. Bickle.** 1985. Site-specific DNA inversion is enhanced by a DNA sequence element in cis. PNAS **82:**3776-3780.
90. **Hughes, K. T., P. Youderian, and M. I. Simon.** 1988. Phase variation in *Salmonella*: analysis of Hin recombinase and hix recombination site interaction in vivo. Genes Dev **2:**937-48.
91. **Hughes, M. A., C. W. Lee, C. F. Holm, S. Ghosh, A. Mills, L. A. Lockhart, S. L. Reed, and B. J. Mann.** 2003. Identification of *Entamoeba histolytica* thiol-specific antioxidant as a GalNAc lectin-associated protein. Mol Biochem Parasitol **127:**113-20.
92. **Huppert, A., R. Olinky, and L. Stone.** 2004. Bottom-up excitable models of phytoplankton blooms. Bul Math Biol **66:**865-878.
93. **Huws, S. A., A. J. McBain, and P. Gilbert.** 2005. Protozoan grazing and its impact upon population dynamics in biofilm communities. J Appl Microbiol **98:**238-44.
94. **Iida, S.** 1984. Bacteriophage P1 carries two related sets of genes determining its host range in the invertible C segment of its genome. Virology **134:**421-434.
95. **Iida, S., J. Meyer, and W. Arber.** 1985. Bacteriophage P1 derivatives unaffected in their growth by a large inversion or by IS insertions at various locations. J. Gen. Microbiol. **131:**129-134.
96. **Iida, S., J. Meyer, K. E. Kennedy, and W. Arber.** 1982. A site-specific, conservative recombination system carried by bacteriophage P1. Mapping the recombinase gene cin and the cross-over sites cix for the inversion of the C segment. Embo J. **1:**1445–1453.
97. **Iida, S., J. Meyer, K. E. Kennedy, and W. Arber.** 1982. A site-specific, conservative recombination system carried by bacteriophage P1. Mapping the recombinase gene cin and the cross-over sites cix for the inversion of the C segment. Embo J **1:**1445–1453.

98. **Jennings, M. P., D. W. Hood, I. R. A. Peak, M. Virji, and E. R. Moxon.** 1995. Molecular analysis of a locus for the biosynthesis and phase-variable expression of the lacto-N-neotetraose terminal lipopolysaccharide structure in *Neisseria meningitidis*. Mol Microbiol **18:**729-740.
99. **Jennings, M. P., Y. N. Srikhanta, E. R. Moxon, M. Kramer, J. T. Poolman, B. Kuipers, and P. van der Ley.** 1999. The genetic basis of the phase variation repertoire of lipopolysaccharide immunotypes in *Neisseria meningitidis*. Microbiology **145:**3013-3021.
100. **Jensen, M. A., S. M. Faruque, J. J. Mekalanos, and B. R. Levin.** 2006. Modeling the role of bacteriophage in the control of cholera outbreaks. PNAS **103:**4652-4657.
101. **Jjemba, P. K.** 2001. The interaction of protozoa with their potential prey bacteria in the rhizosphere. J Eukaryot Microbiol **48:**320-4.
102. **Jones, P. A., N. M. Samuels, N. J. Phillips, R. S. Munson, Jr., J. A. Bozue, J. A. Arseneau, W. A. Nichols, A. Zaleski, B. W. Gibson, and M. A. Apicella.** 2002. *Haemophilus influenzae* type b strain A2 has multiple sialyltransferases involved in lipooligosaccharide sialylation. J Biol Chem **277:**14598-14611.
103. **Jonsson, A. B., G. Nyberg, and S. Normark.** 1991. Phase variation of gonococcal pili by frameshift mutation in *pilC*, a novel gene for pilus assembly. Embo J **10:**477-88.
104. **Jonsson, A. B., J. Pfeifer, and S. Normark.** 1992. *Neisseria gonorrhoeae pilC* expression provides a selective mechanism for structural diversity of pili. Proc Natl Acad Sci U S A **89:**3204-3208.
105. **Joo, J., M. Gunny, M. Cases, P. Hudson, R. Albert, and E. Harvill.** 2006. Bacteriophage-mediated competition in *Bordetella bacteria*. Proc Biol Sci **273:**1843-8148.
106. **Jukes, T. H.** 1980. Silent nucleotide substitutions and the molecular evolutionary clock. Science **210:**973-978.
107. **Jurgens, K., and C. Matz.** 2002. Predation as a shaping force for the phenotypic and genotypic composition of planktonic bacteria. Antonie Van Leeuwenhoek **81:**413-434.
108. **Kamp, D., R. Kahmann, D. Zipser, T. R. Broker, and L. T. Chow.** 1978. Inversion of the G DNA segment of phage Mu controls phage infectivity. Nature **271:**577-80.
109. **Kanaar, R., P. Putte, and N. R. Cozzarelli.** 1988. Gin-mediated DNA inversion: product structure and the mechanism of strand exchange. PNAS **85:**752-756.
110. **Killmann, H., and V. Braun.** 1992. An aspartate deletion mutation defines a binding site of the multifunctional FhuA outer membrane receptor of *Escherichia coli* K-12. J Bacteriol **174:**3479-3486.
111. **Kimura, M.** 1979. The neutral theory of molecular evoultion. Sci Am **241:**98-100.
112. **Kimura, M.** 1977. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. Nature **19:**275-276.
113. **Kimura, M., and T. Ohta.** 1978. Stepwise mutation model and distribution of allelic frequencies in a finite population. PNAS **75:**2868-2872.
114. **Kingsley, R. A., and A. J. Baumler.** 2000. Host adaptation and the emergence of infectious disease: the *Salmonella* paradigm. Mol Microbiol **36:**1006-1014.
115. **Klemm, P.** 1986. Two regulatory *fim* genes, *fimB* and *fimE*, control the phase variation of type 1 fimbriae in *Escherichia coli*. EMBO J **5:**1389-1393.

116. **Knirel, Y. A., N. A. Kocharova, O. V. Bystrova, E. Katzenellenbogen, and A. Gamian.** 2002. Structures and serology of the O-specific polysaccharides of bacteria of the genus *Citrobacter*. Arch Immunol Ther Exp (Warsz) **50:**379-91.
117. **Kojimoto, A., K. Uchida, Y. Horii, S. Okumura, R. Yamaguch, and S. Tateyama.** 2001. Amebiasis in four ball pythons, *Python reginus*. J Vet Med Sci **63:**1365-1368.
118. **Kramer, C., U. Henning, and R. Morona.** 1985. Bacteriophage receptor area of outer membrane protein OmpA of *Escherichia coli* K-12. J Bacteriol **164:**539-543.
119. **Krinos, C. M., M. J. Coyne, K. G. Seinacht, A. O. Tzianabos, D. L. Kasper, and L. E. Comstock.** 2001. Extensive surface diversity of a commensal microorganism by multiple DNA inversions. Nature **414:**555-558.
120. **Kussell, E., R. Kishony, N. Q. Balaban, and S. Leibler.** 2005. Bacterial persistence: a model of survival in changing environments. Genetics **169:**1807-1814.
121. **Kutsukake, K., H. Nakashima, A. Tominaga, and T. Abo.** 2006. Two DNA invertases contribute to flagellar phase variation in *Salmonella enterica* Serovar Typhimurium strain LT2. J Bacteriol **188:**950-957.
122. **Kuwahara, T., A. Yamashita, H. Hirakawa, H. Nakayama, H. Toh, N. Okada, S. Kuhara, M. Hattori, T. Hayashi, and Y. Ohnishi.** 2004. Genomic analysis of *Bacteroides fragilis* reveals extensive DNA inversions regulating cell surface adaptation. PNAS **101:**14919-14924.
123. **Lagow, E., M. M. DeSouza, and D. D. Carson.** 1999. Mammalian reproductive tract mucins. Hum Reprod Update **5:**280-92.
124. **Lawrence, J. G.** 2001. Catalyzing bacterial speciation: correlating lateral transfer with genetic headroom. Syst Biol **50:**479-496.
125. **Leber, A. L.** 1999. Intestinal amebae. Clin Lab Med **19:**601-619.
126. **Lee, S. J., L. K. Romana, and P. R. Reeves.** 1992. Cloning and structure of group C1 O antigen (*rfb* gene cluster) from *Salmonella enterica* serovar montevideo. J Gen Microbiol **138:**305-12.
127. **Lee, S. J., L. K. Romana, and P. R. Reeves.** 1992. Sequence and structural analysis of the *rfb* (O-antigen) gene cluster from a group C1 *Salmonella enterica* strain. J Gen Microbiol **138:**1843-55.
128. **Leroy, A., G. De Bruyne, M. Mareel, C. Nokkaew, G. Bailey, and H. Nelis.** 1995. Contact-dependent transfer of the galactose-specific lectin of *Entamoeba histolytica* to the lateral surface of enterocytes in culture. Infect Immun **63:**4253-4260.
129. **Levin, B. R.** 1988. Frequency dependent selection in bacterial populations. Phil Trans Roy Soc London Ser B **319:**459-472.
130. **Lewis, J. L., and B. T. Forschler.** 2004. Protist communities from four castes and three species of *Reticulitermes* (*Isoptera*: *Rhinotermitidae*). Ann Entomol Soc Am **97:**1242-1251.
131. **Lewis, L. A., M. Gipson, K. Hartman, T. Ownbey, J. Vaughn, and D. W. Dyer.** 1999. Phase variation of HpuAB and HmbR, two distinct haemoglobin receptors of *Neisseria meningitidis* DNM2. Mol Microbiol **32:**977-89.
132. **Ley, R. E., D. A. Peterson, and J. I. Gordon.** 2006. Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell **124:**837-48.
133. **Linehan, S. A., and D. W. Holden.** 2003. The interplay between *Salmonella typhimurium* and its macrophage host--what can it teach us about innate immunity? Immunol Lett **85:**183-192.

134. **Links, G. A.-C. h. t. r., T. d. o. a. a. t. h. l. i. n. p. o. E. coli., , V. A. Barcus, A. J. Titheradge, and N. E. Murray.** 1995. The diversity of alleles at the hsd locus in natural populations of Escherichia coli. Genetics **140:**1187-97.
135. **Links, I. D. C. N. A. S.-., T. s. o. S. infection., M. B. Goldberg, and R. H. Rubin.** 1988. The spectrum of *Salmonella* infection. Infect Dis Clin North Am **2:**571-598.
136. **Links, M. M. J.-., C. o. m. w. t. u. v. p. s. o. t. I. r. a. m. enzymes., , N. E. Murray, A. S. Daniel, G. M. Cowan, and P. M. Sharp.** 1993. Conservation of motifs within the unusually variable polypeptide sequences of type I restriction and modification enzymes. Mol Microbiol **9:**133-143.
137. **Liu, D., A. M. Haase, L. Lindqvist, A. A. Lindberg, and P. R. Reeves.** 1993. Glycosyl transferases of O-antigen biosynthesis in *Salmonella enterica*: identification and characterization of transferase genes of groups B, C2, and E1. J Bacteriol **175:**3408-13.
138. **Liu, D., L. Lindqvist, and P. R. Reeves.** 1995. Transferases of O-antigen biosynthesis in *Salmonella enterica*: dideoxyhexosyltransferases of groups B and C2 and acetyltransferase of group C2. J Bacteriol **177:**4084-8.
139. **Liu, D., N. K. Verma, L. K. Romana, and P. R. Reeves.** 1991. Relationships among the *rfb* regions of *Salmonella* serovars A, B, and D. J Bacteriol **173:**4814-9.
140. **Liu, M., R. Deora, S. R. Doulatov, M. Gingery, F. A. Eiserling, A. Preston, D. J. Maskell, R. W. Simons, P. A. Cotter, J. Parkhill, and J. F. Miller.** 2002. Reverse transcriptase-mediated tropism switching in Bordetella bacteriophage. Science **295:**2091-4.
141. **Loftus, B., I. Anderson, R. Davies, U. C. Alsmark, J. Samuelson, and *e. al.*** 2005. The genome of the protist parasite *Entamoeba histolytica*. Nature **433(7028):865-8.:**865-868.
142. **Lunzer, M., A. Natarajan, D. E. Dykhuizen, and A. M. Dean.** 2002. Enzyme kinetics, substitutable resources and competition: from biochemistry to frequency-dependent selection in *lac*. Genetics **162:**485-499.
143. **Macnab, R. M., and S.-I. Aizawa.** 1984. Bacterial motility and the bacterial flagellar motor. Annu Rev Biophys Bioeng **13:**51-83.
144. **Macpherson, A. J., and T. Uhr.** 2004. Induction of protective IgA by intestinal dendritic cells carrying commensal bacteria. Science **303:**1662-1665.
145. **Makino, S., J. P. van Putten, and T. F. Meyer.** 1991. Phase variation of the opacity outer membrane protein controls invasion by Neisseria gonorrhoeae into human epithelial cells. Embo J **10:**1307-15.
146. **Martinez-Diaz, R. A., S. Herrera, A. Castro, and F. Ponce.** 2000. *Entamoeba* sp. (*Sarcomastigophora*: *Endamoebidae*) from ostriches (*Struthio camelus*) (*Aves*: *Struthionidae*). Vet Parasitol **92:**173-179.
147. **Maskell, D. J., M. J. Szabo, P. D. Butler, A. E. Williams, and E. R. Moxon.** 1991. Phase variation of lipopolysaccharide in *Haemophilus influenzae*. Res Microbiol **142:**719-24.
148. **Matz, C., P. Deines, J. Boenigk, H. Arndt, L. Eberl, S. Kjelleberg, and K. Jurgens.** 2004. Impact of violacein-producing bacteria on survival and feeding of bacterivorous nanoflagellates. Appl Environ Microbiol **70:**1593-1595.
149. **Matz, C., and K. Jurgens.** 2005. High motility reduces grazing mortality of planktonic bacteria. Appl Environ Microbiol **71:**921-929.
150. **Matz, C., and S. Kjelleberg.** 2005. Off the hook--how bacteria survive protozoan grazing. Trends Microbiol **13:**302-307.

151. **McClain, M. S., I. C. Blomfield, and B. I. Eisenstein.** 1991. Roles of *fimB* and *fimE* in site-specific DNA inversion associated with phase variation of type 1 fimbriae in *Escherichia coli*. J Bacteriol **173:**5308-14.
152. **McMahon, S. A., J. L. Miller, J. A. Lawton, D. E. Kerkow, A. Hodes, M. A. Marti-Renom, S. Doulatov, E. Narayanan, A. Sali, J. F. Miller, and P. Ghosh.** 2005. The C-type lectin fold as an evolutionary solution for massive sequence variation. Nat Struct Mol Biol **12:**886-92.
153. **Mead, P. S., L. Slutsker, V. Dietz, L. F. McCaig, J. S. Bresee, C. Shapiro, P. M. Griffin, and R. V. Tauxe.** 1999. Food-related illnesses and death in the United States. Emerg. Infect. Dis. **5:**607-625.
154. **Menge, B. A., B. A. Daley, J. Lubchenco, E. Sanford, E. Dahlhoff, P. M. Halpin, G. Hudson, and J. L. Burnaford.** 1999. Top-down and bottom-up regualtions of New Zealand rocky intertidal communities. Ecological Monographs **69:**297-330.
155. **Meyer, T. F.** 1991. Evasion mechanisms of pathogenic *Neisseriae*. Behring Inst Mitt**:**194-9.
156. **Meyer, T. F., and J. P. van Putten.** 1989. Genetic mechanisms and biological implications of phase variation in pathogenic neisseriae. Clin Microbiol Rev **2 Suppl:**S139-45.
157. **Milkman, R., E. Jaeger, and R. D. McBride.** 2002. Molecular evolution of the *Escherichia coli* chromosome. VI. Two regions of high effective recombination. Genetics **163:**475-483.
158. **Milkman, R., E. Jaeger, and R. D. McBride.** 2003. Molecular evolution of the *Escherichia coli* chromosome. VI. Two regions of high effective recombination. Genetics **162:**475-483.
159. **Mizuochi, T., T. Matthews, M. Kato, J. Hamako, K. Titani, J. Solomon, and T. Feizi.** 1990. Diversity of oligosaccharide structures on the envelope glycoprotein gp 120 of human immunodeficiency virus 1 from the lymphoblastoid cell line H9. Presence of complex-type oligosaccharides with bisecting N- acetylglucosamine residues. J Biol Chem **265:**8519-8524.
160. **Moran, M. D., T. P. Rooney, and L. E. Hurd.** 1996. Top-down cascade from a bitrophic predator in an old-field community. Ecology **77:**2219-2227.
161. **Morona, R., M. Klose, and U. Henning.** 1984. *Escherichia coli* K-12 outer membrane protein (OmpA) as a bacteriophage receptor: analysis of mutant genes expressing altered proteins
. 1984 Aug;159(2):570-8. J Bacteriol **159:**570-580.
162. **Moxon, E. R.** 1992. Molecular basis of invasive *Haemophilus influenzae* type b disease. J Infect Dis **165:**Suppl 1:S77-81.
163. **Munson, R. S., Jr.** 1990. *Haemophilus influenzae*: surface antigens and aspects of virulence. Can J Vet Res **54 Suppl:**S63-7.
164. **Murray, N. E.** 2000. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). Microbiol Mol Biol Rev **64:**412-434.
165. **Murray, N. E., A. S. Daniel, G. M. Cowan, and P. M. Sharp.** 1993. Conservation of motifs within the unusually variable polypeptide sequences of type I restriction and modification enzymes. Mol Microbiol **9:**133-43.
166. **Murray, N. E., J. A. Gough, B. Suri, and T. A. Bickle.** 1982. Structural homologies among type I restriction-modification systems. EMBO J **1:**535-539.

167. **Nelson, K., and R. Selander.** 1994. Intergeneric transfer and recombination of the 6-phosphogluconate dehydrogenase gene (*gnd*) in enteric bacteria. PNAS **91:**10227-10231.
168. **Nelson, K., T. S. Whittam, and R. K. Selander.** 1991. Nucleotide polymorphism and evolution in the glyceraldehyde-3-phosphate dehydrogenase gene (*gapA*) in natural populations of *Salmonella* and *Escherichia coli*. PNAS **88:**6667-6671.
169. **Newbold, C. J., K. Ushida, B. Morvan, G. Fonty, and J. P. Jouany.** 1996. The role of ciliate protozoa in the lysis of methanogenic archaea in rumen fluid. Lett Appl Microbiol **23:**421-5.
170. **Niess, J. H., and H. C. Reinecker.** 2006. Dendritic cells in the recognition of intestinal microbiota. Cellular Microbiol **8:**558-564.
171. **Ohl, M. E., and S. I. Miller.** 2001. *Salmonella*: a model for bacterial pathogenesis. Annu Rev Med **52:**259-274.
172. **Ohta, T., and M. Kimura.** 1969. Linkage disequilibrium at steady state determined by random genetic drift and recurrent mutation. Genetics **63:**229-38.
173. **Okada, S., and D. M. Gordon.** 2001. Host and geographical factors influence the thermal niche of enteric bacteria isolated from native Australian mammals. Mol Ecol **10:**2499-2513.
174. **Palmer, C., E. M. Bik, M. B. Eisen, P. B. Eckburg, T. R. Sana, P. K. Wolber, D. A. Relman, and P. O. Brown.** 2006. Rapid quantitative profiling of complex microbial populations. Nucleic Acids Res **34:**e5.
175. **Parkhill, J., M. Sebaihia, A. Preston, L. D. Murphy, N. Thomson, D. E. Harris, and *e. al.*** 2003. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. Nat Genet **35:**32-40.
176. **Pathan, N., S. N. Faust, and M. Levin.** 2003. Pathophysiology of meningococcal meningitis and septicaemia. Arch Dis Child **88:**601-607.
177. **Patrick, S., J. Parkhill, L. J. McCoy, N. Lennard, M. J. Larkin, M. Collins, M. Sczaniecka, and G. Blakely.** 2003. Multiple inverted DNA repeats of *Bacteroides fragilis* that control polysaccharide antigenic variation are similar to the hin region inverted repeats of *Salmonella typhimurium*. Microbiology **149:**915-924.
178. **Peduzzi, P., and F. Schiemer.** 2004. Bacteria and viruses in the water column of tropical freshwater reservoirs. Environ Microbiol **6:**707-15.
179. **Peng, T. X., A. Moya, and F. J. Ayala.** 1991. Two modes of balancing selection in *Drosophila melanogaster*: Overcompensation and overdominance. Genetics **128:**381-391.
180. **Perch, M., R. Bishop, B. Plikaytis, P. Fields, C. R. Braden, and R. V. Tauxe.** 2003. *Salmonella* surveillance: annual summary. CDC.
181. **Pernthaler, J.** 2005. Predation on prokaryotes in the water column and its ecological implications. Nat Rev Microbiol **3:**537-546.
182. **Petit, C.** 1952. A case of high mutability in *Drosophila melanogaster*. C R Hebd Seances Acad Sci **235:**745-747.
183. **Petit, C.** 1959. Nature of the stimuli responsible for sexual selection in *Drosophila melanogaster*. C R Hebd Seances Acad Sci **248:**3484-3485.
184. **Petit, C.** 1951. Some factors responsible for the advantages of the rare type in *Drosophila melanogaster*. C R Hebd Seances Acad Sci **232:**2482-2482.

185. **Petri, W. A., Jr., R. Haque, and B. J. Mann.** 2002. The bittersweet interface of parasite and host: lectin-carbohydrate interactions during human invasion by the parasite *Entamoeba histolytica*. Annu Rev Microbiol **56:**39-64.
186. **Pillai, D. R., P. S. K. Wan, Y. C. W. Yau, J. I. Ravdin, and K. C. Kain.** 1999. The cysteine-rich region of the *Entamoeba histolytica* adherence lectin (170-kilodalton subunit) is sufficient for high-affinity Gal/GalNAc-specific binding *in vitro*. Infect Immun **67:**3836-3841.
187. **Plasterk, R. H. A., R. Kanaar, and P. Putte.** 1984. A genetic switch in vitro: DNA inversion by Gin protein of phage Mu. PNAS **81:**2689-2692.
188. **Popoff, M. Y.** 2001. Antigenic Formulas of the *Salmonella* Serovars, 8th edition. Institut Pasteur, Paris.
189. **Pozio, E.** 2003. Foodborne and waterborne parasites. Acta Microbiol Pol **52:**83-96.
190. **Pradeep Ram, A. S., D. Boucher, T. Sime-Ngando, D. Debroas, and J. C. Romagoux.** 2005. Phage bacteriolysis, protistan bacterivory potential, and bacterial production in a freshwater reservoir: coupling with temperature. Microb Ecol **50:**64-72.
191. **Rabsch, W., H. L. Andrews, R. A. Kingsley, R. Prager, H. Tschape, L. G. Adams, and A. J. Baumler.** 2002. *Salmonella enterica* serotype Typhimurium and its host-adapted variants. Infect Immun **70:**2249-55.
192. **Rainey, P. B., and M. Travisano.** 1998. Adaptive radiation in a heterogeneous environment. Nature **394:**69-72.
193. **Ravdin, J. I.** 1989. *Entamoeba histolytica*: from adherence to enteropathy. J Infect Dis **159:**420-429.
194. **Reeves, P.** 1993. Evolution of *Salmonella* O-antigen variation by interspecific gene transfer on a large scale. Trends Genet **9:**17-22.
195. **Reeves, P.** 1995. Role of o-antigen variation in the immune response. Trends Microbiol **3:**381-386.
196. **Reeves, P. R.** 1995. Role of O-antigen variation in the immune response. Trends Microbiol **3:**381-386.
197. **Reynolds, H. Y.** 2002. Modulating airway defenses against microbes. Curr Opin Pulm Med **8:**154-165.
198. **Ribeiro-Rosa, E. A., R. Takaki-Rosa, C. V. Pereira, M. F. Gomes-Boriollo, and J. F. Hofling.** 2001. Inter and intra-specific genetic variability of oral Candida species. Rev Iberoam Micol **18:**60-64.
199. **Rich, S. M., R. R. Hudson, and F. J. Ayala.** 1997. *Plasmodium falciparum* antigenic diversity: Evidence of clonal population structure. PNAS **94:**13040-13045.
200. **Richardson, A. R., and S. I.** 2001. Mismatch repair and the regulation of phase variation in *Neisseria meningitidis*. Mol Microbiol **40:**645-655.
201. **Robbe, C., C. Capon, B. Coddeville, and J. C. Michalski.** 2004. Structural diversity and specific distribution of O-glycans in normal human mucins along the intestinal tract. Biochem J **384:**307-16.
202. **Robbe, C., C. Capon, C. Flahaut, and J. C. Michalski.** 2003. Microscale analysis of mucin-type O-glycans by a coordinated fluorophore-assisted carbohydrate electrophoresis and mass spectrometry approach. Electrophoresis **24:**611-21.
203. **Robbe, C., C. Capon, E. Maes, M. Rousset, A. Zweibaum, J. P. Zanetta, and J. C. Michalski.** 2003. Evidence of regio-specific glycosylation in human intestinal mucins: presence of an acidic gradient along the intestinal tract. J Biol Chem **278:**46337-48.

204. **Roche, R. J., N. J. High, and E. R. Moxon.** 1994. Phase variation of *Haemophilus influenza* lipopolysaccharide: characterization of lipopolysaccharide from individual colonies. FEMS Microbiol Lett **120:**279-284.
205. **Rodriguez-Zaragoza, S.** 1994. Ecology of free-living amoebae. Crit Rev Microbiol **20:**225-41.
206. **Ronn, R., A. E. McCaig, B. S. Griffiths, and J. I. Prosser.** 2002. Impact of protozoan grazing on bacterial community structure in soil microcosms. Appl Environ Microbiol **68:**6094-105.
207. **Rosenheim, J. A.** 1998. Higher-order predators and the rgulation of insect herbivore populations. Annu Rev Entomol **43:**421-427.
208. **Rubin, L. G., and E. R. Moxon.** 1983. Pathogenesis of bloodstream invasion with *Haemophilus influenzae* type b. Infect Immun **41:**280-284.
209. **Rubin, L. G., A. Zwahlen, and E. R. Moxon.** 1985. Role of intravascular replication in the pathogenesis of experimental bacteremia due to *Haemophilus influenzae* type b. J Infect Dis **152:**307-314.
210. **Rytkonen, A., B. Albiger, P. Hansson-Palo, H. Kallstrom, P. Olcen, H. Fredlund, and A. B. Jonsson.** 2004. *Neisseria meningitidis* undergoes PilC phase variation and PilE sequence variation during invasive disease. J Infect Dis **189:**402-9.
211. **Samuel, G., and P. Reeves.** 2003. Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. Carbohydr Res **338:**2503-2519.
212. **Sandmeier, H., S. Iida, and W. Arber.** 1992. DNA inversion regions Min of plasmid p15B and Cin of bacteriophage P1: evolution of bacteriophage tail fiber genes. J Bacteriol **174:**3936–3944.
213. **Sandmeier, H., S. Iida, and W. Arber.** 1992. DNA inversion regions Min of plasmid p15B and Cin of bacteriophage P1: evolution of bacteriophage tail fiber genes. J. Bacteriol. **174:**3936–3944.
214. **Schneider, H., H. Fsihi, B. Kottwitz, B. Mygind, and B. E.** 1993. Identification of a segment of the *Escherichia coli* Tsx protein that functions as a bacteriophage receptor area. J Bacteriol **175:**2809-2817.
215. **Schuster, F. L., J. F. De Jonckheere, H. Moura, R. Sriram, M. M. Garner, and G. S. Visvesvara.** 2003. Isolation of a thermotolerant *Paravahlkampfia* sp. from lizard intestine: biology and molecular identification. J Euk Microbiol **50:**373-378.
216. **Segal, E., E. Billyard, M. So, S. Storzbach, and T. F. Meyer.** 1985. Role of chromosomal rearrangement in *N. gonorrhoeae* pilus phase variation. Cell **40:**293-300.
217. **Segal, E., P. Hagblom, H. S. Seifert, and M. So.** 1986. Antigenic variation of gonococcal pilus involves assembly of separated silent gene segments. PNAS **83:**2177-81.
218. **Selander, R. K., P. Beltran, N. H. Smith, R. Helmuth, F. A. Rubin, D. J. Kopecko, K. Ferris, B. D. Tall, C. A., and J. M. u.-. Musser.** 1990. Evolutionary genetic relationships of clones of *Salmonella* serovars that cause human typhoid and other enteric fevers. Infect Immun **58:**2262-2275.
219. **Serkin, C. D., and H. S. Seifert.** 1998. Frequency of pilin antigenic variation in *Neisseria gonorrhoeae*. J Bacteriol **180:**1955-8.
220. **Sesma, M. J., and L. Z. Ramos.** 1989. Isolation of free living amoebas from the intestinal content of reptiles. J Parasitol **75:**322-324.

221. **Shafer, W. M., A. Datta, V. S. Kolli, M. M. Rahman, J. T. Balthazar, L. E. Martin, W. L. Veal, D. S. Stephens, and R. Carlson.** 2002. Phase variable changes in genes *lgtA* and *lgtC* within the *lgtABCDE* operon of *Neisseria gonorrhoeae* can modulate gonococcal susceptibility to normal human serum. J Endotoxin Res **8:**47-58.
222. **Sherr, E. B., and B. F. Sherr.** 2002. Significance of predation by protists in aquatic microbial food webs. Antonie Van Leeuwenhoek **81:**293-308.
223. **Shirazi, T., R. J. Longman, A. P. Corfield, and C. S. Probert.** 2000. Mucins and inflammatory bowel disease. Postgrad Med J **76:**473-8.
224. **Snyder, L. A., S. A. Butcher, and N. J. Saunders.** 2001. Comparative whole-genome analyses reveal over 100 putative phase-variable genes in the pathogenic *Neisseria* spp. Microbiology **147 (Pt 8):**2321-2332.
225. **Som, A., and B. N. Singh.** 2004. Rare male mating advantage for inversion karyotype in *Drosophila ananassae*. Behav Genet **34:**335-342.
226. **Sparling, P. F., J. G. Cannon, and M. So.** 1986. Phase and antigenic variation of pili and outer membrane protein II of *Neisseria gonorrhoeae*. J Infect Dis **153:**196-201.
227. **Spiers, A. J., J. Bohannon, S. M. Gehrig, and P. B. Rainey.** 2003. Biofilm formation at the air-liquid interface by the *Pseudomonas fluorescens* SBW25 wrinkly spreader requires an acetylated form of cellulose. Mol Microbiol **50:**15-27.
228. **Spiers, A. J., A. Buckling, and P. B. Rainey.** 2000. The causes of *Pseudomonas* diversity. Microbiology **146:**2345-2350.
229. **Spiers, A. J., S. G. Kahn, J. Bohannon, M. Travisano, and P. B. Rainey.** 2002. Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. I. Genetic and phenotypic bases of wrinkly spreader fitness. Genetics **161:**33-46.
230. **Spiers, A. J., and P. B. Rainey.** 2005. The *Pseudomonas fluorescens* SBW25 wrinkly spreader biofilm requires attachment factor, cellulose fibre and LPS interactions to maintain strength and integrity. Microbiology **151:**2829-39.
231. **Sprenger, G. A.** 1995. Genetics of pentose-phosphate pathway enzymes of *Escherichia coli* K-12. Arch Microbiol **164:**324-330.
232. **Srikhanta, Y. N., T. L. Maguire, K. J. Stacey, S. M. Grimmond, and M. P. Jennings.** 2005. The phasevarion: a genetic system controlling coordinated, random switching of expression of multiple genes. Proc Natl Acad Sci U S A **102:**5547-51.
233. **Staaf, M., F. Urbina, A. Weintraub, and G. Widmalm.** 1999. Structural elucidation of the O-antigenic polysaccharides from *Escherichia coli* O21 and the enteroaggregative *Escherichia coli* strain 105. Eur J Biochem **266:**241-245.
234. **Stanley, C. M., and T. E. Phillips.** 1999. Selective secretion and replenishment of discrete mucin glycoforms from intestinal goblet cells. Am J Physiol Gastrointest Liver Physiol **277:**G191-200.
235. **Stephens, D. S.** 1985. *Neisseria meningitidis*. Infect Control **6:**37-40.
236. **Stern, A., M. Brown, P. Nickel, and T. F. Meyer.** 1986. Opacity genes in *Neisseria gonorrhoeae*: control of phase and antigenic variation. Cell **47:**61-71.
237. **Stevenson, G., B. Neal, D. Liu, M. Hobbs, N. H. Packer, M. Batley, J. W. Redmond, L. Lindquist, and P. Reeves.** 1994. Structure of the O-antigen of *Escherichia coli* K-12 and the sequence of its *rfb* gene cluster. J Bacteriol **176:**4144-4156.
238. **Storz, J. F., and J. M. Dubach.** 2004. Natural Selection drives altitudinal divergence at the albumin locius in deer mice. Evolution **58:**1342–1352.
239. **Suttle, C. A.** 2005. Viruses in the sea. Nature **437:**356-61.

240. **Theisen, M., M. Borre, M. J. Mathiesen, B. Mikkelsen, A. M. Lebech, and H. K.** 1995. Evolution of the *Borrelia burgdorferi* outer surface protein OspC. J Bacteriol **177:**3036-3044.
241. **Thomson, N., M. Sebaihia, A. Cerdeno-Tarraga, S. Bentley, L. Crossman, and J. Parkhill.** 2003. The value of comparison. Nat Rev Microbiol **1:**11-12.
242. **Thomson, N. R., C. Yeats, K. Bell, M. T. Holden, S. D. Bentley, M. Livingstone, A. M. Cerdeno-Tarraga, B. Harris, J. Doggett, D. Ormond, K. Mungall, K. Clarke, T. Feltwell, Z. Hance, M. Sanders, M. A. Quail, C. Price, B. G. Barrell, J. Parkhill, and D. Longbottom.** 2005. The *Chlamydophila abortus* genome sequence reveals an array of variable proteins that contribute to interspecies variation. Genome Res **15:**629-640.
243. **Tlaskalova-Hogenova, H., R. Stepankova, T. Hudcovic, L. Tuckova, B. Cukrowska, R. Lodinova-Zadnikova, H. Kozakova, P. Rossmann, J. Bartova, and D. Sokol.** 2004. Commensal bacteria (normal microflora), mucosal immunity and chronic inflammatory and autoimmune diseases. Immunol Letters **93:**97-108.
244. **Toscaronicacute, M., and F. J. Ayala.** 1981. Density and frequency dependent selection at the Mdh-2 locus in *Drosophila pseudoobscura*. Genetics **97:**679-701.
245. **Tracey, M. L., and F. J. Ayala.** 1974. Genetic load in natural populations: is it compatible with the hypothesis that many polymorphism are maintained by natural selection? Genetics **77:**569-589.
246. **Uhl, M. A., and J. F. Miller.** 1996. Central role of the BvgS receiver as a phosphorylated intermediate in a complex two-component phosphorelay. J Biol Chem **271:**33176-33180.
247. **Uhl, M. A., and J. F. Miller.** 1996. Integration of multiple domains in a two-component sensor protein: the *Bordetella pertussis* BvgAS phosphorelay. EMBO J **15:**1028-1036.
248. **Vakharia, H., and M. R.** 1996. A genetic approach for analysing surface-exposed regions of the OmpC protein of *Escherichia coli* K-12. Mol Microbiol **19:**881-889.
249. **Vakkari, P., A. Blom, M. Rusanen, J. Raisio, and H. Toivonen.** 2006. Genetic variability of fragmented stands of pedunculate oak (*Quercus robur*) in Finland. Genetica **127:**123-241.
250. **van Deuren, M., P. Brandtzaeg, and J. W. van der Meer.** 2000. Update on meningococcal disease with emphasis on pathogenesis and clinical management. Clin Microbiol Rev **13:**144-166.
251. **van Duijkeren, E., W. J. Wannet, D. J. Houwers, and W. van Pelt.** 2002. Serotype and phage type distribution of *Salmonella* strains isolated from humans, cattle, pigs, and chickens in the Netherlands from 1984 to 2001. J Clin Microbiol **40:**3980-3985.
252. **van Ham, S. M., L. van Alphen, F. R. Mooi, and J. P. van Putten.** 1993. Phase variation of *H. influenzae* fimbriae: transcriptional control of two divergent genes through a variable combined promoter region. Cell **73:**1187-1196.
253. **Van Klinken, B. J., J. Dekker, H. A. Buller, and A. W. A. W. Einerhand.** 1995. Mucin gene structure and expression: protection vs. adhesion. Am J Physiol Gastrointest Liver Physiol **269:**613-627.
254. **van Putten, J. P.** 1993. Phase variation of lipopolysaccharide directs interconversion of invasive and immuno-resistant phenotypes of *Neisseria gonorrhoeae*. EMBO J **12:**4043-4051.

255. **Verma, N., and P. Reeves.** 1989. Identification and sequence of *rfbS* and *rfbE*, which confers antigenic specificity on group A and group D Salmonellae. J Bacteriol **171:**5694-5701.
256. **Wang, L., L. K. Romana, and P. R. Reeves.** 1992. Molecular analysis of a *Salmonella enterica* group E1 *rfb* gene cluster: O-antigen and the genetic basis of the major polymorphism. Genetics **130:**429-43.
257. **Wang, L., D. Rothemund, H. Curd, and P. R. Reeves.** 2003. Species-wide variation in the *Escherichia coli* flagellin (H-antigen) gene. J Bacteriol **185:**2936-2943.
258. **Wasson, K., and R. L. Peper.** 2000. Mammalian microsporidiosis. Vet Pathol **37:**113-128.
259. **Weinacht, K. G., H. Roche, C. M. Krinos, M. J. Coyne, J. Parkhill, and L. E. Comstock.** 2004. Tyrosine site-specific recombinases mediate DNA inversions affecting the expression of outer surface proteins of *Bacteroides fragilis*. Mol Microbiol **53:**1319-1330.
260. **Weinbauer, M. G.** 2004. Ecology of prokaryotic viruses. FEMS Microbiol Rev **28:**127-181.
261. **Weiser, J. N., M. Shchepetov, and S. T. Chong.** 1997. Decoration of lipopolysaccharide with phosphorylcholine: a phase-variable characteristic of *Haemophilus influenzae*. Infect Immun **65:**943-50.
262. **Weiser, J. N., A. Williams, and E. R. Moxon.** 1990. Phase-variable lipopolysaccharide structures enhance the invasive capacity of *Haemophilus influenzae*. Infect Immun **58:**3455-3457.
263. **Weitere, M., T. Bergfeld, S. A. Rice, C. Matz, and S. Kjelleberg.** 2005. Grazing resistance of *Pseudomonas aeruginosa* biofilms depends on type of protective mechanism, developmental stage and protozoan feeding mode. Environ Microbiol **7:**1593-1601.
264. **Wheeler, A. L., P. G. Hartel, D. G. Godfrey, J. L. Hill, and W. I. Segars.** 2002. Potential of *Enterococcus faecalis* as a human fecal indicator for microbial source tracking. J Environ Qual **31:**1286-1293.
265. **Whitfield, C., and I. S. Roberts.** 1999. Structure, assembly and regulation of expression of capsules in *Escherichia coli*. Mol Microbiol **31:**1307-1319.
266. **Whitman, W. B., D. C. Coleman, and W. J. Wiebe.** 1998. Prokaryotes: the unseen majority. PNAS **95:**6578-5583.
267. **Wieltschnig, C., U. R. Fischer, A. K. Kirschner, and B. Velimirov.** 2003. Benthic bacterial production and protozoan predation in a silty freshwater environment. Microb Ecol **46:**62-72.
268. **Wildschutte, H., D. M. Wolfe, A. Tamewitz, and J. G. Lawrence.** 2004. Protozoan predation, diversifying selection, and the evolution of antigenic diversity in *Salmonella*. PNAS **101:**10644-10649.
269. **Worm, B., and R. A. Myers.** 2003. Meta-analysis of cod-shrimp interactions reveals top-down control in ocean food webs. Ecology **84:**162-173.
270. **Xiang, S. H., A. M. Haase, and P. R. Reeves.** 1993. Variation of the *rfb* gene clusters in *Salmonella enterica*. J Bacteriol **175:**4877-4884.
271. **Xiang, S. H., M. Hobbs, and P. R. Reeves.** 1994. Molecular analysis of the *rfb* gene cluster of a group D2 *Salmonella enterica* strain: evidence for its origin from an insertion

sequence-mediated recombination event between group E and D1 strains. J Bacteriol **176:**4357-4365.

272. **Xinyao, L., S. Miao, L. Yonghong, G. Yin, Z. Zhongkai, W. Donghui, W. Weizhong, and A. Chencai.** 2006. Feeding characteristics of an amoeba (*Lobosea*: *Naegleria*) grazing upon *Cyanobacteria*: food selection, ingestion and digestion progress. Microb Ecol.
273. **Ylioja, T., H. Roininen, M. P. Ayres, M. Rousi, and P. W. Price.** 1999. Host-driven population dynamics in an herbivrous insect. PNAS **96:**10735-10740.
274. **Yung, A. P., and M. I. McDonald.** 2003. Early clinical clues to meningococcaemia. Med J Aust **178:**134-137.
275. **Zhu, P., M. J. Klutch, M. C. Bash, R. S. W. Tsang, L. K. Ng, and C. M. Tsai.** 2002. Genetic diversity of three *lgt* loci for biosynthesis of lipooligosaccharide (LOS) in *Neisseria* species. Microbiology **148:**1833-1844.
276. **Zoetendal, E. G., E. E. Vaughan, and W. M. de Vos.** 2006. A microbial world within us. Mol Microbiol **59:**1639-50.
277. **Zubkov, M. V., and M. A. Sleigh.** 1999. Growth of amoebae and flagellates on bacteria deposited on filters. Microb Ecol **37:**107-115.